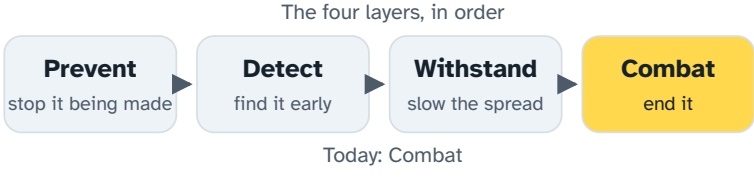
Day 5: Exiting the Pandemic (Combat)

Friday 30 October · discussion 3:00 to 4:50 pm New York

THE DAY MAP

How do you get vaccines and medicines made, tested and delivered fast enough to matter?



Your deep-dive: page 5.4. Why trials take so long, and what the RECOVERY trial did differently. Its readings come first below; the other options follow.

18 readings here, about 244 minutes of reading. Discussion Friday 30 October, 3:00 to 4:50 pm New York, with Conor Spence.

5.1-1	What is a Vaccine? original paper →	3 + 5 min
5.1-2	Vaccines 101: How new vaccines are developed	3 + 6 min

	original paper →	
5.1-3	The Quest for Universal Flu Vaccines original paper →	3 + 12 min
5.1-4	Day Zero Antivirals for Future Pandemics original paper →	3 + 11 min
5.1-5	Accelerating Vaccine Development: The 100 Days Mission original paper →	3 + 14 min
5.1-6	The clinical trial bottlenecks game original paper →	3 + 10 min
5.4-1	From Warp Speed to 100 Days original paper →	3 + 13 min
5.4-2	How the UK found the first effective Covid-19 treatment — and saved a million lives original paper →	3 + 13 min
5.4-3	RECOVERY Trial Wikipedia original paper →	2 + 7 min
5.3-1	How To Reuse the Operation Warp Speed Model original paper →	3 + 9 min
5.3-2	Progress Deferred: Lessons From mRNA Vaccine Development original paper →	2 + 12 min

5.5-1	Promoting versatile vaccine development for emerging pandemics original paper →	3 + 13 min
5.5-2	Calculating the costs and benefits of advance preparations for future pandemics original paper →	3 + 0 min
5.5-3	Expanding capacity for vaccines against Covid-19 and future pandemics: a review of economic issues original paper →	3 + 19 min
5.5-4	COVAX explained original paper →	3 + 9 min
5.6-1	Essential Medicines Supply Chain and Manufacturing Resilience Assessment original paper →	3 + 17 min
5.6-2	COVAX explained original paper →	2 + 7 min
5.6-3	Testing fractional doses of COVID-19 vaccines original paper →	2 + 17 min

WHAT THE COURSE ASKS YOU TO DO

5.2 Rapid MCM development deep-dive

Pick 1 of the following 4 topics to dive deeper into. We've given you some resources on the following pages as a

starting point.

- **Option 1: Faster R&D**
- **Option 2: Clinical trials innovation**
- **Option 3: Better financing and procurement**
- **Option 4: Faster manufacturing and delivery**

You will bring the results of your research to share in your group discussion.

THE READINGS

5.1-1 · CORE · COURSE TIME 7 MIN

What is a Vaccine?

Nucleus Medical Media · 2021

Read the original paper → · the course assigns: the video's own caption transcript (saved subtitle file)

> **A vaccine shows your immune system a safe version of a germ's identifying marker so that memory cells stay behind, ready to destroy the real germ before it can make you sick.**

Average summary

Summarized from the video's caption transcript, not from its pictures. The captions are the video's own written track, so the wording is clean; one term is misspelled in the file and is flagged below.

The video defines **immunization** as the process of becoming immune to, or protected against, a disease, usually by receiving a vaccine. It then walks through how immunity works at all. Immune cells circulate constantly and read substances on the surfaces of cells called antigens. Healthy cells carry different antigens from diseased cells or foreign invaders, so immune cells ignore the healthy ones. When an immune cell meets a germ's antigen, it destroys the germ and then displays that antigen on its own surface, which switches on other immune cells. Some of those recognize the antigen on infected body cells and kill them. Others, plasma cells, make antibodies that lock onto only the germs carrying that specific antigen and mark them for destruction.

Once the infection clears, some of the exposed cells become memory immune cells. If the same germ returns, they are ready, and you do not get sick. The video calls this natural immunity and says that in many cases it lasts a whole lifetime.

The catch is the price of getting it the natural way. Polio can cause permanent paralysis or death. Measles can cause swelling of the brain, which can mean permanent brain damage or death, especially in children under age five. Whooping cough, also called pertussis, can cause pneumonia, slowed or stopped breathing, and death, especially in babies under one year of age. Symptoms are not severe in everyone, but

there is no way to know in advance who will become very ill or die.

Vaccines give the same lesson without the disease. They often contain a small amount of weakened or killed germs; some instead carry genetic material, RNA or DNA, that instructs your own cells to make the germ's antigen. Usually the vaccine is a shot. The process does not make you sick, but it does produce memory cells and antibodies.

The video lists six main types: live attenuated, inactivated, toxoid, subunit and conjugate, mRNA, and viral vector, each with examples. It closes on community protection: when only a few people are vaccinated the risk of an outbreak is high, and when most are, a disease is much less likely to spread. That is community or herd immunity, and it matters most for people who cannot be vaccinated, including people with weakened immune systems, serious allergies, or other health conditions.

Why it's assigned

Unit 5 is the Combat layer, the one about ending a pandemic with medical countermeasures rather than containing it. The course's own framing is that vaccines and therapeutics are more pathogen-specific than the barriers of Unit 4, and that time is everything once a pandemic is running. This video is the floor under that argument: before the unit asks

how to build, test, finance and ship a vaccine in 100 days, it wants everyone holding the same picture of what a vaccine actually is and why a platform like mRNA is different from a killed-germ shot.

In depth

Immunization, and what the immune system already does

The video opens with a sponsor message for the Smart Imagebase, then defines immunization as becoming immune to or protected against a disease, usually by receiving a vaccine. Vaccines, it says, stimulate the immune system so you do not get sick or get an infection.

The explanation starts not with vaccines but with ordinary defence. The organs and cells of the immune system defend the body against harmful germs such as bacteria and viruses. Immune cells circulate constantly and monitor substances on the surfaces of cells called **antigens**. Healthy cells have different antigens from diseased body cells or foreign invaders, so immune cells usually ignore the antigens on healthy cells. When immune cells come across antigens on germs, they destroy the germ.

What happens next is the part that makes vaccination possible. After destroying a germ, the immune cell displays the germ's antigen on its own surface. That display activates other types of immune cells. Some of

those activated cells recognize the antigen on infected body cells and destroy those cells. Others, called plasma cells, make molecules called antibodies. Antibodies travel through the body and lock on only to germs carrying that specific antigen, which marks the germ for destruction, and then other immune cells attack the marked germs.

Memory, and the price of earning it naturally

Once the infection is gone, some of the immune cells that met the antigen become memory immune cells. If the same type of germ infects the body again, the memory cells are ready to destroy it, so you do not get sick. The video calls this natural immunity, and says that in many cases it will last your whole lifetime.

Then it makes the argument that the rest of the unit rests on. The problem with earning natural immunity by having the disease is that some naturally acquired infections cause serious complications or kill. Three examples are given. Polio can result in permanent paralysis or death. Measles can cause swelling of the brain, resulting in permanent brain damage or death, especially in children under age five. Whooping cough, also known as pertussis, can cause complications such as pneumonia, slowed or stopped breathing, and death, especially in babies under one year of age. The video adds the line that makes this more than a list: while symptoms may not be severe in

all people, it is not possible to know who will be affected enough to become very ill or even die.

What is in a vaccine

Vaccines, the video says, often contain a small amount of weakened or killed germs, but some contain genetic material such as RNA or DNA that provides instructions for your body's own cells to make the germ's antigen. Usually you receive a vaccine as a shot. Inside the body, the germ particles in the vaccine teach immune cells to attack those germs. The process does not make you sick, but it does cause the body to make memory cells and antibodies for those germs, so that if the germ turns up later in life the immune system is ready.

The six main types

- **Live attenuated** vaccines use alive but weakened germs. They are most like a natural infection and give strong disease immunity. Examples: measles, mumps, rubella, chicken pox, and the flu nasal spray.
- **Inactivated** vaccines use inactive or killed germs. You may need several doses or booster shots over time. Examples: hepatitis A, flu, polio, rabies.
- **Toxoid** vaccines protect against harmful substances made by germs, called toxins. They use weakened versions of the toxins, called toxoids, and may need boosters to maintain protection.

Examples: the diphtheria and tetanus parts of the DTaP vaccine. (The caption file spells this "diptheria"; the standard term is diphtheria.)

- **Subunit and conjugate** vaccines use only a specific part of a germ and give strong immunity to that key part. They may also require boosters. Examples: the pertussis part of DTaP, and hepatitis B.
- **mRNA** vaccines contain a special type of RNA called messenger RNA, which instructs your cells to make the germ's antigen, triggering an immune response. The video states plainly that these do not carry live virus and do not change your DNA. Examples: certain COVID-19 vaccines.
- **Viral vector** vaccines may use either DNA or mRNA to give cells instructions to make a germ's antigen, with the genetic material wrapped in a different, safe virus that carries the instructions into your cells. These also do not change your DNA. Examples: certain COVID-19 vaccines.

Notice what the last two have in common: nothing of the germ itself is in the vial, only the instruction for one of its parts. That is the property the rest of Unit 5 keeps returning to, because an instruction can be rewritten for a new pathogen in a way that a vat of weakened virus cannot.

Protection beyond the person vaccinated

The video closes on the community. Many of the diseases vaccines prevent spread from person to person. When only a few people in a community are vaccinated, the risk of an outbreak is high; if most people are vaccinated, it is much less likely for a disease to spread. That is called community immunity or herd immunity. It matters especially for people who cannot be vaccinated, which the video lists as people with weakened immune systems, serious allergies, or other health conditions. The final line is that vaccines protect you, your family and your community, and that to find out more you should talk to your healthcare provider.

Words to know

- **Antigen** A substance on the surface of a cell or germ that the immune system reads as an identity tag. It matters here because every vaccine type in the video is a different way of putting one antigen in front of the immune system without the disease attached.
- **Plasma cell** An activated immune cell that manufactures antibodies. It is the production line behind the protection a vaccine is trying to trigger.
- **Memory immune cell** A cell left behind after exposure that stays ready for the same germ. Memory is the whole point of vaccination; everything else is the setup.

- **Live attenuated** Alive but weakened. These are closest to a real infection and give strong immunity, which is also why they are the hardest kind to use safely in a hurry.
- **Toxoid** A weakened version of a bacterial toxin. It shows that sometimes the dangerous thing is not the germ but what the germ makes, so the vaccine targets the poison instead.
- **Herd immunity** Enough people immune that the germ struggles to find a chain of onward hosts. It is what turns individual protection into a population defence, and it is the reason coverage, not just efficacy, matters in a pandemic.

Numbers in this reading

- Six main types of vaccine: live attenuated, inactivated, toxoid, subunit and conjugate, mRNA, and viral vector.
- Under age five: the age group the video names as most at risk of brain swelling, permanent brain damage or death from measles.
- Under one year of age: the age group the video names as most at risk of pneumonia, slowed or stopped breathing, and death from whooping cough.

What the author is unsure about

The video states no uncertainty at all. It is a settled explainer with one honest hedge inside it: while

symptoms may not be severe in all people, it is not possible to know in advance who will be affected enough to become very ill or even die. It gives no efficacy figures, no timelines, no failure rates, and no discussion of how long a given vaccine's protection lasts, so take it as the vocabulary for the unit rather than as evidence about any particular vaccine.

For the discussion

- The video's claim that natural immunity often lasts a lifetime is exactly why the case for vaccination is a case about the price of admission, not about whether immunity works; the question the unit then has to answer is how fast that price can be dropped for a germ nobody has met yet.
- Two of the six types, mRNA and viral vector, deliver an instruction rather than a germ, which is the hinge between this video and the 100-day target the rest of the unit argues over.
- Herd immunity is described here as protection for people who cannot be vaccinated, and that framing is worth holding against the unit's own caveat that a capable adversary would not attack with something an existing vaccine blocks.

5.1-2 · CORE · COURSE TIME 7 MIN

Vaccines 101: How new vaccines are developed

Nature Video · 2020

Read the original paper → · the course assigns: the video's own caption transcript (saved subtitle file)

> **Making a vaccine for a new disease is a long chain of steps, each of which can be sped up but none of which can be skipped, and the newest platforms shorten the chain mainly at the start.**

Average summary

Summarized from the video's caption transcript, not from its pictures. The saved track is the video's own written caption file, so the wording is clean and no technical term came through garbled.

Nature Video walks the whole path from a new disease to a licensed vaccine. First you identify the virus or bacterium causing the disease. Sometimes the pathogen itself can be weakened or deactivated and used directly. When it cannot, the next step is to identify the pathogen's **antigens**, the unique protein or glycoprotein markers a vaccine can be built around. For some pathogens these stay the same; for others, like the flu virus, antigens mutate and must be re-analysed every year.

Then you choose the vaccine platform, which is the design you will use: subunit vaccines that use only part of the pathogen, or newer, more experimental approaches like viral vector and nucleic acid vaccines.

You also choose the route: injection into skin or muscle, orally like the polio vaccine, or through the nose like some flu vaccines. Each route can change the response the vaccine produces.

Testing goes first in cell cultures, then in animals, sometimes including challenge studies where vaccinated animals are deliberately exposed. Humans can take part in challenge studies too, if the disease is not too serious or effective treatments exist. Once the candidate is proven safe in animals, you must show you can manufacture it to a high enough standard and at scale, which is GMP certification. Only then do the three human phases begin: a small safety trial, then hundreds of people to find the dose that triggers a big enough immune response, then thousands to measure effectiveness. That last phase is the slowest, because researchers have to wait for participants to meet the pathogen naturally. The video says it can take decades to gather enough data to be sure.

Even after licensing there are problems: mass production to a high standard is hard, delivery to where the vaccine is needed is hard, and Phase IV monitoring tracks rare adverse reactions that earlier trials missed. The process is very safe but very slow. In an emergency you can cut waiting time on paperwork and run phases at the same time, but steps are never missed out, and working this way

could produce a protective vaccine in as little as twelve to eighteen months.

The last third explains the newer platforms. Nucleic acid vaccines deliver genetic templates rather than antigens, often inside lipid nanoparticles that protect the material and help it enter cells, and some of those particles also act as adjuvants. RNA and DNA are much easier and faster to make in the lab than antigens, and potentially safer than whole-pathogen vaccines, but they are still experimental. Viral vector vaccines put an incomplete segment of the pathogen's genetic material inside a harmless virus, which delivers it; safe vectors already exist and swapping in new genetic material is relatively easy, and because the inserted material is incomplete it cannot replicate and cause disease.

Why it's assigned

Unit 5 is the Combat layer, and the course's claim is that the time between identifying a pathogen and protecting a population is the thing worth attacking. This video lays out that timeline step by step, which is what makes the rest of the unit legible: when the deep dives argue about trial design, financing or manufacturing, they are each arguing about one named box in this chain. It also plants the course's own line about platform vaccines, by showing why a nucleic acid or viral vector design can be rewritten faster than a vaccine built from the pathogen itself.

In depth

Finding something to aim at

The video opens with two questions: how do you create a vaccine to fight a new disease, and why does it take so long. Step one is identifying the virus or bacterium that causes the disease. In some cases the pathogen itself can then be weakened or deactivated and used to make a vaccine, but this is not always possible.

Where it is not, you identify the pathogen's antigens: unique protein or glycoprotein markers that can form the basis of a new vaccine. The video draws the distinction that matters for flu. For some pathogens, these antigens always stay the same. Other pathogens, like the virus that causes flu, can change or mutate their antigens, so these must be re-analysed every year.

Choosing the platform and the route

Once you have the pathogen or its antigens, you decide how to use them to build a vaccine. That decision is called the **vaccine platform**. The video names subunit vaccines, which use only part of the pathogen, and newer or more experimental technologies like viral vector and nucleic acid based vaccines, which it returns to later.

You also decide how the vaccine will be given: by injection into the skin or muscle, orally like the polio vaccine, or through the nose like some flu vaccines. The video notes that each of these options could affect the response the vaccine induces, which is a small line with large consequences for the mucosal approaches discussed elsewhere in this unit.

Preclinical testing and manufacturing standards

With the platform chosen and the vaccine designed, testing begins in cell cultures, then in animal models. This might include challenge studies, where animals that have been given the vaccine are deliberately exposed to the target disease to see whether the vaccine protects them. The video adds that further down the line, if the disease is not too serious or effective treatments are available, humans might take part in challenge studies too.

Before human trials, there is a step that is easy to overlook and that several later readings in this unit treat as a bottleneck: once the candidate is proven safe in animals, you have to prove you can manufacture it to a high enough standard for people and at scale. That is **GMP certification**.

The three phases, and why the third is slow

Human trials normally come in three phases, and all have to be passed.

1. The vaccine is given to a small group of healthy people to test for adverse reactions. This is the safety trial.
2. The vaccine goes to hundreds of people, to work out what dose is needed to trigger a big enough immune response.
3. The vaccine is trialled in thousands of people to see how effective it is.

The third step is often the slowest. Participants are sometimes recruited around the world, and researchers have to wait for participants to come into contact with the pathogen naturally. By tracking how many become infected in each group they can work out how effective the vaccine is at protecting the group. The video's blunt line is that it can take decades to gather enough data to be sure. If you do, the vaccine can finally be licensed.

This waiting-for-infection problem is the exact thing the clinical trials deep dive on page 5.4 is built around, and the thing human challenge trials are proposed to solve.

After licensing

Licensing is not the end. Mass producing a vaccine to a high standard can be a difficult job. Once you have cracked production, you still have to deliver the vaccine to where it is needed. And then there is **Phase IV monitoring**, keeping track of rare adverse

reactions the vaccine might cause even though it passed the earlier trials.

Speeding up without skipping

The video states the trade-off directly: this trial process is very safe, but that comes at a cost, because it is also very slow. In an emergency there are ways to speed things up, such as reducing the time spent waiting for paperwork to be completed, or running different trial phases at the same time. The sentence that follows is the one to hold onto: but steps are never missed out. By working more efficiently, a protective vaccine could be produced in as little as twelve to eighteen months. Advances in biotechnology could speed up production pipelines as well.

Nucleic acid vaccines

Nucleic acid vaccines do not use pathogens or antigens directly. Instead they deliver the genetic templates for the antigens. These DNA or RNA templates can be delivered directly, or inside lipid nanoparticles that help them enter cells and improve their stability after injection. The video describes these as nano-scale fat droplets that encapsulate the RNA or DNA, protecting it as it enters the body. Some lipid nanoparticles can also act as **adjuvants**, molecular triggers that kick start the immune response. Once inside the body, the templates

instruct our cells to start making antigens, which activate the immune response needed to build immunity.

The advantages, in the video's words: RNA and DNA are much easier and faster to produce in the lab than antigens are, and they are potentially safer than vaccines that contain whole pathogens, which means they could hopefully be used by people with weakened immune systems. The caveat, given in the same breath: nucleic acid vaccines are still experimental. Remember that the video is from 2020.

Viral vector vaccines

The other approach inserts an incomplete segment of genetic material from a pathogen inside a harmless virus that does not cause disease. That virus acts as a vehicle, delivering the material to the right place in the body, where it can be translated into proteins and trigger the immune system. The viral vector may even be self replicating, increasing the amount of vaccine in the body.

Viral vector vaccines are also quick to produce: safe vectors have already been established, and it is relatively easy to insert different target genetic material into these established carriers. Because the inserted genetic material is incomplete, they cannot replicate and cause disease, so the video calls them very safe.

The closing note

The video ends on limits rather than promise. There is only so much the process can be sped along. Effective vaccines can save millions of lives, but before they are used they have to be proven to be safe and effective. Looking forward, researchers hope to develop vaccines against non-infectious diseases like cancers, based on chemical markers called tumour neoantigens. The last line: vaccine technology is developing every day and has a bright future, albeit a slow and methodical one.

Words to know

- **Glycoprotein** A protein with sugar attached, often sitting on a virus's surface. It matters here because surface glycoproteins are the markers a vaccine usually teaches the immune system to recognize.
- **Vaccine platform** The underlying design used to present an antigen, as opposed to the antigen itself. This is the idea the whole unit turns on: change the insert, keep the platform.
- **Challenge study** A study in which a vaccinated subject is deliberately exposed to the disease. In animals it is routine; in humans the video allows it only where the disease is not too serious or effective treatments exist.
- **GMP certification** Proof that you can manufacture the product to the required standard and at scale. It sits between animal testing and human trials, so

manufacturing capability is a gate on testing, not just on rollout.

- **Lipid nanoparticle** A nano-scale fat droplet that wraps RNA or DNA, protects it and helps it enter cells. Some also work as adjuvants.
- **Adjuvant** A molecular trigger added to kick start the immune response. The video introduces it in passing; other readings in this unit treat adjuvant design as a frontier of its own.
- **Phase IV monitoring** Tracking of rare adverse reactions after a vaccine is licensed and in use. It is the acknowledgement that trials, however large, cannot see everything.

Numbers in this reading

- Three phases of human trials, all of which have to be passed.
- Phase 1: a small group of healthy people, to test for adverse reactions. The video gives no figure.
- Phase 2: hundreds of people, to find the dose that triggers a big enough immune response.
- Phase 3: thousands of people, to measure effectiveness.
- Decades: how long the video says it can take to gather enough data to be sure in the normal course of things.
- Twelve to eighteen months: what the video says a protective vaccine could take in an emergency, by

working more efficiently rather than by skipping steps.

What the author is unsure about

The video is confident about the shape of the process and careful about the newest parts of it. It states plainly that nucleic acid vaccines are still experimental, and that there is only so much the process can be sped along. Its twelve to eighteen month figure is offered as what emergency working could achieve, not as a measured result; it was published in 2020, and the COVID-19 vaccines went faster than that, which is the gap the rest of this unit examines. It says each delivery route could affect the response a vaccine induces without saying by how much, and it gives no failure rates, costs or efficacy figures anywhere.

For the discussion

- The video's rule that steps are never missed out is the standard every acceleration proposal in this unit has to meet, and it is worth asking which of them genuinely meet it and which just move risk somewhere else.
- The slowest step is waiting for participants to meet the pathogen naturally, which means the Phase 3 bottleneck is not scientific but epidemiological, and that reframes the whole clinical trials debate on page 5.4.

- GMP certification sits before human trials, so manufacturing capacity gates testing as well as rollout, which is a useful point to bring to the manufacturing deep dive on page 5.6.

5.1-3 · CORE · COURSE TIME 25 MIN

The Quest for Universal Flu Vaccines

Ryan Duncombe and Jasmin Kaur · 2025

Read the original paper → · the course assigns: whole piece

> **After nearly a century of flu vaccines that work about four times in ten, several new approaches have finally made a one-shot, all-strains vaccine plausible, though nobody has yet combined them or solved the problem of proving such a vaccine works.**

Average summary

Duncombe and Kaur open with the long history of influenza, from the name itself, which comes from 14th-century Italian phrases meaning influence of the stars and influence of the cold, to the 1918 pandemic that swept six continents and killed between 17 and 100 million people, as much as **5 percent** of the global population. Pandemics in 1957 and 1968 each killed more than one million. For much of modern history, flu pandemics have emerged every 10 to 50 years.

The science caught up slowly. The influenza virus was not isolated until 1933. The first flu vaccines were tested on US military personnel in the early 1940s, with a civilian rollout in 1945, and most seasonal vaccines today are made by a method almost identical to that one: grow the virus in fertilized chicken eggs, inactivate it, and teach the immune system to make antibodies against hemagglutinin, the protein that sticks out of the virus surface.

The results have never been good enough. The seasonal vaccine reduces the risk of infection by just 40 percent on average, against more than 95 percent for smallpox and measles. The 2014-2015 vaccine was only 19 percent effective because of a strain mismatch; even a good year such as 2010-2011 tops out at about 60 percent. The authors give three reasons. Flu mutates up to 1,000 times faster than human genomes, because its copying enzyme cannot proofread. Strains for each season must be picked 6 to 8 months ahead. And the immune system itself gets in the way: antibodies aim at the head of hemagglutinin, which varies most, and the body prefers reusing old antibodies rather than making new ones, a habit called antigenic sin.

The middle of the piece surveys what is being tried. At least 207 universal flu vaccine candidates are in development, 165 preclinical and 42 in clinical trials. Approaches include training the immune system only

on the more conserved hemagglutinin stalk; chimeric antigens that keep one stalk and swap the heads; multivalent mixes, such as Centivax's shot combining hemagglutinin from up to 30 strains; and mosaic nanoparticles such as FluMos-v1, which displays four hemagglutinin proteins at once. Breadth is only half the job; potency and durability are the other half, which brings in inhaled mucosal vaccines and a new generation of saponin adjuvants.

The authors' real argument is in the last section. No one has yet merged these strategies into one vaccine, and in theory a computationally optimized mosaic nanoparticle, given through the nose, with next-generation adjuvants, should be possible. The obstacles are intellectual property, the fear of expensive clinical failure, the stricter bar that applies because vaccines are tested on healthy people, and the awkward problem of proving breadth when only one or two strains circulate each year. Challenge trials could solve that last one, and regulators are warming to them. The closing analogy is the 1920s belief that bacterial infections could never be cured, which collapsed when the first commercial antibiotic arrived.

Why it's assigned

This is the reality check inside the Combat layer. The unit's premise is that broad-spectrum vaccines could be developed, tested and stockpiled ahead of time, and flu is the hardest and best-documented test of

that premise: a hundred years of effort, a pathogen we know intimately, and a vaccine that still misses more than half the time. Reading it next to the course's own caveat, that a capable adversary would not attack with something an existing broad-spectrum vaccine blocks, gives you both halves of the argument: the technology is finally plausible, and even if it arrives it is not the whole answer.

In depth

Plagues and peril: the long record

Influenza has gone by many names, including catarrh, grippe and sweating sickness. The modern name comes from the 14th-century Italian phrases influenza di stelle and influenza di freddo, meaning influence of the stars and influence of the cold. Astrologers believed the disease's periodic return was influenced by the location of the stars. Medical texts, including Avicenna's The Canon of Medicine, linked it to environmental change. Medieval physicians blamed imbalances in the four humours, shifting with the seasons: spring meant an excess of blood, winter an increase in phlegm. Religious texts by Saint Augustine and others read the outbreaks as calls for repentance.

In 1918 influenza swept across six continents and killed between 17 and 100 million people, as much as 5 percent of the global population. The authors footnote that mortality estimates for 1918 vary widely

and have tended to increase over the years. In 1957 and 1968, large influenza pandemics struck again, each killing more than one million people. Similar episodes stretch back to ancient Greece and China, and for much of modern history flu pandemics have emerged every 10 to 50 years.

During the 1918 pandemic most scientists believed bacteria caused influenza. The virus was not isolated until 1933, when an epidemic in London let British researchers Wilson Smith, Christopher Andrewes and Patrick Laidlaw collect human influenza virus from patients' throat washings.

How the modern flu vaccine was built, and why it has not changed much

In 1931, Ernest Goodpasture, a pathologist at Vanderbilt University, and Alice Woodruff, a laboratory technician, worked out how to grow viruses in fertilized chicken eggs, which were protein-rich, nutrient-rich and extremely inexpensive. Around 1940 to 1941, Thomas Francis Jr., a physician and epidemiologist at the University of Michigan, decided to make a vaccine from egg-grown viruses. The US Army initially feared this would cause the virus to mutate and repeat 1918. In response, Francis's team, which included a young Jonas Salk, developed a way to inactivate viruses with formaldehyde so they could be recognized by the immune system without causing disease.

Formaldehyde already had a precedent: Gaston Ramon, a French veterinarian, had used it to inactivate pathogens while developing diphtheria and tetanus vaccines in the late 19th and early 20th centuries. It converts bacterial toxins into toxoids by cross-linking the pathogen's proteins and nucleic acids, preserving what stimulates the immune system while removing the virulence.

That first flu vaccine worked, as most since have, by training the immune system to make antibodies that neutralize **hemagglutinin**, a protein protruding from the virus surface. Antibodies bound to hemagglutinin physically block the virus from attaching to and entering human cells. Tagged viruses then become easy targets for white blood cells such as macrophages and neutrophils.

Most seasonal flu vaccines today are made by a nearly identical method to the one Francis and Salk used in 1941. Better surveillance and forecasting allow trivalent or quadrivalent vaccines, made by injecting multiple viral strains into eggs. The 2024-2025 seasonal vaccines include three viruses: two influenza A strains, H1N1 and H3N2, and an influenza B virus. The World Health Organization selects these strains months in advance using global surveillance data.

Because strains keep shifting, vaccines must be given every year, and efficacy varies a great deal between years. The 2014-2015 seasonal vaccine was only 19

percent effective because of a mismatch between the selected strain and the circulating one. Even in a good year, such as 2010-2011, the vaccine tops out at about 60 percent efficacy. The authors' verdict: after nearly a century, influenza vaccines remain hit-or-miss.

Viral evaders: why the target keeps moving

Strains for each year's vaccine must be predicted a full 6 to 8 months in advance, to give manufacturers time to ramp up production before the season. That is hard when the virus mutates daily.

Flu viruses mutate up to 1,000 times faster than human genomes. The enzyme that replicates influenza's genetic material, RNA polymerase, lacks proofreading, so mutations are not corrected and accumulate quickly. Some weaken the virus; others help it evade immunity.

Antigenic drift is the gradual version: mutations, often in hemagglutinin and neuraminidase, change the virus physically so the immune system struggles to recognize it. Even across strains of influenza A, the structure of hemagglutinin proteins can vary by more than 50 percent, so antibodies that bind one strain may not work against another. Over time, drift produces new strains that partly or completely escape existing population immunity.

Antigenic shift is the abrupt version and the bigger threat: two or more different flu viruses infect the

same host and exchange genetic material, creating an entirely new strain. The 2009 H1N1 swine flu emerged when human, avian and swine strains mixed inside pigs. It was highly pathogenic and at times lethal, causing between 151,700 and 575,400 deaths worldwide during the first year of the outbreak, mostly in Africa and Southeast Asia.

Animal reservoirs make eradication harder still. Influenza A infects bats, cats, dogs, birds, horses and pigs as well as humans, so the virus can persist in animals while humans are uninfected and spill back later. That is why eradicating influenza is much tougher than eliminating a virus that infects only humans, like smallpox, which was eradicated in 1980.

Challenging the immune system

The authors call the immune system, borrowing a line from science writer Ed Yong, arguably the most complex part of the body outside the brain, and they treat its quirks as a bottleneck in their own right.

During an infection, an immune cell coated with antibodies binds a foreign antigen, which sets off a cascade telling the body to make more antibodies of the same kind. These focused antibodies can stick to the original antigen 10,000 times more strongly than the initial cell that met the pathogen.

For influenza, antibodies usually target hemagglutinin because it is the most exposed part of the virus. The

authors picture it as a scoop of ice cream on a waffle cone. The trouble is that the top part, the head, varies widely between strains and mutates fastest, so most antibodies quickly become useless. The stalk region is more conserved across strains.

The second quirk is **antigenic sin**: the immune system's tendency to reuse existing antibodies rather than build new ones for each strain, even when they are a poor match. Once it has made antibodies against H5N1, for instance, it keeps recycling them for newer strains, which undermines updated vaccines. Annual shots still matter, because they boost the supply of old antibodies, and in large enough quantities weakly binding antibodies offer some protection, but not for long. That is the authors' explanation for why a century-old approach cannot deliver lasting immunity.

So a truly universal vaccine has to beat the virus's shapeshifting, deliver high efficacy, and protect against multiple strains at once.

Making universal vaccines: four routes to breadth

None are approved yet. According to the Universal Influenza Vaccine Tracker, at least 207 universal flu vaccine candidates are in development, with 165 in preclinical trials and 42 in clinical trials. A successful universal vaccine must do two things: protect broadly

across strains, and be potent enough to last several years.

- **Stalk-only.** About ten universal candidates train the immune system solely on hemagglutinin stalks rather than heads, because stalks are more similar between strains and mutate more slowly. In one study, stalk-directed responses elicited significantly higher antibody titers than a seasonal comparator against multiple H1 strains as well as an H5 strain.
- **Chimeric antigens.** Mix-and-matched hemagglutinin proteins, none of which naturally infect people, carrying different heads on a single shared stalk. Repeated doses swap in new heads while the stalk stays the same, diluting the response to the head and boosting the broadly protective response to the stalk. Researchers at the Mount Sinai School of Medicine have tested a chimeric vaccine in a Phase I trial and it has now moved into Phase 2. Early results showed some flu protection, but manufacturing these complex proteins is challenging: each protein must be individually engineered, and mixing them can lead to unpredictable stabilities.
- **Multivalent mixes.** Rather than fusing antigens into one molecule, mix them. Centivax is building a vaccine that mixes hemagglutinin proteins from up to 30 different flu strains in a single shot, a more straightforward way to get high antigen numbers

than the 3 to 6 antigens a chimeric vaccine can offer. Antibodies targeting conserved regions are exposed to a 30 times greater dose than those targeting variable regions. Publicly available animal data show a broadly protective antibody response against a panel of 36 strains of flu, including all pandemic strains from the past hundred years and multiple strains not present in the vaccine. It is still in preclinical development.

- **Nanoparticles.** Layering antigens onto nanoparticles mimics the natural structure of viruses and may improve the response, since antibodies bind antigens more effectively when the antigens are anchored rather than floating free. Mosaic nanoparticles display several hemagglutinin proteins on one particle. FluMos-v1 displays four at once and recently wrapped up a Phase I clinical trial; results are not public yet.

Potency and durability

Breadth alone is not enough. One route to longer-lasting immunity is to inhale vaccines rather than inject them. Mucosal vaccines set up defences in the respiratory tract where influenza enters, which also curbs transmission and viral mutation, and being needle-free may raise uptake.

They are hard to make work. The response triggered in the lungs can fade quickly unless it is strong enough to spill over into systemic defences, and the

vaccine has to cross the mucosal barrier and provoke a robust response without excessive inflammation. There is only one FDA-approved intranasal vaccine for any disease: the FluMist Quadrivalent flu vaccine, which uses weakened viruses.

The cautionary case is from 2004, when a study of an intranasal flu vaccine given to people in Switzerland resulted in 46 cases of Bell's palsy, a short-lived paralysis that makes half the face sag. Researchers now believe the cause was the vaccine's **adjuvants**, the additives that act as inflammatory agents to amplify the immune response, rather than the vaccine itself.

New adjuvants are hard precisely because they must inflame enough to boost the response but not so much as to cause harm, and what stimulates one person may over-inflame another. The promising class is saponins, derived from the bark of the Quillaja saponaria tree, native to South America. They have been used in veterinary vaccines for decades and are potent immune activators, though the cruder veterinary preparations can cause serious side effects in humans. Refined saponin adjuvants now appear in the RTS,S and R21 malaria vaccines, in Shingrix, which offers over 90 percent efficacy, and in Novavax's coronavirus vaccine.

The century ahead

Most efforts at universal flu vaccines have achieved only incremental improvements over seasonal vaccines, and, importantly, no one has yet tried merging all these strategies into a single vaccine. In theory it should be possible to build one that uses a computationally optimized mix of hemagglutinin proteins in a mosaic nanoparticle for breadth, is delivered intranasally for mucosal immunity, and contains next-generation saponin adjuvants for potency, and that could even be mixed with non-hemagglutinin antigens to widen coverage further.

Even then the obstacles are not scientific. Intellectual property concerns and the fear of expensive clinical trial failures could deter researchers from launching early-stage tests. Vaccines, unlike most medications, must be tested on healthy people, which drives stricter approval hurdles, especially for new adjuvant formulations. And there is the question the authors flag as the deepest: how do you prove a vaccine defends against multiple strains of flu when only one or two circulate each year?

One answer is the challenge trial, deliberately infecting volunteers with different strains to measure broad-spectrum protection directly. Regulators including the FDA and the European Medicines Agency are increasingly recognizing the value of such trials, and the FDA is exploring more flexible approval pathways for vaccines addressing critical unmet

needs, including accelerated approval relying on immunogenicity data and challenge trial results rather than long-term outcome studies. Easing the regulatory process, the authors argue, may motivate more companies to invest in ambitious designs.

The piece ends with a historical analogy. In the 1920s many doctors believed curing bacterial infections was impossible, a collective resignation that Thomas Hager describes in The Demon Under the Microscope as a belief that medicine could go no further. Influenza and pneumonia alone caused nearly 30,000 deaths among US Army soldiers during World War I, more than half of the 52,000 non-combat deaths recorded. That defeatism dissolved in the early 1930s when a small research team at IG Farben in Elberfeld, Germany, led by the pathologist Gerhard Domagk, combined an azo dye with a sulfonamide side chain to make sulfamidochrysoïdine, sold as Prontosil, the first commercially available antibiotic, effective against both staphylococcal and streptococcal infections. The authors' closing claim is a forecast, not a finding: with enough ingenuity, a one-shot universal flu vaccine will not stay out of reach forever.

The authors note their own backgrounds. Ryan Duncombe previously worked on universal flu and COVID vaccines at Alvea and on vaccine policy at 1Day Sooner, with a PhD in immunology from the University of Chicago. Jasmin Kaur previously led the

pandemic preparedness and human challenge studies program at 1Day Sooner and worked at GSK. The article was published on 26 January 2025.

Words to know

- **Hemagglutinin** The protein sticking out of the flu virus surface that lets it attach to human cells, and the main target of flu antibodies. Its head mutates fast; its stalk does not, which is the whole design problem.
- **Antigenic drift** Gradual accumulation of mutations that changes how the virus looks to the immune system. It is why the vaccine has to be remade every year.
- **Antigenic shift** Two flu viruses infecting one host and swapping genetic material, producing a brand new strain. It is the mechanism behind pandemic flu, as in 2009.
- **Antigenic sin** The immune system's habit of reusing existing antibodies against a new strain instead of making better-matched ones. It puts a ceiling on what an updated shot can achieve.
- **Chimeric antigen** An engineered hemagglutinin with a different head on a shared stalk, designed to focus the immune response on the conserved part.
- **Mosaic nanoparticle** A particle displaying several different hemagglutinin proteins at once, mimicking a virus's physical structure to strengthen and broaden the response.

- **Mucosal vaccine** A vaccine given through the nose or mouth that builds defences where the virus enters. There is only one FDA-approved intranasal vaccine for any disease.
- **Adjuvant** An additive that amplifies the immune response. Getting the inflammation right is the hard part, as the 2004 Bell's palsy cases showed.

Numbers in this reading

- 17 to 100 million: deaths in the 1918 influenza pandemic, as much as 5 percent of the global population. The authors footnote that estimates vary widely and have tended to rise.
- More than one million: deaths in each of the 1957 and 1968 pandemics.
- Every 10 to 50 years: how often flu pandemics have emerged for much of modern history.
- 1933: the year the influenza virus was isolated, from throat washings during a London epidemic.
- Early 1940s and 1945: first flu vaccines tested in US military personnel, then civilian rollout.
- 40 percent: average reduction in infection risk from the seasonal flu vaccine.
- More than 95 percent: effectiveness of smallpox and measles vaccines at preventing disease.
- 19 percent: effectiveness of the 2014-2015 seasonal vaccine, due to strain mismatch.

- About 60 percent: the ceiling in a good year such as 2010-2011.
- Three viruses in the 2024-2025 seasonal vaccine: influenza A H1N1 and H3N2, plus an influenza B.
- 6 to 8 months: how far ahead strains must be predicted so manufacturers can produce in time.
- Up to 1,000 times: how much faster flu viruses mutate than human genomes.
- More than 50 percent: how much hemagglutinin structure can vary even across influenza A strains.
- 151,700 to 575,400: deaths worldwide in the first year of the 2009 H1N1 outbreak, mostly in Africa and Southeast Asia.
- 1980: the year smallpox was eradicated.
- 10,000 times: how much more strongly focused antibodies bind the original antigen than the first cell that met the pathogen.
- 207 universal flu vaccine candidates in development, 165 preclinical and 42 in clinical trials, per the Universal Influenza Vaccine Tracker.
- About ten: candidates using the stalk-only approach.
- Up to 30 strains in Centivax's multivalent shot, against 3 to 6 antigens for a chimeric vaccine, giving conserved-region antibodies a 30 times greater dose.
- 36 strains: the panel against which Centivax's animal data showed a broadly protective response,

including all pandemic strains of the past hundred years.

- Four: hemagglutinin proteins displayed at once by the FluMos-v1 mosaic nanoparticle.
- 46: cases of Bell's palsy in the 2004 Swiss intranasal flu vaccine study, now attributed to the adjuvants.
- Over 90 percent: efficacy of Shingrix, one of the vaccines using a saponin adjuvant.
- Nearly 30,000 of 52,000: US Army non-combat deaths in World War I caused by influenza and pneumonia.

What the author is unsure about

The authors are careful about what is evidence and what is hope. No universal flu vaccine is approved. The stalk-only result they cite is one study of antibody titers, not protection in people. The Mount Sinai chimeric candidate showed some protection in Phase I and manufacturing is hard. Centivax's breadth data are from animals and the vaccine is preclinical. FluMos-v1's Phase I results are not public. They flag that the 1918 death toll is a wide and rising range, and that researchers only now believe the 2004 Bell's palsy cases came from adjuvants rather than the vaccine itself. The whole final section, the combined vaccine and the claim that a one-shot universal vaccine will eventually arrive, is explicitly theoretical and an opinion about the future, supported by an

analogy to antibiotics rather than by data. They also leave open the hardest practical question: how to demonstrate protection against strains that are not circulating.

For the discussion

- Flu is the strongest counterexample to the hope that broad-spectrum vaccines can be stockpiled in advance: a hundred years of work on a familiar pathogen still yields about 40 percent average protection, and the authors' own best case is an untried combination of four separate advances.
- The binding constraint in the final section is regulatory and commercial rather than scientific, which puts this reading in direct conversation with the clinical trials and financing deep dives: how do you test breadth when only one or two strains circulate, and who pays for a trial that might fail expensively?
- Antigenic sin is a point worth raising against any plan that relies on repeated boosting, because the immune system keeps reaching for the antibodies it already has rather than the ones the new strain requires.

5.1-4 · CORE · COURSE TIME 25 MIN

Day Zero Antivirals for Future Pandemics

Brian Wang · 2024

Read the original paper → · the course assigns: whole piece

> **Even a 100-day vaccine arrives too late for a fast pathogen, so Wang argues we should copy the innate immune system and build broad-spectrum antivirals that target what all viruses have in common, are taken before infection, and already sit on shelves on day zero.**

Average summary

Wang opens with two timelines. Ebola tore through West Africa in 2014, killing an estimated **11,325** people in two years, and it took another three years after the outbreak ended for Merck's Ebola vaccine to be approved. For COVID-19, the mRNA vaccines from Pfizer and Moderna took just 326 days from the initial sequencing of the virus to emergency use authorization. CEPI, which raised $2 billion in 2022 from private investors, nonprofits and more than a dozen national governments, wants that down to 100 days.

His argument is that even 100 days can be too slow. After China ended its zero-COVID policy in 2022, estimates suggest Omicron infected upwards of 80 percent of its population, over a billion people, in about one month. Something has to cover the gap, and the candidate is the broad-spectrum antiviral: a

drug that works against many viruses and can be deployed at the first sign of an outbreak, so the response starts on day zero rather than day 100.

So far these drugs have disappointed. Remdesivir, a decade-old Gilead antiviral previously tested against Ebola, got emergency authorization against COVID-19 in May 2020, seven months before the first mRNA vaccines, but manufacturing lagged until October, and one large-scale trial found it reduces the relative risk of death by only about ten percent in hospitalized patients.

Wang's proposal is to learn from the innate immune system, which he says succeeds for three reasons: its tools target the most fundamental properties of viruses, they are prophylactic rather than curative, and they are widespread. Mucus is his model. It blocks viruses from physically reaching cells, which nearly all viruses must do; it works before infection; and the body makes it everywhere a virus might enter.

The middle sections survey work in that direction: porcine gastric mucins that cut infections tenfold in cells; DRACOs and dsCAREs, engineered proteins that make a cell kill itself when it detects double-stranded RNA; and newer prophylactics such as the PCANS nasal spray and the inhaled SHIELD powder, which cut viral RNA in monkeys about 50 times in the lungs and 300 times in the nose. Wang discloses that he

co-founded Panoplia Laboratories, a nonprofit pursuing DNA-encoded versions of this idea.

The last section is about shelves, not science. At the start of the pandemic the world's supply of remdesivir was 5,000 treatment courses, and it took Gilead nine months to reach 500,000, while hydroxychloroquine, already on the market for other diseases, reached about 500,000 off-label prescriptions in weeks. Broad-spectrum drugs struggle to find a pre-existing market because regulators approve drugs for specific diseases. The workarounds, going drug-free or repurposing old drugs, avoid approval but also avoid the human trials that would prove the drugs work.

Why it's assigned

This is the therapeutics half of the Combat layer, and the counterweight to the vaccine readings on the same page. The course's framing is that medical countermeasures are how we exit a pandemic and that speed decides how many die first; Wang's piece argues that even the most ambitious vaccine timeline leaves a gap at the start, and that filling it requires a different kind of product with a different economic problem. It also pairs with the unit's own caveat about a sophisticated adversary: a drug that targets what all viruses must do is harder to design around than a vaccine aimed at one pathogen's antigen.

In depth

The gap at the start of a pandemic

Wang's opening comparison sets the scale. The 2014 West African Ebola outbreak killed an estimated 11,325 people in two years, and the first Ebola vaccine, from Merck, was approved three years after the outbreak ended. COVID-19 was different: mRNA vaccines from Pfizer and Moderna took just 326 days from the initial sequencing of the virus to emergency use authorization. The Coalition for Epidemic Preparedness Innovations, a nonprofit that funds vaccine research and development, raised $2 billion in 2022 from private investors, nonprofit organizations and over a dozen national governments to pursue pandemic vaccines within 100 days.

Then the complication. Even 100 days will not always be fast enough. The Omicron variant spread with alarming rapidity; after China ended its zero-COVID policy in 2022, estimates suggest Omicron infected upwards of 80 percent of its population, over a billion people, in about one month.

Broad-spectrum antivirals, which treat infections caused by multiple viruses, could bridge the wait for a vaccine, and ideally would be deployed at the earliest signs of an outbreak, so that the response begins on **day zero** rather than day 100.

Why the existing drugs disappoint

Remdesivir is Wang's case in point. It is a decade-old broad-spectrum antiviral developed by Gilead and previously tested against Ebola. It was quickly approved for emergency use against COVID-19 in May 2020, seven months earlier than the first mRNA vaccines. But lengthy manufacturing timelines slowed deployment: it was not until October, five months later, that drug manufacturers scaled up enough to meet national demand. And it was not as effective as hoped. One large-scale clinical trial found it only reduces the relative risk of death by about ten percent in hospitalized patients.

The scientific reason the current generation falls short is in the next section. Estimates suggest more than 600,000 different viruses could infect humans, most of them still unknown. Remdesivir and similar drugs work by jamming RNA-dependent RNA polymerase (RdRp), an enzyme nearly all RNA viruses use to copy their genomes, but the structure of RdRp differs between viruses, so many are relatively unaffected. Remdesivir substantially reduces the ability of Ebola virus, coronaviruses and RSV to infect cells, but fails against Chikungunya virus, influenza and some hepatitis viruses.

The innate immune system as a design brief

Wang's organising idea is that the most experienced broad-spectrum antiviral developer is our own innate immune system: the skin, the mucus lining our organs,

the immune cells patrolling tissues, and the antiviral molecules inside cells that disrupt viral replication. He says it has been evolutionarily successful for three reasons, and those three become the specification for the next generation of drugs.

Take mucus. It resembles a dense spider web lining the internal organ surfaces, trapping viruses before they can infect the underlying cells. 1. It targets a **fundamental property** of viruses, that they must physically reach host cells to infect them, so almost all viruses are hindered by it. 2. It is **prophylactic**: it stops cells being infected in the first place rather than trying to curb infection after replication has spiralled. 3. It is **widespread**: cells all along the respiratory, digestive and reproductive tracts produce it constantly, so viruses meet it whatever their portal of entry.

The next generation of broad-spectrum antivirals, Wang argues, should aspire to these mucus-like qualities, so that they work against the next pandemic virus regardless of its identity and can be pulled off shelves immediately at an outbreak's onset.

Early work in that direction

Two strands of research already follow this logic.

In 2012, a research group at MIT isolated the major components of mucus from pig stomachs, called porcine gastric mucins, and tested them against

human papillomavirus, influenza virus and Merkel cell polyomavirus. Cells covered in a one percent mucus solution had ten times fewer infections than control cells covered with a standard buffer. Watching fluorescently labelled virus-like particles under a microscope, they saw the mucus matrix immobilize virus particles. Wang notes that adding such mucins to hygiene products like mouthwash and toothpaste could give regular boosts of antiviral protection.

The second strand targets double-stranded RNA. Human cells have low levels of it, while most viruses produce it as part of their replication cycle, which is also why the innate immune system treats it as a signal that an infection is underway. In 2011, research groups at the MIT Lincoln Laboratory and Boston University fused two proteins: one binds double-stranded RNA, the other triggers apoptosis, or cell death. Inside a cell, the pair works like a quick-trigger suicide pill. If a virus infects the cell, the double-stranded RNA causes the cell to kill itself, cutting off the virus's ability to use that cell's resources. Across enough cells, the infection burns out. Both groups showed these proteins, named DRACOs and dsCAREs, reduced viral replication across ten separate viral families in cells and three separate viral families in mice. Wang footnotes that this count includes two subsequent publications and one public presentation as well as the original papers.

He then applies the brake himself: promising preclinical data is not enough to produce a day zero antiviral, and real-world clinical efficacy does not follow from it. Remdesivir was effective against coronaviruses in lab studies and underperformed clinically.

Why timing is the hardest part

Viral infection grows exponentially. One infected cell typically makes 10 to 100 more virus particles that go on to infect other cells, and within days the number of particles in the body balloons to billions and you feel sick. For COVID-19 and influenza, peak viral load often falls on or around the day of symptom onset, after which the amount of virus declines as the immune system catches up. If you are unlucky, the immune system goes into overdrive even as viral load subsides, damaging the lungs and sending you to hospital.

That dynamic is brutal for antiviral treatments. They work by disrupting replication, so they should be given while the virus, not the immune overreaction, is driving symptoms. But by the time a patient gets treatment, perhaps a week after symptoms began, as happened early in the pandemic when only hospitalized patients were eligible for remdesivir, viral load may already be falling. Starting after the peak inherently limits what the drug can do.

The solution Wang favours is prophylaxis: take the drug before you get sick, so it slows viral growth during the early viral phase. One computational model published in 2020 suggested that antivirals given prior to SARS-CoV-2 infection could reduce peak viral load by more than a million times compared with antivirals given three days post-symptoms.

Converting existing treatments into prophylactics is not simply a matter of timing, though. Remdesivir is given by intravenous injection, and half of its molecules are cleared from the bloodstream every day. Wang footnotes that remdesivir is converted within hours into an active metabolite, GS-441524, which has a half-life of one day. Taking it prophylactically would mean constant clinic visits for infusions of up to two hours, which is hardly practical for a healthy person.

The new prophylactics

Three lines of work aim at drugs that are easier to deliver, longer lasting, or both.

- **PCANS**, reported last year by a research team at Harvard University, is a broad-spectrum antiviral prophylactic nasal spray designed for daily use. Like mucus, it forms a physical barrier in the nasal cavity that traps incoming virus particles. In mice it was safe and effective: PCANS-treated mice survived influenza infections while control mice

died. In a separate study, mice dosed daily for two weeks had no significant side effects, and the spray lingered in their nasal cavities for eight hours after each dose.

- **SHIELD** is made of polymer-based microparticles inhaled as a dry powder. On contact with mucus in the lungs it swells and forms a hydrogel network reinforcing the native mucus layer. Inhaled eight hours before infection with different SARS-CoV-2 variants, it reduced viral RNA in monkeys by about 50 times in the lungs and 300 times in the nose. Safety tests showed no toxicity in mice treated daily for two weeks.
- **DNA-delivered antivirals** could allow much less frequent dosing. DNA delivered to the nose or lungs can express an encoded protein for months to over a year in mice if the vector, the backbone controlling the level and timing of expression, is engineered to be long-lasting. In principle this could give a prophylactic dosed once a season or once a year. DNA-encoded antiviral proteins have already protected mice against influenza or COVID-19, but Wang flags the limit of those experiments honestly: the animals were dosed one to two weeks before infection, whereas a real test of prophylaxis would dose them months or even a year ahead.

He then discloses his own stake. Panoplia Laboratories, the nonprofit he co-founded with

colleagues, is encoding double-stranded RNA-targeting proteins in long-lasting DNA vectors, testing whether candidates reduce viral replication in cells over time, then whether the vectors protect animals from multiple respiratory viruses for months, before moving to the clinic. He calls the goal lofty and names what might be needed: new double-stranded RNA-targeting proteins that react to viral double-stranded RNA without overreacting to the small amount present in human cells, which he describes as a potentially narrow tightrope to walk if the selectivity has to hold for months, and modified vectors that express the proteins at the right levels and only when needed.

Boosting supplies: the part that is not science

The closing section is about markets. At the beginning of the COVID-19 pandemic the worldwide supply of remdesivir was limited to 5,000 treatment courses, and it took Gilead nine months to scale to 500,000. By contrast hydroxychloroquine, already widely prescribed for rheumatoid arthritis, lupus and malaria, reached about 500,000 off-label prescriptions for COVID-19 in just a few weeks. Wang is careful to note that remdesivir was at least somewhat effective while hydroxychloroquine proved a false hope, but the supply difference is the point: hydroxychloroquine had approved indications and established markets, remdesivir had no approval for any commercially relevant indication before the

pandemic, and without a pre-existing revenue stream Gilead had little incentive to manufacture it in quantity.

The structural reason is regulatory. Regulators and consumers favour drugs that treat specific diseases. When regulators review a new drug application they weigh benefits and risks for a specific disease in a particular patient population, and efficacy against other diseases is not usually considered relevant.

Developers are finding ways around this, each with a cost: - **Avoid approval entirely.** PCANS is made entirely from components on the FDA's Inactive Ingredients Database and Generally Recognized as Safe list, so it is classified as a drug-free product, does not require FDA approval, and can be freely purchased online. - **Repurpose an existing drug.** Researchers at Yale University found that neomycin, a common antibiotic in the over-the-counter ointment Neosporin, stimulates innate immunity and has broad-spectrum antiviral activity. Administered into the noses of mice it protected them from both influenza and SARS-CoV-2. Human volunteers who swabbed their noses with Neosporin twice daily for one week had stimulated innate immune systems and tolerated the drug well.

The downside, stated plainly: companies may lack incentives to conduct human clinical trials for drug-free antivirals or repurposed drugs. Drug-free

products can enter the market without such trials, and repurposed drugs are often off-patent, which makes trials a high-cost, low-reward proposition for sponsors. Without human trials, we cannot be certain these antivirals are truly effective, which is why Wang says there remains significant value in navigating the traditional regulatory path.

Panoplia is taking that path. Its strategy is to find market niches underserved by existing antivirals and vaccines, such as immunocompromised patients who lack effective prophylactic options for respiratory infections because vaccines work less well in them. Wang adds the caveat against his own plan: other teams are developing virus-specific prophylactics for the same patients, and if they succeed it would be harder for broad-spectrum antivirals to reach that market, so the indication and population will have to be reassessed on an ongoing basis.

The piece ends where it began. Creating a day zero antiviral will require reinventing what broad-spectrum means: not only viral protein-targeting treatments like remdesivir but fundamental property-targeting prophylactics like PCANS, SHIELD and DNA-encoded DRACOs. Evolution shaped the innate immune system over millions of years and gave us the blueprint; the open question Wang leaves is whether the protections we discover will be widely available in time for the next pandemic.

Words to know

- **Broad-spectrum antiviral** A drug that works against many different viruses rather than one. The appeal is that it can be stockpiled before anyone knows what the next pathogen will be.
- **Prophylactic** Taken before infection to prevent disease, rather than after to treat it. Wang's central argument is that timing, not chemistry, is what limits current antivirals.
- **Innate immune system** The body's fast, general first line of defence: skin, mucus, patrolling immune cells and antiviral molecules inside cells. Wang treats it as a design brief rather than a background fact.
- **RdRp (RNA-dependent RNA polymerase)** The enzyme nearly all RNA viruses use to copy their genomes, and remdesivir's target. Its structure differs between viruses, which is why that class of drug is only partly broad.
- **Double-stranded RNA** A molecule most viruses make while replicating and human cells mostly do not. That difference makes it a plausible target for drugs that do not need to know which virus they are fighting.
- **Apoptosis** Programmed cell death. DRACOs and dsCAREs use it deliberately, killing an infected cell so the virus cannot use it.

- **Viral load** The amount of virus in the body. For COVID-19 and influenza it peaks around the day symptoms start, which is the fact that undermines late treatment.
- **Generally Recognized as Safe** An FDA list of ingredients that do not raise safety concerns. PCANS is built only from such components, which is how it reaches the market without approval, and without trials.

Numbers in this reading

- 11,325: estimated deaths in the 2014 West African Ebola outbreak, over two years.
- Three years: the gap between the end of that outbreak and approval of Merck's Ebola vaccine.
- 326 days: from initial sequencing of the virus to emergency use authorization for the Pfizer and Moderna mRNA COVID-19 vaccines.
- $2 billion: raised by CEPI in 2022 toward delivering pandemic vaccines within 100 days.
- Upwards of 80 percent, over a billion people, in about one month: estimated Omicron infections in China after zero-COVID ended in 2022.
- May 2020 and October 2020: remdesivir's emergency authorization against COVID-19, and the point five months later when manufacturers had scaled to meet national demand. The authorization came seven months before the first mRNA vaccines.

- About ten percent: the reduction in relative risk of death from remdesivir in hospitalized patients, in one large-scale clinical trial.
- More than 600,000: viruses estimated to be capable of infecting humans, most still unknown.
- Ten times fewer infections: in cells covered with a one percent porcine gastric mucin solution, compared with a standard buffer (MIT, 2012).
- Ten viral families in cells and three in mice: the breadth shown for DRACOs and dsCAREs (MIT Lincoln Laboratory and Boston University, 2011), a count that includes two later publications and a public presentation.
- 10 to 100: virus particles typically produced by one infected cell.
- More than a million times: the reduction in peak viral load from antivirals given before SARS-CoV-2 infection versus three days after symptoms, in a 2020 computational model.
- Half: the share of remdesivir molecules cleared from the bloodstream every day; its active metabolite GS-441524 has a half-life of one day; prophylactic use would mean infusions of up to two hours.
- Eight hours: how long PCANS lingered in the nasal cavities of mice after each dose, and also how long before infection SHIELD was inhaled in the monkey study.

- Two weeks: the daily dosing period in the mouse safety studies for both PCANS and SHIELD, with no significant side effects or toxicity.
- About 50 times in the lungs and 300 times in the nose: the reduction in SARS-CoV-2 viral RNA in monkeys given SHIELD.
- One to two weeks: how far ahead of infection animals were dosed in previous DNA-encoded antiviral experiments, which Wang says is too short a test of prophylaxis.
- 5,000 treatment courses, scaled to 500,000 over nine months: the world's remdesivir supply at the start of the pandemic and Gilead's scale-up.
- About 500,000 off-label prescriptions in a few weeks: hydroxychloroquine's reach, because it already had approved indications.

What the author is unsure about

Wang is explicit about the limits of his own case. Preclinical data does not guarantee clinical efficacy, and he uses remdesivir as the warning. The breadth results for DRACOs and dsCAREs are in cells and mice. The PCANS and SHIELD results are in mice and monkeys. The million-fold figure for early prophylaxis comes from a computational model, not a trial. He says his own goal at Panoplia may require going well beyond existing research, that the selectivity needed may be a narrow tightrope to walk, and that the market niche he is aiming at may be taken by

competitors. He also concedes that drug-free and repurposed routes avoid the human trials that would tell us whether the products actually work. He discloses that he is co-founder and Executive Director of Panoplia Laboratories. The piece was published on 26 August 2024.

For the discussion

- Wang's strongest point is that the binding constraint on a day zero drug is the shelf, not the laboratory: remdesivir existed, was authorized seven months before any vaccine, and still could not be made in quantity because it had no market before the pandemic.
- The innate immune system's three properties give a concrete test to apply to any proposed countermeasure in this unit: does it target something all viruses must do, does it work before infection, and is it already everywhere?
- The drug-free and repurposing workarounds get products onto shelves by stepping outside the approval system, which means we gain availability and lose the evidence that the thing works, and that trade is worth naming out loud alongside the course's framing about testing before deployment.

5.1-5 · CORE · COURSE TIME 45 MIN

Accelerating Vaccine Development: The 100 Days Mission

Aishani Aatresh, Nicole Lurie, and Richard Hatchett · 2024

Read the original paper → · the course assigns: whole chapter, via a structured section-by-section account rather than the original text

> **CEPI's case that a vaccine can be ready for initial authorization within 100 days of spotting a new pandemic pathogen, and that getting there needs a change in how preparedness and response relate to each other rather than more of the same preparedness.**

Average summary

Summarized from a structured, section-by-section account of the chapter written by another helper while reading it, not from the original text, so nothing here is a direct quotation. The chapter's five figures, all credited to CEPI, were not captured.

Aatresh, Lurie and Hatchett set out the **100 Days Mission**: vaccines ready for initial emergency authorization and manufacturing at scale within 100 days of recognizing a pathogen with pandemic potential. The benchmark is COVID-19, where the first vaccine was authorized a record 326 days after the genetic sequence was posted, by which time more than 60 million people had been infected and over 1.5

million had died. The authors argue that speed rested on decades of prior work on betacoronaviruses and mRNA, and that a CEPI analysis of COVID-19 innovations suggests using every available innovation next time would cut the timeline by at most about 25 percent, to roughly 250 days. Getting to 100 therefore needs a different approach, not a faster version of the same one.

CEPI was founded after the 2014-2016 West African Ebola epidemic, on the reasoning that vaccines for dangerous but rare pathogens should be developed in advance through Phase II and then tested further during an outbreak. The chapter structures the mission around five areas: prototype vaccines for representative pathogens across viral families; biomarkers for protective immune responses; global capability to characterize pathogens and outbreaks early; rapid manufacture and validation of experimental vaccines; and ready clinical trial infrastructure. It notes that 26 families of viruses are known to infect humans, which is what makes a library of prototype vaccines a tractable idea.

The chapter's organising claim is that outbreak response is outbreak preparedness. Smaller outbreaks build the infrastructure, give practice, reveal gaps and generate data. Its central example of the gap that remains is the 2022 Sudan ebolavirus outbreak in Uganda: investigational vaccine doses

existed in bulk, but 79 days passed from the outbreak declaration before they were filled, finished and delivered, and the outbreak ended before a trial could start, with an incomplete protocol and missing in-country regulatory authorization also in the way.

Three further sections carry the argument. Manufacturing needs a global network of warm facilities that make vaccines routinely and can switch within days or weeks, because factories cannot be woken from dormancy and need market demand between epidemics. Clinical trials need a sustainable global network, a clinical laboratory network and national volunteer registries, with standards set in peacetime. And equity has to be designed in from the start, which CEPI calls systems equity, on the view that scarcity is the enemy of equity.

Why it's assigned

This is the chapter that names the target the whole unit argues about. The Combat layer's premise is that the interval between identifying a pathogen and protecting people decides how many die, and the 100 Days Mission turns that into a number with five named workstreams. Each of the four deep-dive options on pages 5.3 to 5.6 is one of those workstreams: faster R&D, trial innovation, financing and procurement, manufacturing and delivery. Read it as the map of the day, and note that its authors are the people running CEPI, so the chapter is an

argument from an interested party as well as a technical account.

In depth

What CEPI is, and why research belongs inside a response

CEPI was created after the 2014-2016 West African Ebola epidemic, when a vaccine that had reached early development and then been neglected was revived, tested and found safe and effective. Its founders reasoned the world would be better protected against dangerous but rare pathogens if vaccines were developed in advance through Phase II trials and then evaluated further during an outbreak. As a coalition of public, private, philanthropic and civil society organisations, it works on vaccines whose market potential is too small to create commercial incentive, starting from five priority pathogens in the WHO R&D Blueprint: MERS coronavirus, Nipah virus, chikungunya virus, Lassa virus and Rift Valley fever virus.

A founding tenet is that research is part of response, not only of preparedness. CEPI was barely three years old when WHO declared COVID-19 a pandemic in March 2020. Building on its MERS-CoV and rapid response platform programmes, it supported vaccine development, clinical trials and syntheses of real-world evidence, and its development efforts

contributed to the authorization of **eight vaccines**. Recognising that R&D alone could not quell a pandemic, it joined WHO, UNICEF and Gavi to form COVAX, the vaccines pillar of the Access to COVID-19 Tools Accelerator. It also supported studies establishing correlates of protection and immunobridging approaches to regulatory authorization, and funded later research on fractional dosing, mixing doses from more than one producer, efficacy in immunocompromised people, cold chain innovations and broadly protective coronavirus vaccines.

The 2022-2023 outbreaks show how much remains. Sudan ebolavirus appeared in Uganda in late 2022 and Marburg virus in Equatorial Guinea in early 2023. Investigational Sudan virus vaccine doses existed in bulk form, but 79 days passed from the declaration of the outbreak before they were filled, finished and delivered to Uganda. That was faster than previous responses, and still the outbreak ended before a trial could launch. An incomplete protocol and missing in-country regulatory authorization were further barriers. The lesson the authors draw is that preparedness and response are a continuum rather than two states.

The mission, and the arithmetic behind it

Developing and authorizing novel SARS-CoV-2 vaccines in under a year was, in the chapter's judgement, a technoscientific triumph inside a global

response that was inadequate in many other respects. The first vaccine received emergency authorization in a record 326 days from the date the SARS-CoV-2 genetic sequence was posted. By then more than 60 million people had been infected and over 1.5 million had died.

That speed rested on decades of prior research into betacoronaviruses and mRNA platforms, the experience of national biomedical institutions, and established pharmaceutical partnerships. The figure that justifies the mission is a CEPI analysis of COVID-19 vaccine development innovations: using every possible innovation in the next pandemic might cut the time to authorization by at most about 25 percent, to roughly 250 days. Incremental improvement does not get you to 100.

As CEPI articulated it in 2021, the mission is that vaccines should be ready for initial authorization and manufacturing at scale within 100 days of recognition of a pathogen with pandemic potential, when appropriate. From a research perspective it is structured around five areas of the R&D life cycle: 1. Prototype vaccines for representative pathogens across viral families. 2. Biomarkers for protective and robust immune responses. 3. Global capabilities for early characterization of pathogens and outbreaks. A footnote says surveillance and pathogen identification lie beyond the chapter's scope and outside CEPI's

direct purview, though they are a key enabler. 4. Rapid manufacture and validation of experimental vaccines. 5. Ready clinical trial infrastructure that can spring into action to test experimental vaccines.

The mission has been taken up by the G7, the G20 and the US government. The authors are careful about what the number means. A footnote records that other definitions start the clock differently or aim at a different endpoint, for instance 100 days from an emergency declaration rather than from identification of a pathogen with pandemic potential, and that CEPI's 100 days are intended only to reach initial authorization for vaccines, not immediate widespread use and not other medical countermeasures whose constraints differ.

Two elements make this a paradigm shift rather than a call for more preparedness. First, what accelerates development is transformation in a series of overlapping processes, not merely the existence of a vaccine candidate. Second, false alarms, outbreaks that fizzle out, and the data they generate are motivators, practice runs and accelerators in their own right.

Rapid response platforms and vaccine libraries

A rapid response platform is a system or technology using the same basic components, adaptable to different pathogens by inserting new genetic or

protein sequences. The familiar post-COVID example is the mRNA platform used by Pfizer-BioNTech and Moderna, but other platforms and manufacturing processes can be readied too.

The chapter states that **26 families of viruses** are known to infect humans. Developing prototype vaccines against representative pathogens from the families with greatest pandemic potential would expand the toolkit for known pathogens and improve the response to a novel pathogen from the same family. A library of prototype vaccines on selected platforms builds platform-specific experience into which a new sequence can be inserted, and tested prototypes might even be useful as they are if laboratory and animal data suggest enough cross-protection. Such advances could let preclinical work, initial human safety trials and manufacturing scale-up run in parallel, with the degree of simultaneity governed by experience and judgement.

They also argue the intermediate steps have value of their own: assembling the technical repertoire and expressing candidates as plasmids are milestones that convene experts. And as regulators gain experience with a platform, they will likely become more comfortable moving new vaccines or immunoprophylactics, such as monoclonal antibodies and antibody cocktails, into trials with adapted

requirements, by analogy with the annual approval of seasonal influenza vaccines.

The challenges named are substantial. Many constructs from a prototype or pan-family approach cannot be assessed for real-world effectiveness without ongoing transmission, and rigorous trials when they are feasible need well-coordinated mobilisation during an emerging outbreak, which has often fallen short. The investments of time, money and expertise are considerable. At least some vaccines developed will never be used, and there are few commercial incentives to develop vaccines against diseases that may never become serious concerns, which adds political complexity. Intensive engagement with regulators is needed to establish where platform data can accelerate development. Prioritisation across and within viral families is unavoidable and must be decided under what the chapter calls insurmountable uncertainty. And ownership and sharing provisions for information, data and constructs may be highly contested.

Scientific needs, manufacturing and trials

On science, well-characterized assays, biomarkers and correlates of protection are integral to prototype vaccines and the platforms they sit on. CEPI's ambition is a globally distributed scientific infrastructure able to collect and characterize biologic samples early in an outbreak, particularly to develop

animal models and to assess immune responses, paired with upstream work such as using artificial intelligence to optimise antigen design before preclinical testing. It continues to develop an animal model network and a clinical laboratory network, intended both to support development and to build capacity in regions historically disadvantaged in these domains.

On manufacturing, the proposal is a global network of **warm facilities**: plants that routinely produce vaccines for national or regional use and can switch to emergency production on short notice, changing output within days or weeks. Sustaining it means funding creation and maintenance across multiple platforms including mRNA, viral vector and protein subunit. The authors stress two constraints that recur in the manufacturing deep dive on page 5.6: factories cannot be activated quickly from dormancy and need market demand between epidemics to stay viable, and buildings are of limited value without skilled personnel, quality systems and supply chain resilience. CEPI is supporting a network of go-to facilities that produce vaccines regularly and can be called on in an emergency.

On clinical trials, the authors argue that experience with betacoronaviruses points to a possible new regulatory approach. Knowledge from SARS-CoV-1 and MERS-CoV candidates accelerated SARS-CoV-2

development, and regulatory guidance across jurisdictions during the pandemic enabled rapid adaptation of vaccines to emerging variants. With further experience, regulators could plausibly approve initial, carefully monitored use of a vaccine against an emerging pathogen on the basis of fewer but extremely well-supported assessments, as seasonal influenza vaccines are approved now.

That depends on trial infrastructure existing first. CEPI has identified three foundational initiatives: a global clinical trial network sustainable between pandemics; a global clinical laboratory network for faster data readout; and national clinical trial volunteer registries to accelerate enrolment. These rest on standards set in peacetime and refined through smaller crises: accessible master files that reduce administrative overhead, alignment of regulatory standards across jurisdictions, and coordinated global networks able to assess vaccines in and for the populations that need them. In silico models can help predict candidate success and inform dosing choices.

The obstacles are named as frankly as the ambitions: political, ethical and operational barriers to harmonising systems across sites and countries; the need for business models that keep sites and laboratories active between pandemics; governance risks in national registries; personal data and

confidentiality law; historically justified anxieties about studies in people that test novel biomedical technologies against undefined threats; and limited trial registration within the sub-populations most vulnerable to the diseases of concern. The authors argue that shared awareness of essential protections for patients, participants, communities and citizens is what keeps standards from being compromised in the name of acceleration.

The enabling ecosystem, and knowing when to move

The chapter concedes that its focus is vaccines while equitable and fast vaccine development is impossible without infrastructure outside CEPI's mandate. Its list runs from improved biosurveillance and stronger information-sharing norms, through early epidemiological and clinical studies to establish parameters such as R0 and case fatality rate and track how they change, swift use of non-pharmaceutical interventions, faster diagnostics and therapeutics, medical technologies suited to differently resourced settings, stronger delivery capacity including cold chain plus platforms robust to temperature variation, engagement with communities and governments, regulatory standards fitted to a countermeasure's intended use, and mechanisms for monitoring safety and efficacy during rollout. It notes that even a primary health care system where sick

people can be diagnosed and eventually treated is essential yet far from universal.

The activation question gets its own treatment. No amount of preparedness becomes response without clarity about when and how to mobilise. Mapping which early signals would trigger pre-identified **no-regrets levers** across partners lets a coordinated research response begin in time. No-regrets actions are those an organisation or country can and will take once a condition is met even if the outbreak never becomes a Public Health Emergency of International Concern: mobilising funding, collecting samples, activating trial and manufacturing networks, deploying personnel. The authors argue that a research-based response is part of what justifies acting at risk, because the action improves long-term preparedness even where the payoff in one instance is uncertain. Pre-determined off-ramps matter equally, so efforts stay keyed to risk tolerance.

The section closes on the meaning of the number itself. The mission is less about the exact count of days than about aiming at much faster development when an outbreak has pandemic potential, and coalescing around a goal starts a clock on every aspect of a response. The authors caution that every research response differs and preparedness cannot be dictated too rigidly by the past, citing the epidemiologists' adage that if you have seen one

epidemic, you have seen one epidemic. Some assumptions will prove wanting and there will be oversights, both unforeseeable and culpable; what we can do is build basic working checklists, update them after each outbreak, and act on cumulative experience.

Systems equity

Stark global inequity in vaccine distribution was a defining feature of COVID-19. Despite COVAX and the ACT Accelerator, many low- and middle-income countries lacked adequate access when they needed it, while many high-income countries bought a vast share of future doses before those countries had begun procurement conversations. The most coveted mRNA vaccines were hard to transport and use where cold chain was limited, and variants plus limited supply sometimes produced export controls that sequestered vaccines further. By mid to late 2023 the position had inverted: a surplus of doses and of funding to buy them, with very limited demand.

Against the conventional view that speed and scale sit at odds with equity, and that a moonshot means a high-tech race among wealthy countries for their own vaccines, the authors argue that for CEPI speed, scale and equity are inseparable, because **scarcity is the enemy of equity**. They acknowledge that a national government's first responsibility is to its own citizens and that few politicians would be faulted for meeting

it. Their answer is to configure the value chain from the outset so that equity is an output rather than an afterthought, which no single actor can do alone.

The factors they identify as historically delaying development and so worsening access are coordination and clarity of roles; research capacity in outbreak settings; established financing mechanisms for R&D, at-risk manufacturing and procurement so that low- and middle-income countries are not last in line; and the operational infrastructure, legal provisions and political agreements that enable global access.

In practice, CEPI has advanced the first-ever Phase III trial for a chikungunya vaccine candidate and the first in-human Phase I trials for a Nipah vaccine candidate, with several Lassa fever candidates in advanced clinical development. Its Enable research programme assessing the burden of Lassa fever across West Africa is designed to inform late-stage trial design while strengthening local research capacity, so that safety and efficacy are evaluated in the populations regularly affected rather than in relatively high-income travellers. Partners have been funded to advance more thermostable mRNA platforms, and manufacturing partnerships with L'Institut Pasteur de Dakar and Aspen are described as foundations of CEPI's global manufacturing

network and as crucial to increasing the share of vaccines manufactured in Africa, for Africa.

International coordination, and what research cannot settle

Research alone cannot deliver the mission. The chapter names four questions that remain political and economic: how to ensure vaccines developed in these new contexts are sold at affordable prices without price gouging or wealthy countries buying even larger shares; what is required for technological innovations to be shared broadly and faster; how to account for variation within and between regions; and how access and benefit sharing provisions, or alternatives, will be developed, negotiated and enforced in multichannel markets for doses and funding.

These exceed any one organisation. The fora where they are being negotiated are listed as the Intergovernmental Negotiating Body drafting a pandemic instrument under the WHO Constitution, the G20, the G7, the Johannesburg Process for a Medical Countermeasures Network or Platform, and the September 2023 United Nations High-Level Meeting on Pandemic Prevention, Preparedness and Response. The authors argue that growing local capacity matters but so does coordination, so that these efforts serve global public health rather than

aggravating the protectionist fragmentation that infectious disease emergencies tend to produce.

The conclusion restates the frame: emergency vaccine research as a form of preparedness, with preparedness and response each advancing the other, and with consensus on what research should be done built by engaging researchers, policymakers and the public about community and country needs, and not solely in terms of vaccine hesitancy. Pandemics are global events experienced differently around the world, and the goal is to keep the next one from growing beyond a small outbreak.

Words to know

- **100 Days Mission** CEPI's target of having a vaccine ready for initial authorization and scaled manufacturing within 100 days of recognizing a pathogen with pandemic potential. The clock starts at identification, not at an emergency declaration.
- **Rapid response platform** A vaccine technology whose components stay the same while a new genetic or protein sequence is swapped in. It is what makes the 100-day clock imaginable at all.
- **Prototype pathogen** A representative member of a viral family, worked on so that the knowledge transfers to its relatives. The chapter counts 26 families of viruses known to infect humans.

- **Vaccine library** A stock of prototype vaccines built on chosen platforms, ready to receive a new sequence, and sometimes useful as they are if cross-protection is good enough.
- **Immunobridging** Using immune response data to support authorization rather than a full efficacy trial. CEPI supported studies establishing these approaches during COVID-19.
- **Warm facility** A manufacturing plant that produces vaccines routinely and can switch to emergency output within days or weeks. The point is that a dormant factory cannot be restarted quickly.
- **No-regrets lever** An action an organisation commits to taking once a trigger condition is met, even if the outbreak never becomes a formal emergency, such as releasing funding or activating a trial network.
- **Systems equity** CEPI's term for designing the research, development and delivery chain so that equitable access is a natural output rather than an afterthought.

Numbers in this reading

- 100 days: the target, from recognition of a pathogen with pandemic potential to initial authorization and manufacturing at scale.
- 326 days: the record time from posting of the SARS-CoV-2 genetic sequence to the first emergency authorization.

- More than 60 million infected and over 1.5 million dead: the global toll by the time that authorization came.
- About 25 percent, to roughly 250 days: the most a CEPI analysis suggests every available COVID-19 innovation could cut from the timeline next time.
- 2014-2016: the West African Ebola epidemic after which CEPI was created.
- Five priority pathogens at the start: MERS coronavirus, Nipah, chikungunya, Lassa and Rift Valley fever.
- Eight vaccines: those whose authorization CEPI's development efforts contributed to during COVID-19.
- 79 days: from the declaration of the 2022 Sudan ebolavirus outbreak in Uganda to investigational doses being filled, finished and delivered. The outbreak ended before a trial could launch.
- 26: families of viruses known to infect humans.
- Five areas of the R&D life cycle, and three foundational clinical trial initiatives.
- First-ever Phase III for a chikungunya vaccine candidate, and first in-human Phase I trials for a Nipah candidate, both advanced by CEPI.

What the author is unsure about

The authors write as CEPI's chief executive, a CEPI colleague and an academic collaborator, so the

chapter argues for CEPI's own programme; read the claims about its achievements with that in mind. Within that, they are candid. They say the 100 days are a direction more than an exact count, and footnote that other definitions start the clock elsewhere and that the target covers initial authorization only, not widespread use or other countermeasures. They say prioritisation among viral families must be made under insurmountable uncertainty, that at least some vaccines developed will never be used, that many prototype constructs cannot be assessed for real-world effectiveness without ongoing transmission, and that ownership and data-sharing provisions may be highly contested. They concede that trial and manufacturing networks need business models nobody has yet settled, that anxieties about research in people are historically justified, and that the political and economic questions in Section 4 lie beyond what any organisation can answer. They also admit that assumptions will prove wrong and oversights will happen, some of them culpable.

For the discussion

- The strongest single number in the chapter is 79 days: doses existed in bulk for Sudan ebolavirus and the outbreak still ended before a trial could start, which says the binding constraint was fill-

finish, protocol and in-country regulatory authorization rather than science.

- The authors' claim that about 25 percent is the most that incremental innovation can deliver is what turns the 100 Days Mission from a stretch goal into an argument for restructuring, and it is worth asking whether the four deep-dive options together actually close the remaining gap.
- Systems equity rests on the claim that scarcity is the enemy of equity, which is a strong and testable proposition: it predicts that expanding capacity does more for fairness than reallocating a fixed supply, and the financing readings on page 5.5 give the numbers to check it against.

5.1-6 · CORE · COURSE TIME 20 MIN

The clinical trial bottlenecks game

Saloni Dattani · 2026

Read the original paper → · the course assigns: the playable site's walkthrough, conclusion, methodology and reference values. Parts three and four stay locked until you reach the one-year target yourself

> **A playable model of a phase three vaccine trial that sets you the task of cutting a two-year timeline to one year, and whose point is that no single lever gets you there, because fixing one bottleneck simply reveals the next.**

Average summary

Summarized from a capture of the site made on 22 October 2026, covering part one's walkthrough with all six levers, part five's conclusion, the methodology and the downloadable reference values. Part two, the budget game itself, asks for a nickname that is posted to a public leaderboard, so nothing was entered. Parts three, Explore the bottlenecks, and four, Bottleneck river, open only once a player has reached the one-year target themselves, so they are waiting for her to play. The numbers below are the values the game starts from.

Dattani frames it with two claims in tension. Once a drug or vaccine enters human testing it takes an average of **8 to 10 years** to pass through all stages of trials, review and approval, and only if it succeeds at all, so trials are often called a bottleneck to medical progress. Meanwhile others expect rapid AI progress to make all or most diseases curable within ten years. The game is built to test who is right.

The model is a simplified phase three trial, adapted from real HPV vaccine trials. The reference case is the Gardasil FUTURE II trial: about 12,000 participants, about 8 percent recruited each month, true efficacy of 98 percent, an outcome of serious cervical cell changes, and a modeled duration of about two years. The opening state shows 21.1 months.

Six levers can be moved. Sample size, from 500 to 25,000, starting at 12,200. Recruitment rate, 2 to 30 percent of the target per month, starting at 8.3. Disease incidence, 0.05 to 1.5 events per 1,000 participant-months, starting at 0.25. Loss to follow-up, 0 to 5 percent a month, starting at 0.23. True vaccine efficacy, 50 to 100 percent, starting at 98. And the outcome measured, chosen from persistent HPV infection, mild cervical cell changes, serious cervical cell changes (the default) and cervical cancer. Every lever has diminishing returns, and the walkthrough's conclusion is headed that one change is not enough: recruit much faster and you still wait for enough disease cases; have an extremely effective vaccine and you still need enough cases in the control group to prove it.

The game then adds costs, because in the real world you could not give every trial 20,000 participants, instant recruitment and no dropout. The budget version gives 100 coins to spend across the levers. Dattani is open that data on the costs of trial interventions is very limited, so the cost curves make the trade-offs playable without implying the interventions have known prices.

The conclusion is that calling any single feature the bottleneck is misleading. Speeding up trial timelines substantially takes several levers at once, their benefits do not simply add up, and improving the

efficacy of the drugs in the pipeline shifts the bottleneck onto how long diseases take to develop and onto funding, infrastructure and regulation. On AI, she argues the other constraints will not disappear if drug development gets much faster; they will get stronger.

Why it's assigned

The course note tells her to play the game and read the conclusion, and to notice which constraints cannot be fixed at the same time. In the Combat layer, the premise is that the gap between designing a countermeasure and knowing whether it works is where the deaths accumulate, and this is the only reading in the unit that lets her feel the shape of that gap rather than read about it. It is also the direct on-ramp to her chosen deep dive on page 5.4: the levers here are precisely the variables the clinical trials innovation readings propose to change.

In depth

The question the game is built to settle

The framing is a disagreement. Once a drug or vaccine enters human testing it takes an average of 8 to 10 years to get through all stages of clinical trials, review and approval and become available, and only if it succeeds. That is why trials are so often named as the bottleneck to medical progress. On the other side are people expecting rapid AI progress to make all or

most diseases curable within ten years. The game exists to test which picture holds up once you have to move the parts yourself.

The task is concrete: the model is of a **phase three trial**, the last stage before approval, and the player's goal is to cut the modeled timeline from about two years to one year by adjusting a set of levers.

The case, and the real trial underneath it

The case is human papillomavirus vaccines. HPV spreads through sexual contact and can cause cervical, penile, anal, oral, and head and neck cancers. HPV vaccines are highly effective against the types they target, and their trials still take years, which is what makes them a good teaching case.

The reference trial is the phase three Gardasil FUTURE II trial, simplified. As the site states it: - Sample size: about 12,000 participants - Recruitment: about 8 percent recruited each month - True vaccine efficacy: 98 percent - Outcome: serious cervical cell changes - Modeled trial duration: about 2 years

Roughly 12,000 women took part in FUTURE II. Randomising participants to vaccine or placebo, it tested whether the vaccine reduced the risk of serious changes in cervical cells linked to HPV-16 or HPV-18, two of the most harmful types, and found a 98 percent risk reduction. The site's opening default state shows a modeled duration of 21.1 months, and the

methodology's own figure is a 21-month modeled reference.

The six levers, with their ranges and starting values

Each lever has a slider, a plain-language explanation, a timeline bar scaled from 0 to 5 years, an expander offering ways to improve it, and a note on diminishing returns.

1. **Sample size.** The number of participants. Range 500 to 25,000; starting value 12,200. All else equal a larger trial finishes faster, because detecting the difference a vaccine makes requires enough disease in the control group. With too few participants, few controls develop the disease at all, and it is hard to tell whether the vaccine lowered the number further. The improvement slows as larger samples save less additional time.
2. **Recruitment rate.** The percentage of total participants who join per month. Range 2 to 30; starting value 8.3. Real trials do not fill instantly and can take months to reach the target number. Recruitment speed also faces diminishing returns.
3. **Disease incidence.** How often the disease endpoint occurs in the trial population. Range 0.05 to 1.5 per 1,000 monthly; starting value 0.25. You cannot change how often a disease occurs, but you can choose where to run the trial and whom to recruit, and different populations face different

risks. Finding a higher disease rate speeds the trial substantially at first, because cases accumulate much faster.

4. **Loss to follow-up.** The share lost from usable follow-up each month. Range 0 to 5 percent; starting value 0.23. Participants drop out or stop being tested, because they move, lose interest, find visits burdensome or stop responding, and if they do, whether they developed the disease cannot be known, so the trial takes longer. High dropout rates become increasingly costly.
5. **Vaccine efficacy.** The true reduction in outcome risk. Range 50 to 100 percent; starting value 98. This is the vaccine's true underlying efficacy, which researchers do not know in advance. The trial's statistical rule is to wait for enough disease cases to show whether efficacy exceeds 50 percent, so higher true efficacy makes a larger gap between vaccine and control groups and brings a confident result sooner.
6. **Outcome measured.** The endpoint, chosen from four: persistent HPV infection; mild cervical cell changes; serious cervical cell changes, the default, described as abnormal cervical cells linked to HPV-16 or HPV-18, CIN2 or higher grades, and the primary outcome in FUTURE II; and cervical cancer. The reason to vaccinate against HPV is to prevent cancers, but cancer usually takes many years to develop, so waiting for it is far slower.

> Earlier outcomes occur more often and can be confirmed sooner, producing a result much faster, while later outcomes reflect the disease we actually care about and can add years. The choice depends partly on confidence that the earlier outcome predicts the later one; for HPV the chain runs from an infection that does not clear, to mild cell changes, to serious cell changes, to cervical cancer.

One change is not enough

The walkthrough's conclusion carries that heading, and it is the game's central teaching. Changing any one lever makes the trial faster, but diminishing returns set in and another bottleneck appears. Recruiting much faster still leaves you waiting for enough HPV disease cases to confirm the result. An extremely effective vaccine still needs enough cases in the control group to prove effectiveness. The walkthrough then hands over all six levers together, with a button to see how long the trial takes and a control to reset to the starting values.

Why the levers have prices

Under the heading that changing the levers has costs, the site points out that every lever could simply be maxed out, but in the real world it would be very expensive for every trial to have 20,000 participants, recruit extremely quickly and have no loss to follow-

up. So it asks what happens if the levers have costs, and if further improvements cost more, since the last participants recruited are probably the most expensive. It then concedes that data on the costs of clinical trial interventions is very limited: diminishing returns can be assumed, but specific cost-curve data is lacking, which it says is no reason not to turn it into a game.

The parts that are waiting for her

Part two is the budget game proper. It opens by asking what to call you, with a nickname field, and says the nickname will appear on the leaderboard if you design a successful trial. From the methodology, the game gives **100 coins** to spend across the levers, with cost curves attached.

Part three, Explore the bottlenecks, and part four, Bottleneck river, did not open. The methodology explains why: successful combinations are stored in a shared dataset and shown to later users only after those users have reached the one-year target themselves. So those two sections unlock when she plays it, and seeing what other people's successful trials looked like is the reward for getting there.

The model's rules

Participants are recruited gradually and assigned equally to vaccine and control groups. Disease events accumulate over time while some usable follow-up is

lost. The timeline ends when enough events have accumulated to meet the trial's statistical rule.

The statistical rule is that the model tests whether the user-selected true efficacy is greater than 50 percent, at 90 percent power and a two-sided 5 percent significance level. At the reference 98 percent efficacy this needs about **13 expected endpoint events**. That single number is worth sitting with: a two-year trial of 12,000 people turns on thirteen cases.

Fixed assumptions: constant disease rate over time; linear recruitment at a constant share of participants per month; participants lost to follow-up are not replaced; follow-up begins at month seven for each participant.

On coins, the budget game uses 100 for illustration rather than as an estimate of real financial cost, for which there is little data. Most levers share the same diminishing-returns cost curve, scaled per lever. The exception is vaccine efficacy, which uses a technological S curve. For disease incidence the curve runs from the selected outcome's starting rate up to an illustrative 1.5 events per 1,000 participants per month, and spending on incidence stands for the research and trial infrastructure needed to identify and recruit from higher-incidence places or populations. The site states that the cost curves make

the trade-offs playable but do not imply the interventions have known or comparable prices.

Where the reference values come from

Most come from the ClinicalTrials.gov record for NCT00092534, the FUTURE II trial of Gardasil. Its primary publication reports a recruitment window from June 2002 to May 2003, average follow-up of three years, control-group disease incidence of 0.3 per 100 people per year, and 98 percent efficacy against serious cervical cell changes linked to HPV-16 or HPV-18. The site uses 98 percent as the true efficacy. Unlike the original trial, and for simplicity, it applies the same test, efficacy higher than 50 percent at 90 percent power and two-sided 5 percent significance, to every outcome.

The conclusion

The range of successful combinations is the point: it is misleading to call any single feature the bottleneck. When recruitment is slow, speeding it up has a lot of value; once recruitment is fast, other parts of the trial become the bottleneck instead. Even very high efficacy may not be enough on its own.

Substantially speeding up trial timelines therefore takes several levers at once. Their benefits do not simply add up, because the levers interact, and several of them hit diminishing returns too.

Improving the efficacy of the drugs and vaccines in the pipeline really can speed the timeline a lot. But that depends on experimental testing and iteration through clinical trials, and it shifts the bottleneck onto other factors: how long diseases take to develop, and the funding, infrastructure and regulation around trials.

On AI, Dattani sees a lot of room for it to be useful. Beyond drug design and discovery, it could help with participant recruitment, building evidence for surrogate endpoints, better remote testing and more. Her closing argument is the one to bring to the discussion: if AI makes drug development much faster, the other constraints will not disappear, they will become stronger, and we should be thinking now about how to solve them too.

Limitations, in the site's own terms

The model is not a forecast for a proposed clinical trial. It does not model changes in disease incidence over time, geographic variation, adherence, regulatory review, manufacturing, costs, safety outcomes, or ethical constraints on trial design. Results depend on the selected assumptions and should be read as comparisons within the model.

On data: successful combinations are stored in a shared dataset so they can be shown to later users once those users reach the one-year target, and the

chosen nickname, coins spent, model parameters and calculated duration are stored for the chart and leaderboards.

The site's footer credits it as written and designed by Saloni Dattani using GPT 5.6, last updated September 2026, with thanks to Matt Clancy, Nisha Austin, Ruxandra Teslo, Julia Rohrer, Ruben Arslan, Malin Benjes and Simon Grimm for feedback.

Words to know

- **Phase three trial** The large, final stage of testing before approval, which measures whether the product actually prevents disease. It is the stage the game models, and the one that eats the timeline.
- **Endpoint** The specific event a trial counts, such as a persistent infection or a cancer. Choosing an earlier endpoint makes a trial faster and makes its result a step further from the thing you care about.
- **Incidence** How often the endpoint occurs in the trial population, here measured in events per 1,000 participant-months. You cannot change a disease's rate, only where and among whom you run the trial.
- **Loss to follow-up** Participants who stop being tracked, so you never learn whether they got the disease. Every one of them makes the trial longer.
- **Power** The chance a trial detects a real effect if there is one. The game fixes it at 90 percent, which

is part of why the event count is what it is.

- **Diminishing returns** Each further improvement in a lever buying less time than the one before. It is the mechanism that makes a single-lever strategy fail.
- **Surrogate endpoint** An earlier, more frequent marker used in place of the outcome you ultimately care about. Dattani names building evidence for surrogate endpoints as one of the clearest places AI could help.

Numbers in this reading

- 8 to 10 years: average time from entering human testing to approval, if a drug or vaccine succeeds at all.
- Two years down to one year: the game's task.
- 21.1 months: the modeled duration in the site's opening default state. The methodology's reference figure is 21 months.
- Sample size: range 500 to 25,000; starting value 12,200.
- Recruitment rate: range 2 to 30 percent of target per month; starting value 8.3.
- Disease incidence: range 0.05 to 1.5 events per 1,000 participant-months; starting value 0.25.
- Loss to follow-up: range 0 to 5 percent per month; starting value 0.23.

- True vaccine efficacy: range 50 to 100 percent; starting value 98.
- Four outcome options: persistent HPV infection, mild cervical cell changes, serious cervical cell changes (default), cervical cancer.
- Statistical rule: efficacy greater than 50 percent, 90 percent power, two-sided 5 percent significance, requiring about 13 expected endpoint events at 98 percent efficacy.
- Follow-up begins at month seven after a participant's first dose.
- 100 coins: the budget in part two, for illustration rather than real cost.
- Maximum simulated timeline 60 months; success target 12 months.
- FUTURE II reference values as published: trial FUTURE II V501-015, ClinicalTrials.gov NCT00092534; 12,167 participants, 6,087 allocated to vaccine and 6,080 to placebo; recruitment period 12 months at 8.33 percent of target per month; average published follow-up 36 months; registered primary completion date 31 July 2007; 10,903 participants completed base-study follow-up; monthly follow-up loss 0.23 percent of remaining usable follow-up; primary endpoint HPV-16/18-related CIN2 or worse; annual placebo incidence 0.3 participants with the endpoint per 100 person-years, which is 0.25 events per 1,000

participant-months; true vaccine efficacy 98 percent.

- Illustrative incidence rates for the other outcomes: persistent HPV infection 1 event per 1,000 participant-months; mild cervical cell changes 0.29; cervical cancer 0.03, the last taken from the WHO cervical cancer fact sheet.
- Simplified modelling choices: 1:1 allocation, linear recruitment, constant incidence, no replacement after follow-up loss, zero analysis delay.

What the author is unsure about

Dattani is unusually explicit. The model is not a forecast for any proposed trial. It leaves out changes in incidence over time, geographic variation, adherence, regulatory review, manufacturing, costs, safety outcomes and ethical constraints on trial design. Results depend on the chosen assumptions and are only comparisons within the model. She says outright that data on the costs of clinical trial interventions is very limited, so the coin costs are illustrative and do not imply the interventions have known or comparable prices. The efficacy threshold, power and significance level are labelled illustrative model decisions, and so are the interface's maximum values and the 12-month success target. She also simplifies away from the real trial by applying the same statistical test to every outcome. On AI her position is an argument, not a finding: there is room

for it to help, and the other constraints will intensify rather than vanish.

For the discussion

- The game's sharpest fact is that the reference trial's two-year timeline turns on about 13 expected endpoint events, which reframes trial length as a problem of counting rare events rather than of paperwork.
- The outcome lever is the real ethical choice in the model: moving from cervical cancer to persistent infection buys years, and spends them on confidence that the early marker predicts the later disease.
- Dattani's closing claim deserves pressure in the discussion, because it inverts the usual optimism: if AI speeds up drug design, the binding constraints become disease biology, funding, infrastructure and regulation, which are exactly the things the rest of Unit 5 is about.

5.4-1 · CORE

From Warp Speed to 100 Days

Witold Więcek · 2023

Read the original paper → · the course assigns: whole piece, including its seven footnotes

> **We learned to design a COVID vaccine in 65 days and then spent nine months proving it worked, so the route from 326 days to 100 lies in clinical trial design and the incentives around it, not in the laboratory.**

Average summary

Więcek splits the COVID vaccine timeline in two and shows where the time actually went. It took **326 days** to create, test and authorize the first highly effective vaccine against COVID-19. Of those, Moderna needed just 65 days from receiving the genetic sequence on 11 January 2020 to starting the first human trial on 16 March. The other 270 days, just under nine more months, went almost entirely on clinical trials, with a few weeks of regulatory process in November and December.

He is careful not to call that a failure. Modern trials are genuinely complex: doses must be manufactured to quality standards, paperwork filed, predictions made about where infections will occur, insurance negotiated, tens of thousands of patients recruited, and every adverse event reported in detail and almost in real time. The process worked as well as it ever has. But in 2020 over 10,000 people were dying with COVID-19 every day, so the value of speed was enormous, and CEPI now wants the timeline cut to 100 days, against its own estimate that the maximum

achievable under the current system was 250 to 300 days.

The core arithmetic is that a phase 3 trial cannot tell who has been exposed, so it counts infections in each group and must either run long or recruit many. Having twice as many participants halves the expected length. Why not simply recruit twice as many? Logistics and money: it took months to recruit tens of thousands, and trials are already expensive, with a newly announced phase 3 trial for a tuberculosis vaccine expected to cost $550 million.

He then walks through four ways to go faster. Adaptive platform trials evaluate several treatments against a common control and add or drop them as evidence accumulates. Human challenge trials deliberately expose healthy young volunteers, which for COVID could have proved efficacy in weeks with only 100 to 200 participants. Correlates of protection let you infer efficacy from immune response, as with seasonal flu vaccines. And compassionate use lets vulnerable people get an unapproved drug before approval. None of the four changed the COVID timeline, some for technical reasons and some, he argues, because regulators or makers judged them too risky.

The last section is about what trials do not measure. Because manufacturers test one dose, we never learned whether a lower dose or a longer gap would have been better. The UK extended its dose gap to 12

weeks on limited evidence, and a modelling study suggests this reduced UK COVID mortality in 2021 by about 15 percent. Nobody had an incentive to run the trial that would have shown it.

Why it's assigned

This is the anchor reading for her chosen deep dive, clinical trials innovation, and it is the piece that defines the problem the other two readings on this page answer in practice. The Combat layer's claim is that time between pathogen and protection is the thing worth attacking; Więcek shows that in 2020 the design stage was already fast and the testing stage was not, so the remaining gains have to come from how trials are built, run and incentivised. Read it against the clinical trial bottlenecks game on page 5.1: he is naming the same levers in policy language.

In depth

Where the 326 days went

The headline number is the one worth holding: 326 days to create, test and authorize the first highly effective vaccine against COVID-19, when progress on vaccines is normally measured in years.

Inside it, Moderna took 65 days from receiving the genetic sequence of the then-novel coronavirus on 11 January 2020 to designing the vaccine, demonstrating efficacy in cells and animals, and

starting the very first human trial on 16 March. All of that was on a novel vaccine platform never before used at scale.

Then 270 additional days passed before FDA emergency use authorization for the general population. The regulatory process accounted for several weeks in November and December; the majority of the time went on clinical trials.

Więcek's framing of why that is not a scandal is worth keeping. Trial doses must be manufactured to quality and safety standards, paperwork filed with regulators, predictions made about which countries will have the most infections, insurance negotiated, tens of thousands of patients recruited, and every adverse event for every one of them reported in detail and in almost real time. During the pandemic the process worked as well as it ever did and delivered faster than almost anyone expected.

He gives the critics their due too. In 2020, over **10,000 people were dying with COVID-19 every day**, and in fast-spreading epidemics the benefits of vaccination fall away with time, because earlier delivery both protects at-risk individuals and helps build herd immunity. Against that, extreme speed pressure causes errors, and his example is AstraZeneca accidentally giving a lower dose than planned to a large proportion of participants in its trial. In ordinary times that is a minor tragedy and a

business failure; in a pandemic it raises the prospect of an intervention that could save millions being delayed by months because of a small, hard-to-detect accident.

The target now is CEPI's: shorten the timeline from a year to **100 days**. Więcek's own benchmark for how much of that is available from current practice is blunt. The 326 days Pfizer and BioNTech took was close to optimal given the constraints of the current system, and CEPI estimates the maximum achievable speed was 250 to 300 days. In March 2020, most experts regarded an 18-month timeline as optimistic.

The arithmetic that makes phase 3 slow

Human trials begin in phase 1 with a small number of subjects, typically fewer than 100, to establish safety. Phase 2 gives the vaccine to a larger group, typically several hundred, to see whether immune systems respond and at what dose. Outside pandemics these steps take many years. Only then does a promising vaccine move to large-scale phase 3 trials, usually double-blind randomized controlled studies, which establish efficacy.

The structural problem is simple. Because phase 3 trials track participants in the real world, there is no way to tell reliably how many have actually been exposed to the pathogen. So the trial measures how many people were infected in each group, which

means it has to run for a long time or recruit a lot of participants. As Więcek puts it, the math is very simple: having twice as many participants will halve the expected length of the trial.

His footnote gives the real figures. Each COVID-19 trial needed to record about **100 to 200 infections** to confidently conclude high efficacy. For COVID-19 the strategy was to vaccinate the general population and anticipate where many infections would occur. Thanks to that, AstraZeneca was able to deliver first results based on just 11,000 patients in the UK and Brazil, while Pfizer required over 40,000 subjects in six countries.

So why not recruit twice as many? First, logistics: it took the COVID trials months to recruit the tens of thousands they needed. In principle more money would help, but trial costs are already very high. A recently announced phase 3 trial for a new tuberculosis vaccine is expected to cost **$550 million**, and for COVID vaccines the total costs of R&D were in the billions for each successful vaccine.

What 2020 already did

Developers and regulators did work hard to compress the traditional process. Pfizer, Moderna and AstraZeneca took on more risk by merging phases into phase 1/2 or phase 2/3 trials. Więcek is honest that this is not normally attractive to vaccine makers,

since a large proportion of vaccines fail at each phase and separate protocols are easier and less risky than planning all stages at once. On the failure rate he gives a range rather than a figure: historical averages depend on how you count, success rates have been relatively high in the last two decades, but it is reasonable to assume a greater than 50 percent chance of failure at each stage.

Cooperation from country regulators also cut wait times, for example by allowing companies to skip animal testing, and trials benefited from logistical and manufacturing improvements. His footnote credits the research organizations that big pharmaceutical companies outsource trial operations to, down to small improvements like better data cleaning tools, which normally take months. Manufacturing mattered too, since preparing trial batches usually takes months; Pfizer was reported to have cut batch production time from 110 to 60 days through 2020 and 2021.

Four ways to go faster

Adaptive platform trials. Unlike normal trials with fixed protocols, adaptive platform trials allow multiple treatments to be evaluated simultaneously, dynamically adding or removing them in response to what is learned during the trial. They are not used widely because they are complex, not every trial needs

dynamic elements, and regulators' guidance is insufficient.

In 2020, simple adaptive elements were used to design stopping rules for COVID vaccine trials: alongside joint protocols for multiple phases, trials included interim analyses that allowed earlier reporting. The reason phase 3 trials must be so large is that only a fraction of participants get infected; an interim analysis occurs after a certain number of infections, and if the evidence of effectiveness is overwhelming at that point, the trial can conclude. This helped trials report faster, though Więcek says it is hard to say by how much. His footnote explains why the effect was muted: some interim analyses were scrapped or delayed after discussions with regulators, the FDA's summer 2020 safety-data request made others impossible, and rising incidence in autumn 2020 meant trials recorded many more infections anyway.

Platform trials proper evaluate many interventions against a common control group. They are harder to set up and difficult to use for comparing different vaccines against each other, but their benefit is that participants help study the most promising candidates without starting a new trial. For COVID this was not particularly important, because decision-makers largely bet on the winning horses, Moderna, Pfizer and AstraZeneca. His general rule is the useful

part: the lower the chance of success, the larger the benefit of a system that can test many candidates in parallel. Some vaccine platform approaches were tried later in the pandemic and were very effective for finding COVID therapeutics, but to make a difference next time we need experience running them, and to date only a handful of platform trials have ever been tried.

Human challenge trials. HCTs solve the problem of waiting for infections to occur: healthy young volunteers are deliberately exposed to the virus to study the vaccine's protective effect. This is routinely done to study the characteristics of pathogens such as flu. In the case of COVID, an HCT could have proved a vaccine's efficacy in a matter of weeks with only **100 to 200 participants**.

He names two scientific problems. Scientists need to produce an attenuated version of the virus that can safely be given to volunteers, which may take months; the riskier alternative is letting people infect each other in a less controlled way. And vaccines are typically less effective in the elderly because of immune system ageing, the same process that makes the elderly more at risk, so it is hard to extrapolate HCT results to the people we most need to protect.

In 2020, 40,000 volunteers declared their readiness to participate in an HCT, and two of the major COVID vaccine makers explored using them. But there was no

real willingness on the part of regulators to approve vaccines based on HCT data, and the idea was never tested. Więcek discloses that he is a scientific advisor to 1Day Sooner, the nonprofit established in 2020 to connect volunteers wanting to participate in COVID HCTs and to advocate their use.

Correlates of protection. For many diseases we know how immune reactions to a vaccine correlate with protection from disease, which is why a new version of the influenza vaccine can be produced periodically without recruiting tens of thousands of volunteers: measuring immune response in a handful of subjects in the lab is sufficient. For SARS-CoV-2 no such correlates were firmly established when vaccines were being developed. Broad scientific efforts are now in place to establish or improve correlates of protection, not just for coronaviruses but for other threats, and while less preferable, they may offer the first line of defence in future pandemics, at least in some of the main pathogen families.

Compassionate use. Also called expanded access, this lets individuals in the US and EU request some not-yet-approved drugs with a physician's recommendation, a model first used in the early 1990s with azidothymidine, the first HIV drug. In summer 2020 some argued for using COVID vaccines in care homes this way, and Dr Deborah Birx, the White House Coronavirus Response Coordinator,

encouraged vaccine makers to use the pathway, though none did.

Why none of them mattered in 2020

All four were floated and none had an impact. Więcek separates the reasons carefully, which is the most useful paragraph in the piece.

Some were not technologically feasible at that stage: an attenuated virus would have taken months to develop for HCTs, and the process was too early for correlates of protection to play a role. Some did not seem viable under a decentralized R&D ecosystem: while many regulators are generally supportive of platform trials, it is hard to imagine the main vaccine producers pooling together to start a joint trial of their vaccines.

But in a couple of cases, he argues, feasible solutions were simply discouraged by regulators or vaccine makers because they were seen as too risky. One was compassionate use for the elderly in summer 2020. The other was the FDA's September 2020 guidance that vaccine makers should not submit for emergency use authorization until there were **two months of safety data on at least half of trial participants**, apparently regardless of the size of the trial, which he says makes the reasoning hard to understand. That guidance in turn made the adaptive elements of trials

largely irrelevant, since trials were then very likely to run until November.

What we never measured

The final argument is the most original. Successful development drove tens of billions of dollars of profit for the vaccine makers, with most revenue captured by the first few companies to get approved and scale manufacturing. His footnote makes the winner-take-most point vivid: in the US, 97 percent of doses delivered have been Pfizer or Moderna, and in the EU 90 percent; Novavax reported comparable efficacy only several weeks later, in January 2021, by which point wealthy countries were already delivering the other shots, and Novavax's market share in US and EU vaccinations is about 0.025 percent.

Against those profits, the benefits dwarf the costs. By one estimate, in 2020 the social value of a COVID vaccine was over 100 times higher than its cost. Więcek's own footnote explains where that comes from: in late 2020 he was part of a group of economists that calculated the total benefit of projected vaccine supply as **$17 trillion**, the total economic cost including ongoing impacts on GDP, education and health that could be averted, equal to **$5,800 per course of vaccine**, against the $6 to $40 manufacturers were charging at the time. Even a conservative calculation looking only at ongoing GDP losses would price the course at $2,900. That gap

justified supporting research and mass production of many vaccines in parallel, as Operation Warp Speed did, a form of insurance by increasing the number of shots on goal. By another estimate, the entire OWS would have paid for itself by shortening the pandemic by **12 hours**.

Vaccine approval depends on showing the product meets the regulator's criteria of individual safety and efficacy, and after authorization we keep monitoring safety, effectiveness in the real world, and durability. But important questions remain unanswered by that process. Does the vaccine hold up against variants? Was the dose tested in phase 3 optimal? Vaccine response changes with age, yet under cost and time pressure manufacturers typically test only one dose, the one they think will work best for an average person.

He thinks that happened twice during the pandemic. First, in early 2021 the UK extended the gap between two doses from the three to four weeks tested in clinical trials to 12 weeks, letting more people get some protection in the first crucial months. A modelling study suggests this reduced COVID mortality in the UK in 2021 by about **15 percent**, and many other countries facing shortages followed. Second, some trial data suggested it might have been beneficial to lower the doses of the Moderna and Pfizer vaccines, which could be done simply by

drawing less volume into the syringe. Doing so would reduce efficacy only a little while allowing much faster vaccination. The idea was backed by immune response data and subsequently tested in many trials, but not implemented outside the lab for the first course of vaccination, though Moderna did reduce the dose for boosters.

In both cases, opponents were reluctant because comprehensive clinical trial evidence was missing, and Więcek's point is that under the current framework it is no wonder the evidence did not exist. In the race to develop a vaccine, why would a producer spend additional hundreds of millions of dollars to test a slightly less efficacious version of their own product? He notes the irony that in AstraZeneca's case it was likely the initial dosing error that inspired subsequent trials of a lower dose, because patients accidentally given the lower dose did not differ significantly in their immune responses from those given the full dose. Data was eventually generated, and in both cases seemed to confirm that extending gaps and lowering doses were feasible, but all of it came too late. In the end experts had to make uncertain decisions about lowering doses, extending gaps and timing boosters without complete trials.

The path to 100 days

The success of the COVID vaccines depended on advances in the core science of vaccine making. To go

faster still, Więcek argues, we have to look beyond the lab bench to the science of clinical trials, and each option requires thinking about incentives and the regulatory framework that shapes them. These developments must happen before the next pandemic arrives. None of them alone is guaranteed to get us to 100 days, but all of them together may increase the likelihood. His closing line is an economic one: COVID proved decisively that it is economically viable to pay for a wide range of solutions.

Words to know

- **Phase 3 trial** The large randomized study that establishes whether a vaccine actually prevents disease, by comparing infections between vaccinated and control groups. Its size is set by how many infections it needs to count.
- **Adaptive trial** A trial whose protocol can change as results come in, for instance by stopping early at a pre-set number of infections. It is how COVID trials reported sooner than planned.
- **Platform trial** A trial evaluating many interventions against one shared control group, so new candidates can be tested without starting from scratch. Only a handful have ever been run.
- **Interim analysis** A planned look at the data partway through, which can end the trial if the evidence is overwhelming. The FDA's safety-data request made several of them pointless in 2020.

- **Human challenge trial** A study in which volunteers are deliberately infected, so you do not have to wait for natural exposure. For COVID it could have shown efficacy in weeks with 100 to 200 people.
- **Correlate of protection** A measurable immune response that reliably predicts protection, letting you skip a large efficacy trial. It is why flu vaccines can be updated each year without recruiting thousands.
- **Compassionate use (expanded access)** A route for an individual to obtain a not-yet-approved drug on a physician's recommendation. First used in the early 1990s for the HIV drug azidothymidine.
- **Shots on goal** The insurance logic of funding many candidates in parallel because you do not know which will work, which is what Operation Warp Speed bought.

Numbers in this reading

- 326 days: from the start to the authorization of the first highly effective COVID-19 vaccine.
- 65 days: from Moderna receiving the genetic sequence on 11 January 2020 to the first human trial on 16 March 2020.
- 270 days: the remaining time to FDA emergency use authorization, just under nine months, most of it clinical trials.
- Over 10,000: people dying with COVID-19 every day in 2020.

- 100 days: CEPI's target. 250 to 300 days: CEPI's estimate of the maximum achievable under the current system. 18 months: what most experts in March 2020 considered optimistic.
- Fewer than 100 subjects in phase 1; typically several hundred in phase 2.
- 100 to 200 infections: what each COVID-19 trial needed to record to conclude high efficacy confidently.
- 11,000 patients in the UK and Brazil for AstraZeneca's first results; over 40,000 subjects in six countries for Pfizer.
- Greater than 50 percent: a reasonable assumption for the chance of failure at each trial stage, though Więcek stresses the historical average depends on how you count.
- $550 million: the expected cost of a newly announced phase 3 trial for a tuberculosis vaccine. COVID vaccine R&D ran into the billions for each successful vaccine.
- 110 to 60 days: the reduction Pfizer was reported to have achieved in the time to produce a batch of its vaccine.
- 100 to 200 participants, in a matter of weeks: what a human challenge trial for COVID could have required.
- 40,000: volunteers who declared readiness to participate in an HCT in 2020.

- Two months of safety data on at least half of trial participants: the FDA's September 2020 requirement before an EUA submission.
- Three to four weeks extended to 12 weeks: the UK's dose gap change in early 2021, which a modelling study suggests cut UK COVID mortality in 2021 by about 15 percent.
- Over 100 times: how much higher the social value of a COVID vaccine was than its cost in 2020.
- $17 trillion: the total benefit of projected vaccine supply calculated by Więcek's group in late 2020, equal to $5,800 per course against the $6 to $40 manufacturers were charging; a GDP-only calculation gives $2,900.
- 12 hours: the shortening of the pandemic that would have paid for the entire Operation Warp Speed, by one estimate.
- 97 percent of US doses and 90 percent of EU doses from Pfizer or Moderna; Novavax's share about 0.025 percent after reporting comparable efficacy several weeks later in January 2021.

What the author is unsure about

Więcek flags his own interest: he is a scientific advisor to 1Day Sooner, which advocates for human challenge trials, and he was part of the economist group that produced the $17 trillion figure. On the evidence he is careful. He says it is hard to say how much interim analyses actually sped up reporting. He

gives the trial failure rate as a reasonable assumption rather than a measured number, and says the historical average depends on how you count. The 15 percent mortality reduction from the UK's longer dose gap comes from a modelling study, not an experiment. He says he finds the reasoning behind the FDA's September 2020 safety-data request hard to understand, which is a criticism rather than a finding about intent. He states plainly that none of his four options alone is guaranteed to reach 100 days, offering them together as something that may increase the likelihood. And he notes the real limits of HCTs: an attenuated virus takes months to produce, and results in healthy young volunteers do not extrapolate well to the elderly, who are the people most in need of protection.

For the discussion

- The single most useful fact to bring is the split: 65 days to design the vaccine, 270 days to prove it, which means the remaining time is in trial design and incentives rather than in science.
- The dose question is the sharpest critique in the piece and the easiest to state: no manufacturer in a race has any reason to spend hundreds of millions proving that a cheaper, slightly weaker version of its own product would save more lives, so that evidence only ever arrives late.

5.4-2 · CORE

How the UK found the first effective Covid-19 treatment — and saved a million lives

Dylan Scott · 2021

Read the original paper → · the course assigns: whole article, from the Internet Archive's 2021 copy of the Vox page

> **Two Oxford researchers designed a trial simple enough that a ward nurse could run it in 20 minutes, plugged it into the NHS, and within three months had the world's first treatment proven to cut COVID deaths.**

Average summary

Scott's story starts with a phone call in June 2020. Peter Horby had just seen the first data from the UK's Recovery Trial on dexamethasone, a cheap, widely available steroid, and called his Oxford colleague Martin Landray. His summary was four words: "Holy shit, it works." On 16 June, after verifying the data, the trial announced that a drug had finally been shown to reduce COVID-19 mortality. Dexamethasone became part of the standard of care worldwide, and a UK government estimate published in March 2021 concluded the drug had so far saved **22,000 lives in the UK and an estimated 1 million worldwide**.

The point of the piece is how the trial got built. Horby and Landray first met face to face on 5 March 2020,

when Landray had never worked on infectious diseases and Horby had never run a trial of more than about 200 patients. They had a proposal by 10 March, pitched it to England's chief medical officer Chris Whitty, and he and his counterparts in Scotland, Wales and Northern Ireland wrote to NHS hospitals asking them to take part. More than 175 hospitals eventually joined, and within two weeks of that first meeting more than 1,000 patients had enrolled.

What made that possible was the design and the system it ran in. The design came from Oxford's 1980s heart attack trials, which enrolled more than 140,000 patients by being simple enough for ordinary ward staff to run. Recovery copied it: recruitment started when a patient walked through the door, consent was one page instead of the usual five or six, randomisation was automatic, the NHS's electronic system generated the prescription, and the trial drug became just another name on the drug chart. Nurses could be trained to run it in about 20 minutes.

The system was the NHS. One organisation, roughly 1,250 hospitals, staff who are government employees, and a single shared record system. The largest hospital system in the US has fewer than 200 facilities, and American trials had to coordinate across disconnected systems; Regeneron ended up bringing its antibody cocktail to Recovery after its own

US trial enrolled only a few hundred patients, while the UK arm would include more than 9,000.

Scott gives the critics room. Simplicity costs precision: the trial can answer whether a patient lived or died, not whether a drug suits a diabetic in their seventies. It has been accused of science by press release. The Berlin epidemiologist Tobias Kurth questions whether results stopped early are signals or settled findings, noting that tocilizumab saves one life for every 25 patients treated and that more data could in principle make that link disappear. Horby's answer is that they had confidence in the data and would not sit on it.

Why it's assigned

This is the worked example for her deep dive. Where Więcek's piece names the levers that could shorten trials, this is the one case where somebody pulled several of them at once, in the middle of a crisis, and produced a result that changed clinical practice everywhere within days. For the Combat layer it also widens the frame: the first thing that reduced COVID mortality was not a vaccine but a cheap old steroid, found because somebody built the machinery to look.

In depth

The moment, and what it was worth

Horby had just seen new data from the Recovery Trial, an experiment enrolling tens of thousands of patients at dozens of hospitals to investigate COVID-19 therapies. He and Landray had dreamed up the trial programme in March 2020. Nearly three months later, more than 35,000 people were dead in the UK and more than 100,000 in the US, and they were still hunting for the world's first lifesaving treatment.

The data showed that dexamethasone had lowered the likelihood of death for COVID-19 patients who required oxygen or a ventilator in hospital. On 16 June, after verifying it, the trial announced the finding. Dexamethasone became part of the standard of care across the world, and a UK government estimate published in March 2021 concluded the drug's use had so far saved 22,000 lives in the UK and an estimated 1 million lives worldwide.

Scott sets the context plainly. As the pandemic exploded, medical science was starting from scratch. The long-range goal was a vaccine, but millions could and would die before a shot was approved and distributed, so the best bet for preventing deaths in the meantime was to find existing drugs that could treat COVID-19. Doing that required enormous capacity: testing many therapies at once, which means recruiting thousands of patients. In the US, as Carl Zimmer wrote in the New York Times in January 2021, many trials for COVID antivirals were doomed

from the start, too small and poorly designed to produce useful data.

Recovery was the opposite: massive and simple. Scott calls it the most effective programme in the world for delivering urgently needed research results. Pharmaceutical companies sought partnerships with it because they saw it as the most reliable way to find out whether their drugs could help. Since June 2020 it has identified a second drug that improves mortality, **tocilizumab**, and has also performed the valuable service of showing that some once-promising candidates do not work.

He is careful about the national frame. Britain struggled in many important ways and has suffered among the most cases and deaths per capita of any country. Recovery is the exception.

Where the design came from

The road to Recovery started in the 1980s, when a group of Oxford scholars, dissatisfied with the lack of treatments for heart attacks, imagined a trial that could test different interventions at a scale of perhaps 10,000 to 15,000 patients. For a trial that big to work, it had to be simple enough that nurses and doctors could try out the treatments as part of their normal care routine.

The system began when a patient was admitted. The research team told the medical staff which treatment

pack to use, containing either a placebo or a treatment, with nobody, including the researchers, knowing which. Staff reported whether the patient lived or died, plus notable events such as a blood clot or a stroke. That research, the **International Studies of Infarct Survival**, enrolled more than 140,000 patients and identified several treatments that significantly cut heart attack deaths, including a combination of an anti-clotting enzyme and aspirin, a regimen still in use today. It became famous in the field and attracted younger researchers like Landray, who came to Oxford in 2000 to work with the team.

But nobody had set up a very large trial programme in the middle of an emergency, and previous efforts had failed. Horby had lectured on the missed research opportunities of the SARS crisis. He had tried to start trials during the 2009 H1N1 flu pandemic, and says it took researchers three weeks to draft the proposals, which then spent two more weeks in front of an ethics review board before approval; the trials that made it through enrolled just a few hundred patients, nowhere near enough for robust results. He calls it a failed model for epidemic infections.

Ten days from email to proposal

Landray, meanwhile, had spent his career figuring out how to run very large trials, and had never given infectious disease outbreaks much thought.

On 28 February 2020, when there were still just 20 confirmed COVID-19 cases in the UK, he emailed Jeremy Farrar, head of the Wellcome Trust. It seemed clear to him that SARS-CoV-2 would spread fast and that trials would be needed, and that in the short term doctors would need an existing treatment that also worked against COVID-19. Farrar suggested he contact Horby, who had secured financial support for a multi-drug trial in China; by the time that financing was approved the Chinese outbreak had slowed while UK cases were taking off.

When they sat down together in Oxford on **5 March**, they realised they each had one half of the equation. Horby had the protocols for a multi-drug investigation during a pandemic; Landray had been thinking about how to establish a massive trial in the UK. Landray's summary of the mismatch is the detail worth remembering: Horby had never done a clinical trial over about 200 patients, and Landray had never worked on infectious diseases.

They had a proposal drawn up by 10 March. With Farrar's support they pitched it to England's chief medical officer, Chris Whitty, who bought it. He and his peers in Scotland, Wales and Northern Ireland sent a missive to NHS hospitals urging them to take part, and more than **175 hospitals** eventually agreed. Duncan Browne, who runs the trial at a hospital in Cornwall, says they were able to get going within days

because of the NHS, and that with a fragmented approach they would still be arguing about it.

Within two weeks of Horby and Landray's first meeting, more than 1,000 patients had enrolled.

What it looked like on the ward

Browne and Fiona Hammonds, the lead research nurse at the Cornwall hospital, had both been working on diabetes before the pandemic. Neither had worked on infectious diseases. Browne says the lack of experience did not matter, because it is really about the process of doing research.

The operational details are the heart of the piece: - Recruitment started as soon as a patient walked through the hospital doors, with UK emergency rooms stocked with pamphlets and posters advertising the trial. - Within three or four hours of testing positive and being admitted, an eligible patient was approached by a member of the research team and asked whether they would like to take part. Many wanted to. More than 35,000 patients enrolled over the trial's first year. - Consent was a **simplified one-page form** instead of the more typical five or six, shorn of much of the fine print. - The team entered the patient into the trial's computer algorithm, which randomly assigned them to one of the trials. The NHS's electronic system then generated a prescription for the experimental drug for half the

patients; the other half served as the control group. - The trial drug became just another name on the patient's drug chart, given by nurses in the normal course of care. Browne says it took no more time for nurses and doctors on the ward than conventional treatment. - Nurses could be trained to run the trial in about **20 minutes**. They logged the discharge date and any side effects or adverse events, and the information uploaded automatically to an online NHS database where an independent panel monitored it.

Choosing what to test

A group of experts formally recommended which treatments to include. Some candidates were obvious because they were already being used in COVID care; hydroxychloroquine, promoted by then-President Donald Trump as a possible cure, was one.

Others, like dexamethasone, were educated guesses. There was limited research on dexamethasone and viral infections, some of it suggesting the drug could help and some indicating it could be dangerous at high doses. Prominent experts warned against testing it, Horby and Landray say, because the potential adverse effects made it unethical. The investigators saw it the other way round. There was already an existing proposal to set up a steroid trial whenever the next flu pandemic happened, and they thought the uncertainty was precisely the reason to run a trial. Horby's version: there was a lot of controversy about

it, which probably tells you it is a good drug to test, and a trial was the only way to settle divided opinion.

In one year the trial examined nine repurposed drugs, a convalescent plasma and a newly developed antibody cocktail, and found two drugs that reduce mortality in seriously ill patients: dexamethasone and, later, tocilizumab.

One patient's experience

Paul Buckler, 46, was one of about 4,000 patients enrolled in the tocilizumab trial. He lives near Southampton and had been careful during the pandemic, because his father has chronic obstructive pulmonary disease, but was infected in early November. He first noticed he could not smell his hand sanitizer, tested positive, and after a few days started coughing and lost his appetite. His home monitor showed low blood oxygen; his doctor told him to get to a hospital.

The morning after his first night there, a nurse and a Recovery staff member pitched the trial to him. They reassured him they knew the drugs well, including what side effects to expect, and gave him material to read, though he admits he did not examine the fine print. By that afternoon he had been selected for the tocilizumab trial. He never knew whether he got the placebo, and he started feeling better within days.

In February 2021 the results were published: tocilizumab, largely when taken alongside dexamethasone, reduced mortality for patients receiving oxygen or on a mechanical ventilator, and helped patients with severe symptoms be discharged more quickly. Scott notes this was another victory for the premise of going big, because earlier, smaller trials had failed to show any benefit from tocilizumab. Landray's line about the speed of adoption is the one everyone quotes: results published at lunchtime, practised by tea time.

The criticisms

Scott gives the critics a full section, and even they open with praise for the trial's scope and efficiency, reaching for words like brilliant and stellar. Tobias Kurth, a Berlin-based epidemiologist who focuses on study design, says the approach was really good given that COVID was new and nobody knew how to treat it.

His objections are about what simplicity costs: - **Precision.** Big numbers are good at answering basic questions, such as whether the patient lived or died, but do not provide the detail or certainty researchers are used to. Horby concedes it directly: asked at the WHO whether tocilizumab would be clinically appropriate for a diabetic in their seventies, his answer was that he could not say, because they did not look at that specific subgroup. - **Science by press release.** The trial has at times been labelled

with that phrase, which has stained other findings during the pandemic. Announcing initial results without a full data set could erode trust among the experts whose acceptance is needed for widespread adoption. - **Signal versus result.** For Kurth the core methodological challenge is how the trial conveys certainty. There is a large difference between a signal, an association that may indicate but does not prove cause and effect, and a demonstrable causal link, and he says it is still hard to know which one we are getting. - **Stopping too soon.** Recovery data is published when the independent monitoring panel decides it has seen enough. Kurth's worry is that the call might come too early, and that an apparent effect could disappear with more data. His example: tocilizumab saves one person who would otherwise have died for every **25 patients** who take it, and it is at least conceivable that with more data that link could evaporate.

Horby and Landray knew they were taking that risk, and judged that they had a responsibility to publish clinically relevant information that could save lives, because even a small effect means thousands of deaths averted given how many people were hospitalised. Horby's answer: they had confidence in their data, and asks whether they should sit on it while hundreds of thousands of people are in hospital.

Why the UK and not the US

The NHS is Scott's explanation. It made it possible to launch with the urgency the moment required. Derek Lowe, who writes about drug discovery, told him the US could conceivably have set up a similar programme, and did not. Instead, American trials had to coordinate among disconnected hospital systems. The NHS is a single system of roughly **1,250 hospitals**; the largest hospital system in the US includes fewer than 200 facilities and most are much smaller. American efforts struggled to reach Recovery's scale and were slow to deliver results.

The consequence is that even American companies came to the UK. Regeneron, a US biotech that developed a cocktail of COVID-19 antibodies, initially ran a small trial with America's biodefence research agency, BARDA, then realised it needed a much bigger sample to get reliable results on deaths. When it was ready for wider testing in September it partnered with Recovery. The US trial had enrolled just a few hundred patients; the UK trial would eventually include data from more than 9,000 people. Leah Lipsich, a vice president at Regeneron, says working with a big-basket trial was the way to go for meaningful results, and that the NHS is an enormous linked network seeing, in effect, every COVID-19 patient in the country, which she does not think can be said of anywhere in the US.

Scott notes the UK's wider research record too: two other trial programmes ran alongside Recovery, one for primary care patients and one for the critically ill; Browne's hospital took part in the Siren study tracking antibodies in hospital workers; British scientists led the world in sequencing coronavirus genomes, which was critical to identifying new variants; and the UK vaccinated about the same share of its population as the US, with almost half of those doses the AstraZeneca vaccine designed by Oxford scientists.

He ends on the ward. Hammonds, the nurse in Cornwall, found comfort in the first dexamethasone result, because she could then tell patients she was recruiting that one thing had already been found that worked and would help them. Her recruiting line, once dexamethasone was standard: we know what the best treatment is and you are receiving it, would you like one or two more?

The article was published on 26 April 2021 as part of Vox's six-part Pandemic Playbook series, with photographs by Alicia Canter and support from the Commonwealth Fund.

Words to know

- **Repurposed drug** An existing, approved medicine tested against a new disease. Nine of them went through Recovery in its first year, which is why results could change practice immediately.

- **Dexamethasone** A cheap, widely available steroid, the first drug shown to reduce COVID-19 deaths, in patients needing oxygen or a ventilator.
- **Tocilizumab** An anti-inflammatory drug, the second Recovery finding, which reduced mortality largely when given alongside dexamethasone.
- **Convalescent plasma** Blood plasma from recovered patients, tested in Recovery as one of its arms.
- **Science by press release** The criticism that initial results are announced before the full data set is published. Recovery has been accused of it, and defends the practice on the grounds that delay costs lives.
- **Signal versus result** The distinction between an association that may indicate a cause and a demonstrated causal link. It is Kurth's central methodological worry about stopping trials early.
- **Big-basket trial** Regeneron's phrase for a trial large and open enough that an outside company can add its candidate and get a reliable answer.

Numbers in this reading

- 22,000 lives in the UK and an estimated 1 million worldwide: the UK government's March 2021 estimate of lives saved by dexamethasone.
- 16 June 2020: the date the dexamethasone result was announced, after verification.

- More than 35,000 dead in the UK and more than 100,000 in the US: the toll by the time the result arrived, nearly three months after the trial was conceived.
- 28 February 2020: Landray's email to Jeremy Farrar, when the UK had just 20 confirmed cases.
- 5 March and 10 March 2020: the first Horby-Landray meeting and the finished proposal.
- More than 175 hospitals: NHS hospitals that joined the trial.
- More than 1,000 patients enrolled within two weeks of that first meeting; more than 35,000 patients enrolled over the first year.
- Three or four hours: the window after a patient tested positive and was admitted before the research team approached them.
- One page instead of the typical five or six: the simplified consent form.
- About 20 minutes: the time needed to train a nurse to run the trial.
- Nine repurposed drugs, one convalescent plasma and one newly developed antibody cocktail, examined in one year; two of them reduced mortality.
- About 4,000: patients enrolled in the tocilizumab arm, one of them Paul Buckler, 46.
- One in 25: patients treated with tocilizumab who survive and otherwise would not have, the figure

Kurth uses to argue the effect could in principle be smaller than it looks.

- More than 140,000 patients: enrolment in the 1980s International Studies of Infarct Survival, the design Recovery was modelled on, which had been imagined at 10,000 to 15,000.
- About 200 patients: the largest trial Horby had run before this one.
- Three weeks to draft proposals plus two more weeks at the ethics board, for trials that enrolled just a few hundred patients: Horby's account of the 2009 H1N1 attempt.
- Roughly 1,250 NHS hospitals, against fewer than 200 facilities in the largest US hospital system.
- A few hundred patients in Regeneron's US trial, against more than 9,000 in the UK arm.

What the author is unsure about

Scott reports the uncertainty rather than carrying it himself, and he does it well. The million-lives figure is a government estimate, not a measurement. The critics' objections are given at length: that simplicity trades away precision, that announcing results before publishing full data risks eroding expert trust, that it is hard to tell a signal from a demonstrated result, and that stopping when the monitoring panel has seen enough could mean stopping before an apparent effect would have disappeared. Horby concedes he cannot answer subgroup questions. The investigators'

defence is a judgement about responsibility under uncertainty rather than a methodological rebuttal. Scott also notes that falling UK hospitalisations, likely driven by the vaccine rollout, had slowed enrolment by the time he wrote.

For the discussion

- The fastest thing in this story is not the science but the paperwork: a one-page consent form, a 20-minute training, automatic randomisation and a drug on the normal chart are what let a trial enrol 1,000 patients in two weeks.
- The trade-off Kurth names is the real question for trial innovation: Recovery bought speed and scale by answering only whether people lived or died, which is why it could not tell a clinician anything about the diabetic in their seventies in front of them.
- The structural comparison is uncomfortable and worth raising directly: a single national system with shared records and salaried staff produced the world's answer, while the better-funded, more fragmented system had to send its own companies abroad to get one.

5.4-3 · CORE

RECOVERY Trial Wikipedia

Wikipedia contributors · not given (encyclopedia article, as captured)

Read the original paper → · the course assigns: whole article

> **The reference account of what the RECOVERY trial actually is: an open-label, multi-arm adaptive trial in UK hospitals measuring death at 28 days, which found four treatments that work, six that do not, and stated the limits of its own headline result.**

Average summary

The Randomised Evaluation of COVID-19 Therapy, or **RECOVERY Trial**, is a large-enrolment clinical trial of possible treatments for people admitted to hospital in the United Kingdom with severe COVID-19 infection, later expanded to Indonesia, Nepal and Vietnam. It has tested ten interventions on adults: eight repurposed drugs, one newly developed drug and convalescent plasma.

Three design features define it. It is a randomized controlled trial. It is open label, meaning both the people receiving treatment and the clinicians know which treatment is being given. And it is **multi-arm adaptive**, so new treatments can be added as it progresses and other arms closed to new enrolment once results are in. Enrolment is automatic once a patient is in: randomisation has three parts, so a patient might be allocated none, one, two or three

treatments, with any treatment that is contraindicated, positively indicated or unavailable excluded from their randomisation, and a second randomisation if their disease progresses. The primary objective is reliable estimates of the effect of study treatments on all-cause mortality at 28 days after first randomisation. The protocol was developed in March 2020 and designed to minimise the administrative load on hospital staff facing overwhelming admissions.

It is run by the Nuffield Departments of Population Health and of Medicine at the University of Oxford, supported by the National Institute for Health Research, and led by Peter Horby and Martin Landray as co-chief investigators. By July 2020 it was running at 176 NHS hospitals, and by March 2021 it had enrolled more than 40,000 participants. Estimated primary completion was December 2021 and estimated study completion December 2031.

The results divide cleanly. Four treatments produced positive results: low-dose dexamethasone, tocilizumab, the antibody cocktail REGN-COV2, and baricitinib. Six were closed after being shown ineffective: lopinavir-ritonavir, hydroxychloroquine, azithromycin, convalescent plasma, aspirin and colchicine. Dimethyl fumarate was found in January 2024 not to improve outcomes significantly.

The dexamethasone result, announced in June 2020, cut the death rate by one third in people needing

ventilators and one fifth in those on oxygen therapy, with no benefit and the possibility of harm in those who did not require oxygen. The article is unusually careful about the limits of that finding, which are set out below.

Why it's assigned

Paired with the Vox article on the same page, this is the technical half of her deep dive. Vox tells the story of how the trial was built in ten days and what it changed; this tells her exactly what was tested, in whom, with what numbers, and with what caveats. For the Combat layer it also makes the design visible: the multi-arm adaptive structure and the single simple endpoint are the concrete answer to the trial-speed problem Więcek sets out, so it is the place to see an adaptive platform trial that actually ran.

In depth

What kind of trial it is

RECOVERY is a large-scale randomized controlled trial of treatments for people hospitalised in the UK with severe COVID-19, later extended to Indonesia, Nepal and Vietnam. It has tested ten interventions on adults: eight repurposed drugs, one newly developed drug and convalescent plasma.

It is an **open label** study, meaning both the people receiving the treatment and the attending clinicians

know which treatment is being administered. That is a deliberate simplification; it removes the need for blinding machinery on a ward in crisis.

It is a **multi-arm adaptive** trial: new treatments can be added as it progresses, and other treatment arms closed to new enrolment once results have been produced. The article notes that setting the trial up very fast was crucial for a fast-developing pandemic, and quotes Martin Landray in March 2021 saying it has set a new standard for what can be delivered, not just for pandemics, and that it would be a travesty to return to a situation where it sometimes takes years to get a trial off the ground.

Enrolment and goals

When people hospitalised with COVID-19 are enrolled, they are automatically randomized to receive trial treatments. If any treatment is contraindicated or positively indicated for that patient, or is not available, that treatment is excluded from their randomisation. The main randomisation stage has three parts, so patients might be allocated none, one, two or three of the trial treatments, and if their disease progresses there may be a second randomisation.

The primary objective is to provide reliable estimates of the effect of study treatments on **all-cause mortality at 28 days** after first randomisation. The trial protocol was developed in March 2020, and the

design minimises the administrative load on hospital staff, who at the time were facing the prospect of overwhelming numbers of COVID-19 admissions.

The arms

As of 25 June 2021, the treatments allocated at random to hospitalised people with severe COVID-19 were baricitinib, an immunomodulatory drug used in rheumatoid arthritis; dimethyl fumarate, adults only, early phase assessment; infliximab, adults only; and high-dose dexamethasone, for adults with hypoxia only.

Children with PIMS-TS could also be allocated low-dose dexamethasone, a steroid that reduces inflammation; intravenous immunoglobulin; tocilizumab, an anti-inflammatory; or anakinra, another anti-inflammatory.

Treatments previously included that obtained positive results: low-dose dexamethasone, tocilizumab, REGN-COV2 (a cocktail of two anti-viral monoclonal antibodies), and baricitinib.

Treatments previously included and closed to new entrants after being shown ineffective: lopinavir-ritonavir, an HIV medication; hydroxychloroquine, used to treat malaria and rheumatism; azithromycin, an antibiotic; convalescent plasma, blood plasma from people who have recovered from COVID-19 and which may contain antibodies against SARS-CoV-2;

aspirin, a commonly used blood thinner; and colchicine, a drug used for gout.

Who runs it, and at what scale

The trial is run by the Nuffield Departments of Population Health and of Medicine at the University of Oxford and supported by the National Institute for Health Research. It is led by Peter Horby and Martin Landray, who serve as co-chief investigators. By July 2020 it was in progress at **176 NHS hospitals** in the UK, involving many thousands of health professionals.

It began in March 2020. As of March 2021 it had enrolled more than 40,000 COVID-19 participants admitted to UK hospitals. The estimated primary completion date was December 2021 and the estimated study completion date December 2031, which is a useful reminder that a trial that changes practice within days can still be scheduled to run for a decade.

The dexamethasone result and what followed

In June 2020, preliminary results published in a preprint showed that low-dose dexamethasone treatment reduced the death rate by **one third** in hospitalised people needing ventilators due to severe COVID-19, and by **one fifth** in people treated with oxygen therapy. There was no benefit, and the possibility of harm, among people who did not require

oxygen. The preliminary report was subsequently published in the New England Journal of Medicine.

The speed of adoption is documented in detail. Six days before the preprint, the results were announced in a news release. A UK Therapeutic Alert was issued the same day, and all the Chief Medical Officers in the United Kingdom exceptionally recommended an immediate change of UK-wide clinical practice, in advance of publication of any final paper. Demand for dexamethasone surged after the preprint appeared.

Based on the preliminary, unpublished results, the US National Institutes of Health COVID-19 Treatment Guidelines Panel recommended dexamethasone in patients on mechanical ventilation or requiring supplemental oxygen, and recommended against it for those not requiring supplemental oxygen. Japan, Taiwan and South Africa granted specific approval for the drug as part of standard medical care. A preprint study, awarded Health Data Research UK Open Access Publication of the Month for August 2020, found the discovery would save approximately **650,000 lives globally over the course of six months**.

The limits of that result

This section is the reason the article is worth reading alongside the Vox piece. John Fletcher, research editor at the BMJ, noted limitations and causes of

concern in the dexamethasone results: - Around a third of patients in the trial were still in hospital at the end of the 28-day trial period, so their final outcomes were not known. - As an immunosuppressive drug, there are fears dexamethasone could make the illness worse and prolong the infection in patients whose immune system has not yet overreacted and caused inflammation. - The preprint authors themselves warned that the results are consistent with possible harm in patients who did not require oxygen at enrolment. The trial observed a **22 percent increase in mortality** in these patients, with a rate ratio of 1.22 and a 95 percent confidence interval of 0.93 to 1.61, p=0.14, though the article notes this observation may still be due to chance. - Responding to the publication, the WHO emphasised that dexamethasone should only be used for patients with severe or critical disease, under close clinical supervision, stating there is no evidence it works for patients with mild disease or as a preventative measure, and that it could cause harm.

The result was then replicated: a systematic review published in September 2020 by the WHO's Rapid Evidence Appraisal for COVID-19 Therapies working group confirmed the initial finding.

The other results, with their numbers

- **Hydroxychloroquine.** In June 2020 the trial determined there was no clinical benefit from its

use in people hospitalised with COVID-19.

- **Lopinavir-ritonavir.** In June 2020 the chief investigators reported no clinical benefit in 1,596 people hospitalised with severe COVID-19 over 28 days of treatment. Published in The Lancet on 5 October 2020.
- **Azithromycin.** In December 2020 the investigators announced no benefit. 2,582 patients received azithromycin and 5,182 were randomised to usual care alone; 28-day mortality in both groups was 19 percent. Preliminary results were released as a preprint the same day.
- **Convalescent plasma.** In January 2021 the investigators announced closure of this arm. Among 1,873 reported deaths from 10,405 patients, there was no significant difference in the primary endpoint of 28-day mortality: 18 percent with convalescent plasma against 18 percent with usual care alone, risk ratio 1.04, 95 percent confidence interval 0.95 to 1.14, p=0.34. Full results were published on 14 May 2021.
- **Tocilizumab.** It significantly reduces the risk of death in hospitalised patients with severe COVID-19. 2,022 patients allocated to tocilizumab were compared with 2,094 who received standard hospital care: 596, or 29 percent, of the tocilizumab group died within 28 days, against 694, or 33 percent, in the usual care group, a rate ratio of 0.86 with a 95 percent confidence interval

of 0.77 to 0.96, p=0.007, an absolute difference of 4 percent. Receiving tocilizumab also increased the chance of discharge from hospital within 28 days. This supported the earlier findings of the Remap-Cap trial on tocilizumab for patients in intensive care, and extended those findings to a wider group of patients.

- **Dimethyl fumarate.** A study published in January 2024 by the trial's chief investigators found it did not significantly improve clinical outcomes in hospitalised COVID-19 patients.

Words to know

- **Open label** Neither patients nor clinicians are blinded to which treatment is given. It removes a layer of machinery, at the cost of a layer of protection against bias.
- **Multi-arm adaptive trial** A trial that can add new treatment arms while running and close arms once an answer is in. It is why ten interventions could pass through one trial.
- **All-cause mortality at 28 days** The primary endpoint: whether the patient was alive 28 days after first randomisation, from any cause. Its simplicity is what allows ward staff to run the trial.
- **Rate ratio and confidence interval** The ratio of event rates between groups, with the range of values consistent with the data. The tocilizumab

rate ratio of 0.86 with a 0.77 to 0.96 interval is what makes that result convincing.

- **PIMS-TS** Paediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2, the children's condition for which the trial had a separate set of allocations.
- **REGN-COV2** A cocktail of two anti-viral monoclonal antibodies, the one newly developed product among the trial's interventions.
- **Convalescent plasma** Plasma from recovered patients, which may contain antibodies against the virus. The trial closed this arm after finding no mortality difference.

Numbers in this reading

- Ten interventions tested on adults: eight repurposed drugs, one newly developed drug and convalescent plasma.
- 28 days: the window for the primary endpoint, all-cause mortality after first randomisation.
- Three parts to the main randomisation, so a patient may receive none, one, two or three trial treatments, with a possible second randomisation on progression.
- 176 NHS hospitals by July 2020. More than 40,000 participants enrolled by March 2021.
- March 2020: trial start and protocol development. December 2021: estimated primary completion.

December 2031: estimated study completion.

- Dexamethasone: death rate reduced by one third in ventilated patients and one fifth in those on oxygen therapy; no benefit and possible harm in those not needing oxygen.
- 22 percent increase in mortality among patients not requiring oxygen at enrolment, rate ratio 1.22, 95 percent confidence interval 0.93 to 1.61, p=0.14, possibly due to chance.
- About one third of trial patients were still in hospital at the end of the 28-day period, so their final outcomes were unknown.
- Approximately 650,000 lives globally over six months: the estimated saving from the dexamethasone discovery, in a preprint study.
- Six days: the gap between the news release of the dexamethasone result and the preprint.
- Lopinavir-ritonavir: no benefit in 1,596 people over 28 days.
- Azithromycin: 2,582 treated against 5,182 on usual care alone; 28-day mortality 19 percent in both.
- Convalescent plasma: 1,873 deaths among 10,405 patients; 18 percent against 18 percent; risk ratio 1.04, 95 percent confidence interval 0.95 to 1.14, p=0.34.
- Tocilizumab: 2,022 allocated against 2,094 on usual care; 596 (29 percent) died against 694 (33 percent); rate ratio 0.86, 95 percent confidence

interval 0.77 to 0.96, p=0.007; absolute difference 4 percent.

What the author is unsure about

As an encyclopedia article it states facts and attributes judgements, but it is careful where the underlying science is uncertain, and it devotes a whole section to the limitations of its own headline result. It records that a third of trial patients had unknown final outcomes at 28 days; that there are fears dexamethasone could worsen illness in patients whose immune systems have not yet overreacted; that the preprint authors warned of possible harm in patients not needing oxygen, while noting the 22 percent mortality increase may be due to chance given the confidence interval crosses 1 and p=0.14; and that the WHO restricted its recommendation to severe or critical disease under close supervision. It also notes which results were announced by news release or preprint before peer-reviewed publication, and which were later replicated, in the dexamethasone case by the WHO's REACT working group in September 2020. Because it is a collaboratively edited article rather than a primary source, treat the figures as pointers to the cited publications rather than as the publications themselves.

For the discussion

- The adaptive design is what let one trial answer ten questions, and six of those answers were negative, which is a result worth defending out loud: knowing that hydroxychloroquine, azithromycin, convalescent plasma, aspirin, colchicine and lopinavir-ritonavir do not work freed clinicians from giving them.
- The strongest caution in the whole deep dive is in the dexamethasone subgroup: the same drug that cut deaths by a third in ventilated patients showed a 22 percent increase in mortality in patients not needing oxygen, with a confidence interval wide enough that it might be chance, and the WHO built its guidance on that uncertainty rather than around it.
- The trial changed UK-wide practice before any final paper existed, on a preprint and a news release, which is either the model for emergency evidence or the thing that will go wrong next time, and the two readings on this page disagree about which.

5.3-1 · CORE

How To Reuse the Operation Warp Speed Model

Arielle D'Souza · 2023

Read the original paper → · the course assigns: whole piece

> **Operation Warp Speed worked because of four reusable features and several that only a crisis can supply, and telling them apart is what decides whether a future "Warp Speed for X" makes sense at all.**

Average summary

D'Souza starts with the way the name is now used. Operation Warp Speed has joined the Manhattan Project and Apollo as shorthand for an ambitious government programme, and policymakers call for a Warp Speed for Alzheimer's or anything else. Her argument is that OWS succeeded because a particular environment allowed a particular design, and for most problems an OWS model would not pass a cost-benefit test.

The results first. OWS was a public-private partnership that deployed the Moderna and Pfizer-BioNTech vaccines across the United States in 2020, and the Johnson & Johnson vaccine in 2021. Between December 2020 and May 2021 it is estimated to have saved the lives of **140,000 Americans**, along with $1.8 trillion in economic benefits. Between 2020 and 2021, COVID-19 cost about $26 billion a day; OWS cost around $13 billion, roughly twelve hours' worth of those daily costs.

She then separates the features that were specific to COVID-19 from the ones that travel. Three were

specific. The emergency environment let the team cut through red tape, activate pre-planned Emergency Support Functions, and report directly to Senior Advisor to the President Jared Kushner, with authority to decide without White House clearance. Military logistics, under General Gustave Perna, mapped every manufacturing step and its failure points, built in redundancy, deployed the Army Corps of Engineers to expand sites and invoked the Defense Production Act for tubing, bioreactors and syringes. And manufacturing ran in parallel, with factories equipped for mass production before clinical trials were finished and purchase commitments negotiated with six industry partners at once.

Four features, she argues, generalise. A clear commercialization goal, which for OWS was at least one safe and effective vaccine manufacturable at scale and distributed nationally by the end of 2020, and which did not require a scientific breakthrough. Market-shaping mechanisms, including advance purchase agreements that paid even if a vaccine never received emergency use authorization. A whole-of-government strategy, including Product Coordination Teams and the practice of putting government staff inside manufacturing plants. And streamlined regulation, where the FDA allowed combined and concurrent trial phases and cut the usual wait between Phase II and Phase III review from up to six months to around three weeks.

Her closing test is about fit. An OWS approach suits problems with a verifiable deliverable and positive spillovers that markets underprovide, such as carbon capture, a universal coronavirus or flu vaccine, or a resilient electrical grid. It does not suit novel consumer products, where nobody knows in advance what people want.

Why it's assigned

Page 5.3 is the first deep-dive option, faster research and development, and this is its anchor: the clearest account of how the United States actually compressed vaccine timelines in 2020, written to be reusable rather than celebratory. For the Combat layer it does two things. It names the specific mechanisms behind the speed everyone cites, and it draws a line between what a crisis buys you and what policy can buy in advance, which is the same line the 100 Days Mission has to cross.

In depth

What OWS did, and what it cost

Strictly speaking, OWS was a public-private partnership that resulted in the deployment of two vaccines, Moderna and Pfizer-BioNTech, across the United States in 2020, along with the Johnson & Johnson vaccine in 2021. Between December 2020 and May 2021 it is estimated to have saved the lives

of 140,000 Americans, along with $1.8 trillion in economic benefits.

D'Souza highlights the cost-benefit arithmetic directly. Between 2020 and 2021, COVID-19 resulted in a $26 billion loss a day. OWS cost around $13 billion, or around twelve hours' worth of COVID-19's daily costs.

The features that only a crisis supplies

An emergency environment. OWS's agility came largely from the crisis itself, which let it cut through red tape. Logistical parts of the response relied on Emergency Support Functions, pre-planned structures that set out interagency coordination during disasters; ESF 8 was activated early and helped multiple federal agencies coordinate public health and medical services support. The OWS board drew staff from the White House and multiple executive branches, met regularly with and reported directly to Senior Advisor to the President Jared Kushner, who reported to President Trump, and was authorized to make decisions without White House clearance, avoiding standard bureaucratic channels. D'Souza's point is that this level of White House support usually only occurs during an emergency, and that the professional and personal commitment from industry, government and the military likely could not be sustained outside a crisis.

Military logistics. General Gustave Perna, previously the Army's Deputy Chief of Staff, led logistics with a military approach: work through anticipated implementation challenges, map out responsibilities, and build in redundancy at every level. Tasks were sorted into development, manufacturing or distribution challenges. The team invoked the **Defense Production Act** to allocate key equipment such as tubing, bioreactors and syringes. While therapeutic candidates were still in trials, the team mapped each step of the expected manufacturing chain along with its potential failure points. The Army Corps of Engineers deployed to expand manufacturing sites. Even mailing vaccine vials was tested repeatedly ahead of time.

Manufacturing in parallel. Because OWS optimised for a tight timeframe, it used a parallel task structure: manufacturing tasks usually done in series were done simultaneously. Factories were equipped for mass production before clinical trials of different candidates were even completed. Specialized packaging was developed and neighbourhood vaccination sites were set up at the same time. The team negotiated parallel orders with six different industry partners. Taking on additional redundancy costs is what bought speed.

The four features that travel

1. A clear commercialization goal. OWS had an explicit target that was measurable and verifiable: develop at least one safe and effective COVID-19 vaccine that could be manufactured at scale and distributed across the country by the end of 2020. That clarity made partnerships with the private sector radically simpler. With advance purchase agreements in place, industry partners focused on development while the government, principally through HHS and the Department of Defense, provided logistical support for manufacturing and distribution.

She adds a point that matters for anyone proposing a Warp Speed for something else: the strategy did not involve much novel or experimental research. It relied on industry's existing technology. Although COVID-19 accelerated mRNA technology and produced the first approved mRNA vaccines, research and development of mRNA therapeutics had been maturing since the late 1980s, the SARS-CoV-2 sequence was made publicly available on **January 10th, 2020**, and mRNA vectors had already been used successfully in animals before OWS. The government acted as an active partner creating a market for existing technology, not as a subsidizer of new science.

Her test, stated plainly: for many problems people want to solve this way, a clear commercialization goal simply does not exist, because the deliverable is underspecified, would require a major scientific

advance, or could look like many different things. Policymakers should use OWS tools where the ask from government to the private sector is explicit, narrowly bounded, and achievable without major scientific breakthroughs.

2. Market-shaping mechanisms. To demonstrate sustained commitment, the government used financial incentives that reduced the private sector's investment risk. OWS supported six vaccine candidates across three platforms: mRNA (Moderna and Pfizer-BioNTech), replication-defective live-vector (AstraZeneca and Johnson & Johnson), and recombinant-subunit-adjuvanted protein (Novavax and Sanofi/GSK).

Other than with Pfizer, OWS signed pre-purchase contracts committing to buy vaccines even if those vaccines later did not receive Emergency Use Authorization from the FDA. In Moderna's case OWS provided additional funding, because it could not assume the same risks as its larger competitors. Without those guarantees, the private sector would have been slower to make the investments needed. **Advance purchase agreements** reassured companies of a guaranteed market before trials had produced safety and efficacy data.

Her recommendation for a future partnership: be a flexible partner rather than a risk-averse subsidizer, because one industry partner's needs may differ

entirely from the next's, and consider what deliverables government can call into being simply by committing to provide a market, whether through pre-purchase agreements, milestone payments or innovation prizes.

3. A whole-of-government strategy. OWS required coordination across agencies and between government and industry that is rare outside a crisis. The team built a clear organizational hierarchy with channels for constant communication, and worked to function as a cohesive integrated group rather than a task force of people from other agencies. Leaders discouraged team members from invoking their original agency affiliation in meetings.

The organisational innovation was the **Product Coordination Team**. Each PCT had a team leader and specialists from industry, manufacturing and government agencies, and was formed to address industry needs immediately without running an interagency process for each need. Every OWS therapy and vaccine had a corresponding PCT working with its developing company through development, testing and distribution. Apart from the FDA, every US government agency had to inform a PCT leader before interacting with a manufacturer, which simplified information flows. PCTs held daily check-ins and conference calls to get clarification and secure clearances rather than waiting on inbox responses.

The other innovation was the persons in plants model: officers from the Army, the Biomedical Advanced Research and Development Authority and the National Institutes of Health were deployed to manufacturing plants across the country, giving the PCTs constant updates on challenges on the ground. D'Souza recommends that an OWS 2.0 implement this kind of on-the-ground operational support where possible, while recognising it may be too expensive to justify outside an emergency.

She illustrates what happens without such coordination using the FastGrants funders' catalogue of roadblocks delaying a pan-variant COVID-19 vaccine until at least 2024: groups need to run primate trials, then human trials, then ramp manufacturing and distribution, and are frequently tripped up by things outside their control, any one of which can set them back months. Her example is a group whose monkeys were delayed by US Customs, pushing the start of their primate trial back until September. Her conclusion is that an emergency model needs someone with the authority to tell Customs to expedite the monkeys, even if doing so falls outside normal protocol, which argues for a mission-oriented group rather than an interagency task force.

4. Streamlined regulation. OWS ran a coordinated effort to meet regulatory requirements in a time-

sensitive way without compromising safety standards. The FDA issued guidance under the Public Health Service Act allowing concurrent and combined trial phases and, in some cases, preclinical animal studies while human trials were under way. Normally animal studies come first and approval follows the conclusion of all human phases; here the structure was condensed, and Phase III trials were initiated during Phase I trials to determine efficacy in an ongoing pandemic.

Review of Phase II data was heavily prioritized. In traditional FDA review between Phase II and Phase III, therapeutics can be backlogged up to **six months** before receiving a green light; for OWS vaccines this was expedited to around **three weeks**.

When selecting candidates, OWS chose products with the highest potential of receiving an EUA. The FDA independently issued EUAs based on Phase III data, and even after securing them, companies continued their clinical trials. Additional measures included preparing for Phase III trials as soon as companies filed Investigational New Drug applications, recruiting large populations for Phase III, and stockpiling doses before FDA approval so they could be deployed immediately afterwards.

The generalisable lesson is that OWS identified the most time-consuming red tape and procedural barriers and increased communication between

companies and regulators. For another Warp Speed-like effort, the challenge is to find the equivalent barriers in a different field, and she names purely procedural regulations such as NEPA, which mandates an arduous process without demanding particular outcomes.

What could be next

The US government does not typically address challenges at warp speed, and OWS was a temporary crisis model from the start. The team almost chose the name Manhattan Project II; the name they picked emphasised speed and innovation while evoking the military influence central to the approach. Through OWS, the government used aggressive financial and regulatory commitments to reduce private sector uncertainty and let mature research be commercialized faster.

But the model does not fit every challenge. It would not necessarily accelerate novel consumer products such as smartphones or electric vehicles, where it is not clear in advance what consumers want or need. By contrast, for problems like carbon capture, where a verifiable amount of carbon must be pulled out of the atmosphere, an OWS model may be more valuable. The approach also needs to be applied to problems with positive spillovers that lack adequate market incentives: OWS's vaccines benefited not just the person receiving them but everyone made less likely

to be infected, and federal purchase commitments were essential to de-risk development by companies that did not know in advance whether their vaccines would work. Her closing examples of plausible targets are a universal coronavirus or flu vaccine and a resilient electrical grid.

Words to know

- **Operation Warp Speed** The US public-private partnership that funded, de-risked and distributed COVID-19 vaccines in 2020 and 2021. The piece treats it as a design to be taken apart rather than a slogan.
- **Emergency Support Function** A pre-planned structure setting out how federal agencies coordinate during a disaster. ESF 8, for public health and medical services, was activated early in the pandemic.
- **Defense Production Act** US legal authority to direct industrial production and allocate scarce materials. OWS used it for tubing, bioreactors and syringes.
- **Advance purchase agreement** A commitment to buy a product if it is produced, made before anyone knows whether it will work. OWS paid under most of these even if the vaccine never received authorization.
- **Emergency Use Authorization** A temporary FDA authorization to use an unapproved product in an

emergency. Companies continued their trials after receiving one.

- **Product Coordination Team** A single team per product, combining industry, manufacturing and government specialists, empowered to resolve needs without running an interagency process each time.
- **Persons in plants** Officers from the Army, BARDA and NIH posted inside manufacturing facilities to report problems on the ground in real time.
- **NEPA** A US law requiring environmental review procedures without mandating particular outcomes. D'Souza names it as an example of purely procedural regulation that a future programme might need to streamline.

Numbers in this reading

- 140,000: American lives OWS is estimated to have saved between December 2020 and May 2021.
- $1.8 trillion: the economic benefits attributed to OWS over the same period.
- $26 billion a day: the cost of COVID-19 between 2020 and 2021.
- Around $13 billion: the cost of OWS, which she puts at around twelve hours' worth of COVID-19's daily costs.
- Six industry partners across three platforms: Moderna and Pfizer-BioNTech on mRNA;

AstraZeneca and Johnson & Johnson on replication-defective live-vector; Novavax and Sanofi/GSK on recombinant-subunit-adjuvanted protein.

- January 10th, 2020: the date the SARS-CoV-2 sequence was made publicly available.
- Late 1980s: when research and development of mRNA therapeutics began maturing, well before the pandemic.
- Up to six months, cut to around three weeks: the FDA review backlog between Phase II and Phase III, normally and under OWS.
- 2024: the earliest a pan-variant COVID-19 vaccine was expected, according to the FastGrants funders' catalogue of roadblocks she quotes.

What the author is unsure about

D'Souza presents her evidence base openly: Paul Mango's book on OWS, written by one of its leaders, plus interviews with OWS staff, which means the account leans on insider testimony. The lives-saved and economic-benefit figures are given as estimates without a source or method. Her central claims about which features generalise are judgements, not findings, and she says so by framing them as what policymakers should do. She is explicit that persons in plants may be too expensive to justify outside an emergency, that the commitment levels seen in 2020 probably cannot be sustained in peacetime, and that

for most problems an OWS model would not pass a cost-benefit test. One number to hold loosely: she puts OWS's cost at around $13 billion, while other readings in this unit give $18 billion for the programme. Keep both figures as stated by their authors rather than picking one.

For the discussion

- The sharpest reusable point is that OWS commercialized mature technology rather than funding new science, which means any proposed Warp Speed for X should be tested first on whether the deliverable is explicit, bounded and achievable without a breakthrough.
- Her Customs-and-monkeys example is the clearest argument in the unit for mission authority over interagency coordination: speed failed in small, mundane places that no single agency owned.
- The regulatory compression is worth looking at closely, because cutting a Phase II review from up to six months to about three weeks is a change in queue position, not in standards, and that distinction is the one the course keeps asking us to hold.

5.3-2 · CORE

Progress Deferred: Lessons From mRNA Vaccine Development

Tim Hwang · 2024

Read the original paper → · the course assigns: whole piece

> **mRNA vaccines could have arrived years earlier, and what held them back was not missing science but three fixable frictions: a reputation problem, an unprofitable market, and discoverers who were not suited to selling their own discovery.**

Average summary

Hwang's question is simple and uncomfortable. The idea of using mRNA as a drug was proposed in **1988**. Katalin Karikó and Drew Weissman published the key discovery, that modifying mRNA lets it enter a cell without being rejected, in 2005. Operation Warp Speed then resolved most remaining design and manufacturing problems within months. If the science was largely in place, why did the world wait for a pandemic?

His answer names three structural frictions. First, **perceived viability**. DNA vaccines had promised much in the 1990s HIV effort and delivered little, culminating in Merck's 2007 trial failure, and mRNA looked worse still: fragile, hard to handle, apparently impossible to manufacture at scale. That reputation cut funding, academic interest and corporate support. Second, vaccines as an unprofitable field. Vaccines cost between $200 and $900 million to create, pass

clinical trials 6 to 11 percent of the time, and take an average of 10.71 years; before COVID-19 the fastest turnaround ever was Merck's mumps vaccine in 1967, at four years. Demand is unpredictable, margins thin, and many infectious diseases hit poorer countries hardest. Third, specialization: Karikó and Weissman were superb researchers and poor promoters. Their 2005 paper was rejected by leading journals as an incremental improvement, they received only two speaking invitations in the following years, and Penn licensed their patents to a small lab equipment company, Cellscript, for $300,000, while their own startup ran out of money and closed in 2009.

The recommendations follow the diagnosis. Increase the variance of research funding, with mechanisms such as golden tickets that let one enthusiastic reviewer fund a project against consensus, and last-shot funding for high-impact areas that are losing support. Address market failures in the scientific marketplace, by building common test beds, as OWS did when the FDA allowed adaptive parallel trials, and by funding the whole portfolio rather than the favourite, as OWS did across four platform technologies. And support research entrepreneurialism, by training science operators who partner with labs the way a chief operating officer would, and by making university technology transfer more flexible, noting that fewer than 5 percent of

technology transfer offices produce net positive returns for their university.

Hwang closes by generalising. These frictions are not unique to mRNA; neural networks followed a strikingly similar path. He argues for a metascientific playbook that raises the odds of capturing overlooked opportunities earlier, in any field.

Why it's assigned

Page 5.3 is the faster research and development option, and this is the historical half of that case. Where D'Souza's piece explains how Operation Warp Speed compressed the last year, Hwang explains why the clock started so late, and his claim is the stronger of the two for the Combat layer: the binding constraint on mRNA was resourcing, not knowledge. If that is right, then the biggest available gain in future pandemics comes from funding and institutional design in peacetime, which is exactly what the unit's financing and R&D arguments turn on.

In depth

The puzzle

mRNA vaccines are now being applied to HIV, MERS, malaria, tuberculosis and other conditions, and can be developed and manufactured faster and more cheaply than traditional vaccines. Yet the concept was first proposed in 1988, Karikó and Weissman's critical

discovery was published in 2005, and that paper itself built on observations dating to the 1960s. Much of the groundwork was available years before COVID-19. That Operation Warp Speed resolved many issues of vaccine design and manufacturing within months is, for Hwang, the telling fact: it suggests the primary limitation was resourcing rather than fundamental barriers of understanding or technology.

The stakes of the delay are not abstract. He footnotes an estimate by Bjoerkheim and Tabarrok that within US nursing homes alone, starting vaccinations five weeks earlier could have saved on the order of 14,000 lives, and ten weeks earlier 40,000.

Part I: a capsule history

The HIV detour. Traditional vaccine development introduces a weakened version of a virus so the immune system can gain experience. HIV broke that paradigm twice over: it compromises the immune systems it infects, and it mutates so fast that cultivating weakened versions risked producing something more virulent. Recombinant DNA offered an alternative, producing only harmless fragments of the virus, and subunit protein vaccines inserted selected DNA segments into a medium such as yeast. Through the 1990s the results from Genentech, MicroGeneSys and Chiron were disappointing, because HIV's changing structure made it hard to isolate envelope proteins the immune system could

reliably recognize. That work would later underpin the Novavax and GlaxoSmithKline COVID-19 vaccines.

Merck took a different route, using an adenovirus to carry DNA into cells so the body would produce HIV structural proteins and trigger a T cell response. Monkey trials in 1998 looked promising, a human trial began in 2004, and in **2007** the results were disastrous: the candidate had no effect, and some evidence suggested it had worsened infections among some recipients. Parallel NIH trials found the same. Merck shut down the trial and its entire HIV vaccine program.

The chilling effect. The failures produced a chill that reached well beyond HIV. The viral vector approach Merck had used would later form the basis of the Oxford-AstraZeneca and Johnson & Johnson COVID-19 vaccines, but researchers carrying that work forward faced years of skepticism. Dan Barouch and Jaap Goudsmit continued exploring adenovirus approaches at Crucell, later part of Johnson & Johnson, through the 2000s; negative feedback from management meant Goudsmit worked while keeping management in the dark, hiding results in lab notebooks under generic labels such as Assay Validations.

mRNA suffered the same chill, with an extra handicap. In the normal function of a cell, DNA sits in the nucleus and mRNA carries a temporary blueprint

outside it, which raised the possibility of generating the proteins needed for an immune response without entering the nucleus at all. Robert Malone, at the Salk Institute, proposed using mRNA as a drug as early as 1988, and Jon Wolff and collaborators showed in 1990 that mRNA could be directly used to produce proteins. Even so, Dan Wattendorf, who led the DARPA ADEPT program supporting mRNA work from 2010, found that most industry experts he consulted were skeptical, given their poor experience with viral vector and subunit protein vaccines. One article called it a fool's errand. Matt Winkler, founder of one of the earliest mRNA lab supply companies, said that if you had asked him then whether you could inject RNA into someone for a vaccine, he would have laughed in your face. The extra handicap was physical: mRNA is highly fragile, degrades in heat and is easily destroyed by common enzymes, which made it hard to work with and raised doubts about producing it at scale.

The researchers who kept going. Karikó and Weissman, at the University of Pennsylvania, investigated mRNA as a route to an HIV vaccine, publishing in 2000 that mRNA could produce the desired HIV proteins and in 2005 that modifying a component called uridine allowed mRNA to enter a living cell without rejection. Funding proposals to the NIH and elsewhere were denied, and the pivotal 2005 paper was rejected from leading journals for representing only an incremental improvement.

Unable to raise enough investment or afford Penn's licensing fees, they lost control of the patents, which the university licensed to Cellscript, a small lab equipment company, for $300,000.

Others worked in parallel. Eli Gilboa at Duke published in 1996 that mRNA could treat cancer tumours in mice; further funding from the dean of Duke's medical school was denied in 2005, on the sense that betting on mRNA was dangerous given the unsolved problems of stabilising the molecule and delivering it to cells. That research nonetheless inspired Ugur Sahin and Ozlem Tureci to found **BioNTech** in 2008 to explore mRNA as a cancer vaccine strategy. In 2010, Luigi Warren and Derrick Rossi showed mRNA could reprogram human skin cells into embryonic-like stem cells, work that became the basis for **Moderna**, which pivoted in 2013 to vaccines and significantly improved the lipid encasement technology that delivers mRNA to cells.

Skepticism persisted right up to the pandemic. A 2016 Nature Biotechnology article compared Moderna and its CEO Stéphane Bancel to Theranos and Elizabeth Holmes. By 2019 BioNTech faced a cash crunch as investors worried about its failure to bring any drug to market, and its October 2019 IPO raised little over half its original target.

The pandemic. At the start of COVID-19, mRNA vaccines remained a largely untested concept, with

promising laboratory results and some early human successes but no mass deployment. What made them attractive was speed: traditional methods require cultivating a weakened or deactivated virus over months in large vats of cells, while mRNA can be synthesized in a few hours from a gene of interest.

Even then, mRNA was not the favourite. The viral vector approach taken by Sarah Gilbert and Adrian Hill at Oxford was initially considered more promising because it was more mature, with experimental trials against MERS, flu and Zika suggesting it was effective and apparently safe. It was not until clinical trial results raised the possibility of harmful side effects for the Oxford-AstraZeneca vaccine, and showed lower efficacy than the Moderna and BioNTech vaccines, that mRNA was considered the superior option.

Operation Warp Speed, which allocated $18 billion, supported both. It accelerated mRNA in two ways. Upfront purchase commitments and research grants underwrote a major expansion in manufacturing capacity, which mattered especially for a startup like Moderna with no infrastructure to produce at scale, and that funding also drove technical progress, forcing resolution of practical problems from cooling to quality control. Second, it allowed the usual clinical trials to be combined and run in parallel, and allowed companies to manufacture candidates before they were proven safe and effective. The accelerated

process also resolved uncertainty fast: an untested, speculative option was identified as the most promising approach within months. The first American to receive a COVID-19 vaccination outside a clinical trial did so on December 13, 2020, with a Pfizer/BioNTech shot.

Part II: the three frictions

Perceived viability. mRNA progressed slowly through the 2000s and 2010s partly because it was widely thought to have an extremely low chance of success. Hwang is fair about this: given what was known, the assessment was arguably reasonable. Related HIV approaches had produced years of lacklustre results, and mRNA's fragility created practical problems at every stage.

The strength of the skepticism shows in its persistence. Economists preparing recommendations to the White House on allocating funds across COVID-19 vaccine methodologies were told by some that investment plans positing a role for mRNA were fanciful. Hwang argues the effect was compounding. Limited funding meant slower progress; slower progress fed doubts about the commercial viability of Moderna and BioNTech; and a thin evidence base made it hard to justify further investment. He calls this second effect epistemological uncertainty: it was simply difficult to get data on whether mRNA vaccines

worked at all or how they compared with alternatives, because clinical trials were scarce.

Vaccines as an unprofitable field. It is no accident, Hwang writes, that the leading mRNA companies were startups rather than established pharmaceutical firms. Vaccines can cost between **$200 and $900 million** to create and pass clinical trials only 6 to 11 percent of the time. The average vaccine requires a development timeline of **10.71 years**, and before COVID-19 the fastest turnaround ever, Merck's mumps vaccine in 1967, took four years. Profits are uncertain, because infectious diseases come and go and the market may vanish before the product arrives, and because many infectious diseases disproportionately hurt poorer nations, creating pressure for low prices.

His sharpest evidence is Merck's decision during the pandemic. It chose not to prioritize a COVID-19 vaccine partly because an earlier, expensive Ebola effort had won approval only after the 2014 outbreak faded; executives felt COVID-19 might similarly disappear. Even Moderna and BioNTech did not launch with vaccines as the goal, initially focusing on mRNA for a wide range of therapeutics. The conclusion Hwang draws is that if a broad research area is perceived as unprofitable, paradigm-changing research may never be commercialized at all.

Research versus entrepreneurship. Many groups were working on similar problems, often with more

resources than Karikó and Weissman, and Hwang notes it is not implausible another team could have made the uridine discovery. As it happened they did, and both appear to have been less suited to promoting and fundraising around their work. He quotes colleagues describing Karikó as brilliant but abrasive and Weissman as reserved to the point of seeming off-putting.

He traces two concrete costs. Their work failed to become common knowledge even among researchers working on mRNA therapeutics: Luigi Warren, working on cell reprogramming at Harvard Medical School in 2007 and 2008, struggled with exactly the problem Karikó and Weissman had solved years earlier, and was on the verge of giving up until a stray reference from an assistant professor at a nearby lab put him on the right track. In the several years after their 2005 paper, the pair received only two invitations to speak about the work. Second, they could not negotiate commercial terms around their intellectual property. Penn owned the patents and made the licensing decisions; years of negotiation for a license to their own discovery for their startup RNARx failed, the company burned through its limited NIH small-business funding and shut down in 2009, and Cellscript, which obtained the exclusive license, was introduced to Penn by Karikó herself. The contrast is Rossi, whose work was named one of Time magazine's top ten medical breakthroughs of 2010,

with Moderna launching that year and raising $240 million from AstraZeneca by 2013.

Part III: three recommendations

Hwang is clear there is no silver bullet, and offers three institutional changes that might have raised the odds.

1. Increase the variance of research funding. Pessimism about mRNA was grounded more in theoretical considerations and analogies to other approaches than in conclusive scientific failure. He proposes mechanisms that encourage funders to make higher-variance, heterodox bets: **golden tickets** that let a reviewer who feels strongly fund a project against peer consensus, and last-shot funding deliberately offered to potentially high-impact areas seeing declining funding and researcher activity. His evidence is who actually did fund mRNA during the lean years. DARPA's willingness, Wattendorf says, came from an organizational norm giving programme managers free rein. The Bill and Melinda Gates Foundation gave a $20 million grant to Moderna in 2016 and $55 million to BioNTech in 2019, partly from Gates's personal interest in vaccine technologies and partly because the foundation deliberately scanned adjacent scientific disciplines for overlooked methods.

2. Address market failures in the scientific marketplace. Established pharmaceutical companies had the talent, money and production know-how to accelerate mRNA and did not, because the risk of vaccine production and approval plus an inconstant market reduced the incentive. Since you cannot run an Operation Warp Speed every time market incentives slow high-impact innovation, Hwang highlights two transferable mechanisms. First, common test beds: in OWS the FDA permitted clinical trials to run in parallel and adaptively, with new treatments entering and ineffective ones exiting as evidence accumulated, which let everyone learn what worked far faster and lowered the cost of resolving uncertainty. Second, funding the entire portfolio in high-impact areas, knowing many bets will fail. OWS did not know in advance that mRNA would win; it made guarantees across four different platform technologies to mitigate the risk of failure on safety, efficacy, manufacturability or scheduling.

3. Support research entrepreneurialism. One objective of science policy, he argues, should be ensuring society can rapidly benefit from a breakthrough even when its originators are not placed to promote or commercialize it. But researchers should be left to research: the literature documents a trend toward specialization and larger scientific teams, and that is likely to continue. His two interventions are a trained class of **science**

operators who partner with labs and look after the practicalities of turning early research into a sustainable reality, something like a chief operating officer, building on models such as ARPA-E's tech-to-market advisors and ARPA-H's Project Accelerator Transition Innovation Office; and more flexible university technology transfer. On the latter he notes that Penn's rigid demand for a high upfront licensing fee helped kill RNARx, and that fewer than 5 percent of technology transfer offices have net positive returns for their university, which creates persistent pressure to cash out as soon as possible.

Lessons

Hwang's closing argument is that these frictions are not peculiar to mRNA. Low perceived viability, commercial disincentives and uneven researcher access to entrepreneurial resources appear across fields. Neural networks offer a highly parallel pattern: a major innovation reputationally tainted by early high-profile disappointments, returning to prominence decades later. That commonality suggests other high-impact innovations may be slowed right now by the same issues, that comparative work across fields would be valuable in building a taxonomy of recurring structural problems, and that there may be a metascientific playbook of interventions that raise the probability of capturing overlooked opportunities. He notes that bringing

mRNA vaccines to market only after a pandemic began had a cost measurable in lives lost, both during the pandemic and before it.

Words to know

- **mRNA** Messenger RNA, the temporary blueprint a cell copies from DNA and uses to make proteins. Its fragility is the practical reason it was dismissed for decades.
- **Viral vector** A harmless virus used to carry genetic material into cells. Merck's failed HIV vaccine used one, and so did Oxford-AstraZeneca and Johnson & Johnson.
- **Subunit protein vaccine** A vaccine made from selected fragments of a pathogen, produced by inserting DNA into a medium such as yeast. The 1990s HIV versions disappointed and the approach matured anyway.
- **Chilling effect** The way a high-profile failure suppresses funding and interest in a whole adjacent area, regardless of that area's merits. It is Hwang's main mechanism for the delay.
- **Golden ticket** A funding mechanism letting a single enthusiastic reviewer fund a proposal against the consensus of their peers, to raise the variance of what gets funded.
- **Adaptive platform trial** A trial in which treatments enter and leave as evidence accumulates. Hwang frames it as government-built

shared infrastructure that lowers everyone's cost of resolving uncertainty.

- **Technology transfer office** The university unit that licenses patents arising from its research. Fewer than 5 percent of them return more than they cost, which shapes how hard they bargain.
- **Science operator** Hwang's proposed role: a trained professional who handles promotion, fundraising and business-building for a lab, so researchers can keep researching.

Numbers in this reading

- 1988: when using mRNA as a drug was first proposed, by Robert Malone at the Salk Institute.
- 1990: Jon Wolff and collaborators showed mRNA could be directly used to produce proteins.
- 2005: Karikó and Weissman's uridine-modification paper, rejected by leading journals as an incremental improvement, and built on observations dating to the 1960s.
- $300,000: what Penn licensed their patents to Cellscript for.
- 2009: the year their startup RNARx shut down.
- Two: speaking invitations they received in the several years after the 2005 paper.
- 1998, 2004, 2007: Merck's promising monkey trials, the start of its human HIV trial, and the

disastrous results that closed its HIV vaccine programme.

- 2008 and 2010: the founding of BioNTech and the launch of Moderna; Moderna pivoted to vaccines in 2013 and raised $240 million from AstraZeneca that year.
- October 2019: BioNTech's IPO, raising little over half its original target.
- $18 billion: what Operation Warp Speed allocated, in Hwang's account, across four platform technologies.
- December 13, 2020: the first American vaccinated outside a clinical trial, with Pfizer/BioNTech.
- $200 to $900 million: the cost of creating a vaccine.
- 6 to 11 percent: how often vaccines pass clinical trials successfully.
- 10.71 years: the average vaccine development timeline.
- Four years: the fastest pre-COVID vaccine turnaround, Merck's mumps vaccine in 1967.
- $20 million to Moderna in 2016 and $55 million to BioNTech in 2019: Gates Foundation grants during the lean years.
- Fewer than 5 percent: university technology transfer offices with net positive returns.
- 14,000 and 40,000: lives Hwang's footnoted estimate suggests could have been saved in US

nursing homes if vaccinations had started five or ten weeks earlier.

What the author is unsure about

Hwang's argument is a counterfactual, and he treats it as one. He says the cumulative impact of negative funder and researcher views was likely significant though difficult to measure, and frames the alternative history as what would have been likely rather than certain. He concedes the skepticism was arguably a reasonable scientific assessment on the facts known at the time. He allows that even with better perceptions of mRNA, the number of companies willing to commercialize it would probably have stayed small because of the broader vaccine market. He notes it is not implausible another team would have made the key discovery. And he states that there is no single silver bullet policy intervention, offering his three changes as things that may have raised the chances rather than as guarantees. He also calls for further comparative work across fields before treating his playbook as general.

For the discussion

- The strongest version of Hwang's claim is that OWS did not invent anything; it paid for things that were already known, which makes the policy question how to pay for the next overlooked technology before the emergency rather than during it.

- Merck's reasoning about COVID-19, that the market might disappear as it had after Ebola, is the clearest single illustration in the unit of why vaccine development is an incentives problem before it is a science problem.
- His third friction is the most uncomfortable and the most testable: a discovery sat unused partly because the people who made it were not promoters, which suggests a cheap intervention, embedding operators in labs, with an unusually high expected return.

5.5-1 · CORE

Promoting versatile vaccine development for emerging pandemics

Joshua T. Monrad, Jonas B. Sandbrink and Neil G. Cherian · 2021

Read the original paper → · the course assigns: whole paper

> **If you pay for vaccines one pathogen at a time, you systematically underfund the platforms and prototype research that would let you respond to the pathogen nobody has met yet.**

Average summary

Monrad, Sandbrink and Cherian start from an observation about 2020. The first SARS-CoV-2 vaccines with phase 3 efficacy data, from NIH-Moderna, BioNTech-Pfizer and Oxford-AstraZeneca, were all **platform-based**, and all drew on work done before the virus existed. Their argument is that the funding system which produced that advantage did so almost by accident, because it pays for vaccines against named pathogens rather than for the capability to make vaccines quickly.

They define two complementary routes to faster pandemic vaccines. A vaccine is platform-based if a nearly identical mechanism, device, delivery vector or cell line is used for multiple target vaccines, which covers viral vectors, recombinant protein expression systems and nucleic acids, and distinguishes them from pathogen-specific approaches such as whole-inactivated or live-attenuated viruses. The **prototype pathogen** approach means doing countermeasure research on a representative member of a viral family, knowing the prototype may never emerge as a threat but assuming the techniques will apply to close relatives.

The economics section explains why private money falls short. Preventive products tend to be less lucrative than curative therapeutics even at equal social value, and vaccines create positive externalities by preventing transmission, so demand sits below the

socially optimal level. Vaccine sales rarely form a large share of revenue for leading producers, with a few exceptions such as Pfizer's multivalent pneumococcal vaccine and Merck's nonavalent HPV vaccine. The problem is worse for pathogens affecting lower-income populations and for sporadic epidemic pathogens where sustained demand is not guaranteed.

Then comes the argument that makes this paper distinctive. Existing fixes, push funding for early work and pull mechanisms such as **advance purchase commitments** for late-stage work, still value a project by the harm one named pathogen would do. But platform and prototype research creates positive technological spillovers whose value exceeds the value of any single countermeasure, so pathogen-specific financing structurally undervalues exactly the work that buys pandemic preparedness. Their worked example is MERS-CoV: fewer than three thousand cases were reported between 2012 and 2019, mostly in Saudi Arabia, with transmission apparently limited to close-contact healthcare settings, so any 2019 valuation of a MERS vaccine would have had a low ceiling. Yet prior MERS work supplied the prefusion-stabilised spike design and the safety data that accelerated SARS-CoV-2 vaccines.

Their solutions: extend pull mechanisms to include a term for spillover value, with larger rewards for the

first demonstration of safety and efficacy on a new platform or prototype pathogen; price dual-use risk into funding decisions; support precompetitive collaboration between firms; and have governments and public institutions do more of this work directly.

Why it's assigned

Page 5.5 is the financing and procurement option, and this is the paper that states the market failure most precisely. The Combat layer assumes that broad-spectrum and platform vaccines could be developed and stockpiled ahead of time; this explains why they mostly are not, and why the gap is not a failure of science or will but of how rewards are calculated. It is also the natural companion to the 100 Days Mission chapter on page 5.1, because the prototype pathogen and platform libraries CEPI proposes are exactly the things this paper says nobody has an incentive to fund.

In depth

The two routes to a faster pandemic vaccine

The paper opens with the COVID-19 timeline. Work on SARS-CoV-2 vaccine candidates began only after the novel strain was sampled and its genome sequenced in early January 2020. When the WHO declared a pandemic in March, many projections suggested regulatory authorisation and distribution could take several years. Less than 12 months later, the first

patients were being vaccinated outside clinical trials. The authors credit a concerted research effort, unprecedented public and private investment and overwhelmingly high global incidence, and then add the factor they care about: a set of vaccinological insights and technologies developed before the novel coronavirus was even identified.

Their working definition of a **platform** is a candidate in which an underlying, nearly identical mechanism, device, delivery vector or cell line is used for multiple target vaccines. That covers viral vectors, recombinant protein expression systems and nucleic acids, and sets them apart from the pathogen-specific approaches, whole-inactivated or live-attenuated viruses, that underlie most currently licensed vaccines.

The **prototype pathogen** approach means conducting countermeasure research for prototype pathogens, understanding that the prototype may not emerge as a threat but assuming the techniques would apply to closely related microorganisms. Such research targets prototypic members of viral families, genera or species known to pose risks to humans, prioritised by expected structural similarity to the next potential pandemic pathogen, and may involve identifying conserved antigen epitopes so insights transfer. It can also inform multivalent or universal vaccines that target a conserved antigen. The authors focus on viruses because of their potential for globally

catastrophic pandemics, while noting that emerging bacterial, parasitic and fungal pathogens matter too, particularly given antimicrobial resistance, and that the general approach can apply to them.

Why the market underprovides vaccines

For-profit companies conduct a large share of pharmaceutical R&D, especially late-stage trials, and a set of well-understood economic factors push private investment in infectious disease vaccines far below the socially optimal level.

- Preventive products tend to be less lucrative than curative therapeutics, even for products with the same potential to generate social value.
- Vaccines create powerful positive externalities by preventing community transmission, which means demand for them is often lower than the socially optimal level.
- Consequently, with the exception of a few blockbusters such as Pfizer's multivalent pneumococcal vaccine and Merck's nonavalent human papillomavirus vaccine, both with billion-dollar markets, vaccine sales rarely form a large share of revenue for leading producers.
- The problem is worse for pathogens affecting lower-income populations with less purchasing power, and for sporadic emerging epidemic pathogens where sustained demand is not

guaranteed and investment is particularly risky commercially.

The existing responses are financing by governmental agencies such as BARDA and the NIH, supranational bodies such as the European Commission, and philanthropies such as the Gates Foundation and Wellcome Trust, plus CEPI, founded after the 2014-2015 Ebola outbreak as a public-private partnership to coordinate funding for vaccines against epidemic diseases.

Those come in two kinds. **Push funding** means up-front grants for promising projects, suited to preclinical and early clinical work. **Pull mechanisms** reward output, through prizes or market guarantees, and suit late-stage trials and product development. Most programmes will involve phase 3 trials requiring significant capital, though the authors note that for some pathogens such as Nipah and MERS-CoV, outbreaks may never be large enough for conventional efficacy trials, so licensure may rest on animal data.

The standard pull mechanism is the **advance purchase commitment**, in which a funder precommits to buying doses conditional on successful licensing and production. It worked for pneumococcal vaccines and was used for SARS-CoV-2 candidates, including through COVAX.

The flaw in the advance purchase commitment

Here the paper makes a point that is easy to miss and hard to forget once seen. An APC ties the developer's reward to the quantity of doses procured. For an epidemic vaccine deployed early, when case numbers and geography are still limited, success may contain the outbreak before it grows, so the vaccine quells its own market because fewer doses are needed. The standard APC therefore does not strongly incentivise bringing a product to market proactively to pre-empt a pandemic. They quote Christopher Snyder and colleagues to the effect that when the best social outcome produces the worst economic outcome for the developer, there is a strong argument for a non-market intervention.

The fix they endorse is to delink revenue from doses used and set the reward to match the social value of developing the vaccine, counting both health and economic benefits. Snyder, Kendall Hoyt and colleagues from CEPI proposed a variation of the APC in which the maximum paid for successful development derives from the total social value of deploying the vaccine.

Spillovers, and why pathogen-by-pathogen funding fails

Even a social-value-based pull mechanism has a blind spot. The total social value of advancing a platform or doing prototype pathogen research exceeds the direct value of having a countermeasure for any one

pathogen, because the knowledge created produces positive technological spillovers. The authors are precise that they mean spillover in the economics sense, positive externalities of technological development, not zoonotic spillover.

Platform spillovers. Platform approaches can offer fast sequence-based design, testing, and rapid scalable manufacturing. Because vaccines on the same platform may differ only in the sequence encoding the target antigen, preclinical and early clinical work on one pathogen informs the general optimisation, formulation and delivery of the platform, so platform-based candidates can reach clinical testing faster. Phase 3 efficacy data on a novel platform tells you not only about that pathogen but about a range of pathogens with similar correlates of protection. Manufacturing facilities built for one nucleic acid vaccine may be repurposed for related vaccines, and large-scale manufacturing experience translates directly into speed and cost for the next vaccine on the same platform. The viral vector approach underpins two approved Ebola vaccines, Merck's recombinant vesicular stomatitis virus vaccine and Johnson & Johnson's human adenovirus 26 and modified vaccinia virus Ankara vaccine, and the Oxford-AstraZeneca chimpanzee adenovirus vector vaccine; nucleic acid platforms underpin NIH-Moderna and BioNTech-Pfizer. The Oxford-AstraZeneca vaccine and the Inovio and Imperial

College London candidates were built on platforms CEPI had funded over the preceding years.

The authors add two cautions. Different platforms suit different pathogens depending on tissue tropism and whether humoral or cellular immunity matters, so a broad portfolio must be maintained: RNA and viral vector vaccines can induce robust cellular responses that may be needed against enveloped viruses with complex pathogenesis such as poxviruses or filoviruses, while virus-like particles might suit neutralising antibody responses against non-enveloped viruses. And antigen-specific concerns such as disease enhancement or induction of autoimmunity must be addressed for each new vaccine, largely independently of previous experience with the same platform.

Prototype pathogen spillovers. Their central case is the coronavirus chain. Prior work on human betacoronaviruses confirmed that the spike protein is a prime target for neutralising antibodies. Earlier work on respiratory syncytial virus, parainfluenza virus and MERS-CoV showed that prefusion-stabilised immunogens may improve neutralising antibody responses against enveloped viruses, and before the pandemic researchers at the US NIH identified two proline substitutions, the **2P** substitutions, that stabilise the spike protein of MERS-CoV, SARS-CoV and HCoV-HKU1 in its prefusion conformation. That

stabilisation was applied to the SARS-CoV-2 spike, and the resulting S-2P prefusion-stabilised spike is used in candidates from NIH-Moderna, BioNTech-Pfizer, Novavax and Johnson & Johnson. A lipid nanoparticle formulation similar to that of mRNA-1273 had been tested for a MERS-CoV S-2P mRNA vaccine in mice, and the mouse model used in that study, based on CRISPR/Cas9-induced mutations in the dipeptidyl peptidase 4 receptor, had originally been developed for MERS-CoV. Clinical work transferred too: safety data from a phase 1 human trial of a chimpanzee adenovirus platform vaccine encoding the MERS-CoV spike enabled fast commencement of large-scale trials of the ChAdOx1-based AZD1222 vaccine for SARS-CoV-2, and prior work on a DNA plasmid vaccine for MERS-CoV aided Inovio's SARS-CoV-2 candidate.

The authors also note that earlier Flaviviridae work on yellow fever, Japanese encephalitis and dengue accelerated Zika vaccine candidates, and that efficacy trials against prototype pathogens may be critical for identifying the correlate of protection for a viral family, which then informs platform choice, antigen selection, formulation and delivery for a novel relative.

They add an important qualification about what makes a spillover public. Insights such as the promise of the spike glycoprotein as an antigen or information about animal models are likely to be published widely

and benefit everyone, which is a clear case for public intervention. But that depends on the work being accessible. Public funding of prototype pathogen work should promote openly accessible and translatable insights as far as practicable, while taking judicious advantage of proprietary intellectual property. Even where a proprietary insight mainly benefits the originating organisation, the research is still worth subsidising, because society benefits from having developers better prepared to respond.

Finally, they note the prototype approach is not limited to novel platforms: insights on antigen design, attenuation, optimisation, purification and manufacturing can transfer for live attenuated and inactivated vaccines too, with a synonymous codon replacement attenuation approach applied to both influenza A and RSV as the example.

The MERS illustration

The clearest statement of the thesis is a thought experiment. Suppose a funder in 2019 contemplated supporting a MERS-CoV vaccine, a virus already identified as a target for prototype research on the Coronaviridae family by US NIH researchers. If the funder ignored generalisable insights, they would underinvest. The authors are careful to say MERS's high fatality rate and history of significant hospital-acquired outbreaks warranted support on their own, and that both the WHO and CEPI rightly made it a

priority pathogen. But the epidemiology could have argued against larger investment: fewer than three thousand cases were reported between 2012 and 2019, the majority in Saudi Arabia, and human-to-human transmission appeared limited to close-contact healthcare settings, so any valuation of a MERS vaccine had its upper bound set by an apparent lack of truly pandemic potential. With hindsight, that work proved highly valuable for SARS-CoV-2, and any financing scheme that ignored the complementarity undervalued it.

The same logic applies to a general funder choosing between two projects for a WHO priority pathogen. If the decision is made purely to address that pathogen, the funder offers the same incentive for a pathogen-specific approach as for a platform-based one, despite the latter's spillovers. And the more useful a platform is across viruses, the larger the gap between private and social value, so the stronger the market failure and the better the case for non-market incentives.

Looking forward, they name Nipah virus, highlighted by US NIAID researchers as a prototype for the Henipavirus genus in the Paramyxoviridae family and already backed by CEPI, where regular outbreaks with some human-to-human transmission and the high mutational rate of RNA viruses raise the possibility of a future pandemic from a more transmissible relative.

Solutions

Stronger commercial incentives. Pathogen-specific funding should be complemented by efforts that maximise collective ability to address novel unknown pathogens. They point to CEPI's model, which requires potential grantees to propose plans for generating preclinical safety, immunogenicity and efficacy data for three different pathogens in addition to a phase 1 trial for at least one selected antigen, and which encourages grantees to develop their platform vaccines against known prototype pathogens. For pull mechanisms they propose modifying the Snyder social-value APC by extending the derivation to include a term for the indirect value of technological spillovers, quantified by considering the relative advantages of platforms, the pandemic characteristics of different pathogens, and the concrete potential for shortening timelines or raising the probability of success. They accept an exact quantitative valuation may be complex but argue relative merits can be evaluated and incentives set accordingly, and suggest rewards should be greater for the first time any manufacturer demonstrates safety and efficacy in phase 3 for a given platform or prototype pathogen.

They also fold in biosecurity. Funders should consider potential **dual-use concerns**: certain viral vector research, particularly work engineering viruses to evade antivector immunity, may produce insights with dual-use risk, which is a negative externality borne by

all of society rather than by the researcher. By explicitly incorporating that risk into funding decisions, for example by announcing larger rewards for robust approaches, funders can incentivise safer development, and they cite work by Sebastian Farquhar and colleagues showing dual-use risks can be quantified either through a market-based insurance model or a centrally commissioned risk assessment.

Beyond markets. Governments and philanthropists can support precompetitive collaborations between industry stakeholders, pooling expertise on shared advances such as improved assays, animal models and manufacturing processes, with the Biomarkers Consortium, the Patient-Reported Outcomes Consortium and the Innovative Medicines Initiative as precedents. The WHO is well placed to play a normative role in facilitating consortia and knowledge sharing, as are civil society organisations advocating equitable access, and public sequence repositories such as GISAID, where early SARS-CoV-2 sequences were shared, can accelerate the start of global development.

Governments can also act directly. Publicly financed universities already generate non-proprietary immunology and vaccinology insights, and public institutions have led development through all stages: the Public Health Agency of Canada designed the

rVSV Ebola vaccine later licensed to NewLink Genetics and acquired by Merck before authorisation in the EU, North America and Africa; the public Instituto Butantan in Brazil has been central to late-stage development of affordable dengue and influenza vaccines; and public agencies including the US NIH and the Chinese Wuhan Institute of Biological Products were central to SARS-CoV-2 vaccines. Public involvement can also promote equitable access through volume and pricing commitments for developing countries.

Conclusion

Biomedical pandemic preparedness needs investments in platforms and prototype pathogens that account for knowledge gains aiding vaccine development across multiple pathogens. The authors note that established vaccine types have real advantages, with known safety and lasting protection profiles from large-scale campaigns against measles, yellow fever, polio and influenza, and greater manufacturing and distribution capacity at least in the short term. They stress that accelerating R&D must be complemented by regulatory innovation to facilitate platform-based licensure, development of multivalent or universal vaccines, and work on equitable global access, plus regulatory innovation to allow emergency authorisation on the basis of animal data and, in some cases, controlled human infection models where

conventional trials would be unethical or infeasible. Realising the potential of platform vaccines also means solving manufacturing and delivery problems such as nucleic acid thermostability, cold chain storage and devices for DNA vaccine electroporation. Their closing claim is that investments enabling faster countermeasure development can provide net economic benefits on the order of trillions of dollars.

They declare that Neil Cherian is a Scientific Officer at CEPI, and that the other two authors have no competing interests.

Words to know

- **Platform vaccine** A vaccine in which nearly identical machinery is reused across targets, with only the antigen sequence changing. The reusability is the whole source of the spillover value the paper is about.
- **Prototype pathogen** A representative virus studied so that the techniques transfer to its relatives, accepting that the prototype itself may never cause an outbreak.
- **Push funding** Up-front grants that pay for research whether or not it succeeds, suited to early-stage work.
- **Pull mechanism** A reward for output, such as a prize or a purchase guarantee, suited to late-stage development.

- **Advance purchase commitment** A promise to buy doses if the vaccine is licensed and produced. Its weakness is that an early, successful epidemic vaccine shrinks its own market.
- **Technological spillover** Value created by research that accrues to others rather than the funder, in the economics sense rather than the zoonotic one. It is the paper's central concept.
- **Correlate of protection** A measurable immune response that predicts protection. Establishing one for a viral family through prototype work tells you what a vaccine for a novel relative must achieve.
- **Prefusion stabilisation and the 2P substitutions** A structural trick that locks a viral surface protein in the shape that provokes the best antibodies. Developed on MERS-CoV and relatives, then used in four SARS-CoV-2 vaccines.
- **Dual-use concern** Research whose results could be misused, creating a risk society bears rather than the researcher. The paper argues funders should price it in.

Numbers in this reading

- Fewer than three thousand MERS-CoV cases reported between 2012 and 2019, the majority in Saudi Arabia, with human-to-human transmission apparently limited to close-contact healthcare settings.

- Three different pathogens: the number for which CEPI requires grantees to plan preclinical safety, immunogenicity and efficacy data, in addition to a phase 1 trial for at least one antigen.
- Two proline substitutions: the 2P change that stabilises the prefusion spike, identified on MERS-CoV, SARS-CoV and HCoV-HKU1 before the pandemic and used in the NIH-Moderna, BioNTech-Pfizer, Novavax and Johnson & Johnson candidates.
- December 2020: when phase 3 efficacy data existed for the NIH-Moderna, BioNTech-Pfizer and Oxford-AstraZeneca vaccines, all platform-based, which the authors call a record speed by historical standards.
- Less than 12 months: from the March 2020 pandemic declaration, when projections ran to several years, to the first patients vaccinated outside trials.
- Trillions of dollars: the order of potential net economic benefit the authors attribute to investments enabling more rapid countermeasure development.

What the author is unsure about

The authors are explicit that platform technologies do not yet have the established track record of classical inactivation or attenuation approaches, and that DNA vaccines have only been licensed for veterinary use to

date despite decades of research. They say it remains to be seen whether any SARS-CoV-2 vaccine will provide long-term protection. They insist a broad portfolio of platforms must be kept because the relative usefulness of each varies by pathogen, and that antigen-specific risks such as disease enhancement or autoimmunity must be assessed afresh for every new vaccine regardless of platform experience. On their own proposal, they concede that obtaining an exact quantitative valuation of spillovers may be complex and fall back on relative comparison. They acknowledge that established vaccine types retain advantages in proven safety, durability and manufacturing capacity. And one author is a Scientific Officer at CEPI, whose funding model the paper holds up as a model approach.

For the discussion

- The sharpest single idea here is that an epidemic vaccine used early quells its own market, which means the standard advance purchase commitment pays least exactly when the vaccine does most good.
- MERS is the test case to bring: on 2019 evidence, fewer than three thousand cases in seven years, nobody could have justified the investment on that pathogen's own expected harm, and that investment is part of why COVID-19 vaccines arrived when they did.

- Their dual-use proposal is a direct bridge back to Unit 2, because it asks funders to price the risk that platform research creates, rather than treating safety and speed as separate conversations.

5.5-2 · FURTHER

Calculating the costs and benefits of advance preparations for future pandemics

Rachel Glennerster, Christopher M. Snyder and Brandon Joel Tan · 2023

Read the original paper → · the course assigns: from the abstract and introduction, the conceptual sections on programme design, the results and international sections, the conclusion, and every section heading

> **Future pandemics are expected to cost the world about $712 billion a year forever, and spending $60 billion now to build vaccine capacity before the next one would return more than $500 billion over waiting.**

Average summary

From the abstract, introduction, Section 4 on programme design, Sections 6 and 7 on results and on national versus international programmes, the conclusion, and all section headings. The modelling sections and appendix were not read line by line.

Glennerster, Snyder and Tan do two calculations. First, what future pandemics will cost. They combine Marani and colleagues' estimates of how often pandemics of each intensity arrive, drawn from a long series running from 1600 to the present and following a generalized Pareto distribution, with economic estimates of mortality, lost output and lost learning. Their answer: the expected present value of the stream of global social losses from future pandemics is **$17.8 trillion**, which at a 4 percent social discount rate is $712 billion every year going forward, more than the US Department of Defense budget in their baseline year. Under their most optimistic scenarios expected losses still exceed $350 billion annually, and under pessimistic but plausible ones they rise above $2.1 trillion.

Second, what it would cost to blunt that. Their programme spends about **$60 billion up front** and $5 billion a year thereafter to expand production capacity and supply-chain inputs, enough for 24 billion vaccine courses per year, which would let the world fully vaccinate 70 percent of the global population within six months of a new virus. Compared with the status quo, which they deliberately model as a sophisticated programme that still invests at risk but only once the pandemic has started, advance investment produces an expected net present value of $735 billion, a gain of more than

$500 billion, by reaching the coverage target about seven months earlier.

The design argument matters as much as the numbers. Most of the two-year lag between detecting COVID-19 and producing enough vaccine for 70 percent of the world was not discovery or approval but scaling production afterwards. Capacity can be built before approval, in parallel with trials, which they call at-risk investment; ordinarily that is wasteful, since vaccine programmes failed in clinical trials 70 percent of the time over the past two decades, but the value of speed justifies the waste. They also argue for contracting on capacity rather than on doses, because dose contracts put countries in a queue and capacity contracts speed the queue up.

They say no to the obvious objection that COVID-19 already prepared us. At peak the world made 580 million doses of mRNA vaccine a month, which would take over a year and a half to cover 70 percent of the population, and some of that capacity is being shut down as demand falls. Their closing line is that allowing it to be dismantled means we are failing to learn the lesson of COVID-19.

Why it's assigned

This is the Further reading on the financing and procurement page, and it is the one that puts a price on the whole argument of the Combat layer. The unit

keeps claiming that speed saves lives; this paper converts that into an expected loss per year, a programme cost, and a net present value, which is the form the claim has to take to move a treasury. It also supplies the one number the rest of the unit leans on without stating: how large the gap is between what preparation costs and what a pandemic costs.

Words to know

- **Net present value** The value today of a stream of future costs and benefits, after discounting. Everything in this paper is expressed this way so that a one-off investment can be compared with losses stretching into the future.
- **Social discount rate** The rate at which future social costs and benefits are converted into today's terms. The baseline here is 4 percent; at 2 percent the programme's gain rises above $1 trillion.
- **At-risk investment** Building manufacturing capacity in parallel with clinical trials rather than after approval, accepting that it is wasted if the vaccine fails. It is the mechanism that buys months.
- **Generalized Pareto distribution** The statistical shape the historical record of pandemic intensities follows, used here to put a probability on pandemics of different severities.
- **Positive spillover** A benefit one country's investment confers on others. Advance capacity

reduces a country's competition for scarce in-pandemic capacity, leaving more for everyone else.

- **Free riding** Waiting for others to invest and benefiting anyway. It is why the authors expect advance investment to be below optimal without an international agreement.
- **Kept warm** The G7's term for maintaining manufacturing capacity in working order between pandemics, which the paper's annual fee is designed to pay for.

Numbers in this reading

- $17.8 trillion: the expected present value of global social losses from future pandemics, equal to $712 billion every year going forward at a 4 percent discount rate, which they note is more than the US Department of Defense budget in the baseline year.
- More than $350 billion annually under the most optimistic scenarios; over $2.1 trillion annually under pessimistic but plausible ones.
- $13.8 trillion: the reduction in economic output from COVID-19 relative to pre-pandemic forecasts by 2024, per the IMF.
- 7 to 13 million: estimated excess deaths from COVID-19. $10 to $17 trillion: estimated loss of future productivity and earnings from school disruption.

- 2 percent of the world's population killed and GDP reduced by 6 percent: estimates for the 1918 flu. 30 percent of Europe's population: the Black Death.
- $1.3 trillion: the social value of accelerating capacity for 1.5 billion annual courses by just three months, per Castillo and colleagues.
- $60 billion up front and $5 billion a year: the programme's cost, funding capacity for 24 billion vaccine courses a year, enough to vaccinate 70 percent of the global population within six months.
- $735 billion: the expected net present value of the programme, a gain of more than $500 billion over the status quo, achieved by reaching target coverage about seven months earlier. The results section gives the gain as $504 billion.
- $32 billion: the reduction in expenditure during the next pandemic, because less capacity is left to install when short-run costs are high. Net of rental income from using the capacity between pandemics, expected expenditures are $27 billion higher and gross benefits $539 billion higher.
- Roughly two years: the lag between COVID-19 being detected and enough vaccine existing to fully immunise 70 percent of the world, most of it scale-up after approval rather than discovery or approval.
- 70 percent: the failure rate of vaccine programmes in clinical trials over the past two decades, which is what makes at-risk investment risky.

- $3 trillion: the global benefit if the Covid vaccine capacity available by April 2021 had been available three months earlier, per Athey and colleagues.
- $6 to $40 per course: what Covid-19 vaccines sold for, against an estimated $5,800 social value of a course of annual capacity in early 2021.
- 580 million doses a month: peak global mRNA production, which would take over a year and a half to cover 70 percent of the world with a full course.
- Always greater than $288 billion: the NPV gain under any single parameter change tested, including halving the probability of a pandemic arriving, cutting the value of a statistical life to $1 million, raising the discount rate to 6 percent, or setting the social cost of public funds at 30 percent.
- Nearly $1.3 trillion greater than the counterfactual if the truncation of pandemic intensity is relaxed; more than $1 trillion at a 2 percent discount rate.
- $113 billion up front for 45 billion annual courses raises the NPV gain to $705 billion; $30 billion for 12 billion courses gives an NPV of $534 billion, $303 billion more than the counterfactual.
- United States alone: $77 billion NPV, a gain of $61 billion, or $183 per capita. Brazil alone: $19 billion NPV, a gain of $15 billion, or $70 per capita.

What the author is unsure about

The paper is a working paper that has not been peer reviewed by the NBER board, and the authors say the views are their own and not the IMF's. They treat their own estimates as conservative and stress-test them hard: they report that the programme's advantage survives halving the arrival probability, zeroing out mortality losses entirely, and large changes to the discount rate or the social cost of public funds, and that reversing the conclusion would require extreme values, such as cutting the probability of pandemic arrival by a factor of 20. They are open that they chose an exogenous capacity level rather than optimising it, because a much larger programme might hit constraints outside the model, from government budgeting and international cooperation to a steeper marginal cost curve than their linear specification assumes. They acknowledge that full participation by all countries may not be achievable without an international agreement, that an insurance-style allocation carries moral hazard risk, and that redistribution from rich to poor countries has a limit beyond which participation stops being incentive compatible. Their model does not allow for differences in pandemic severity or timing across countries. And they note that contracts on reserve capacity would need monitoring, citing a case during COVID-19 where reserve capacity failed.

For the discussion

- The number worth carrying into the room is $712 billion a year of expected pandemic losses against a $60 billion up-front programme, because it reframes preparedness spending as cheap insurance rather than as a large expenditure.
- Their sharpest structural point is the difference between contracting on doses and contracting on capacity: dose contracts fight over a position in the queue, and capacity contracts make the queue move faster for everyone, which turns a negative spillover into a positive one.
- The closing observation is the most provocative one to raise: mRNA capacity built during COVID-19 is being converted or shut down as demand falls, which is exactly the behaviour their model says is most expensive, and they call it failing to learn the lesson.

5.5-3 · CORE

Expanding capacity for vaccines against Covid-19 and future pandemics: a review of economic issues

Susan Athey, Juan Camilo Castillo, Esha Chaudhuri et al. · 2022

Read the original paper → · the course assigns: sections I to V, with all three tables. Sections VI, VII and VIII were not read

> **Vaccine capacity was worth about $6,200 per annual course to the world and about $6 to $40 to the firms making it, and that gap is the whole reason public money has to buy capacity before anyone knows which vaccine works.**

Average summary

Summarized from a structured account of the article written by another helper while reading it, not from the original text, so nothing here is a direct quotation. It is also partial: sections I to V were read in full, including all three tables, while section VI on programme design, section VII on dose stretching and section VIII, the conclusion, were not read and are not summarized here. What is known of those three comes only from the paper's own roadmap in its introduction.

The authors review the economic case for public policy to accelerate vaccine supply in a pandemic. Their tagline is spending billions to save trillions. Each month of COVID-19 caused tens of millions of hospitalisations, hundreds of thousands of deaths and hundreds of billions of dollars of lost output, and more once school closures, job separations and indirect health harms are counted. Vaccines contained the worst of it, and investment was still too low.

Their core diagnosis is a gap between social and commercial incentives. Vaccine deals ran at $4 to $60 per course against an estimated social value of capacity of around **$6,200 per annual course**, and even a higher price creates no incentive for speed if it is the same whether doses arrive in a month or a year.

Section II makes the case with a deliberately simple example: a pandemic causing $1 trillion of harm a month for two years, ended once 70 percent of the population is covered. Baseline capacity of 3.2 billion courses a year avoids $14.0 trillion. Investing at risk, in parallel with trials rather than after them, brings availability three months sooner and adds $3 trillion. Doubling capacity cuts the time to 70 percent coverage from 21 months to 10.5 and adds more again. Detailed estimates from Castillo and colleagues give similar answers: $18.8 trillion for baseline at-risk capacity, with at-risk timing worth about $1,000 per annual course and doubling worth about $870 per additional course.

Section III explains why firms will not do this. A firm paid $40 billion for a billion courses earns the same whether it delivers in one month or twelve, while the faster rate needs twelve times the capacity investment. Deadlines slip, as in the EU's suit against AstraZeneca. Their preferred instruments are contracting on capacity rather than doses, and push

funding covering about 85 percent of investment so firms keep some stake.

Section IV asks how large a portfolio to fund. Operation Warp Speed spent $18 billion on six candidates, an order of magnitude less than the $70 billion on 15 to 20 candidates the authors' own 2020 op-ed recommended. Using an o-ring model where a candidate fails if anything fails at the disease, platform or subcategory level, they build portfolios with a greedy algorithm that maximises the chance at least one succeeds. Section V compares model to reality: investing at recommended levels would have given the world 7.12 billion annual courses of approved vaccines rather than roughly 3 billion, worth $1.7 trillion more.

Why it's assigned

This is the quantitative centre of the financing and procurement option. The Combat layer argues that speed is what saves lives, and this paper converts speed into money twice over: once as the social value of a course of capacity, and once as the portfolio decision a government has to make before it knows which vaccine works. It also supplies the diagnosis the other readings on this page assume, that a fixed price per dose provides almost no incentive to deliver sooner, which is the specific market failure a purchase guarantee alone does not fix.

In depth

Section I: the scale of harm and the core diagnosis

Each month of the pandemic caused tens of millions of hospitalisations, hundreds of thousands of deaths, and hundreds of billions of dollars in lost economic output, with the damage much greater once human capital losses from school closures and job separations, and indirect health harms such as impaired mental health and worsened comorbidities, are counted.

Vaccines were vital, and though development speed and manufacturing scale-up were unprecedented, public and private investment was likely still suboptimal given the social benefits of accelerating vaccination. Supplies were extremely tight early on even in high-income countries: the European Union sued AstraZeneca for delaying shipments to the EU while filling orders in its home country, the United Kingdom. Coverage in lower-income countries lagged well into the pandemic, with many failing to vaccinate even **5 percent** of their populations by the end of 2021. The authors note that surveys found more willingness to take COVID-19 vaccines in some lower-income countries than in some higher-income ones, which suggests scarce supply rather than scarce demand drove low initial uptake.

The diagnosis: the shortage traces to the gap between social and commercial incentives for capacity. Political and social pressure cap vaccine prices far below full social value. Prices in vaccine deals ran $4 to $60 per course, against an estimated social value of capacity of around $6,200 per annual course. And even higher prices would not supply full social incentives if the price is the same whether courses arrive in a month or a year.

Their headline preview is that **scale equals speed**: where capacity is the rate-limiting factor, doubling it halves the time to supply a given quantity, so cumulating losses stop that much sooner. Capacity also improves equity. With a year-long queue, doubling capacity cuts the wait of someone one month into the queue by two weeks, and cuts the wait of someone at the back by six months.

They set their own scope limits honestly. Vaccine capacity is one piece of a larger puzzle, since vaccinating someone requires manufacture, delivery and administration, and vaccines are one defence among surveillance, diagnostics, testing, drug therapies, protective equipment and social distancing. They focus on vaccines for their long record of high benefit-cost ratios and because administration may interfere less with normal life, and they narrow to production capacity because those decisions sit with manufacturers whose commercial interests may need

inducements, and because capacity carries more economic risk and lead time than other parts of a campaign. Their notion of capacity includes facility space plus enough staff to ramp up production and maintain quality control.

Section II: valuing capacity investment

The argument is built on a deliberately simple numerical example, illustrated in Figure 1. Assume the pandemic causes **$1 trillion of harm a month** from all sources and lasts two years. Assume a vaccine ends it once 70 percent of the population is covered, the herd immunity threshold many epidemiologists assumed early in COVID-19, a target the field later moved away from in favour of the widest possible coverage as more transmissible strains proved less preventable by vaccines than serious illness was. Assume the vaccine is approved at date 0 and takes three further months to install global capacity.

The figure divides the harm avoided into four regions: - **Region A**: harm avoided if the vaccine reduces monthly harm in proportion to coverage up to 70 percent, as with random rather than prioritised distribution. The example assumes enough capacity to cover 70 percent by the end of month 24, which is 21 months after approval. - **Region B**: additional harm avoided by better targeting, such as prioritising medical personnel or the elderly, or concentrating supply where infection is rising fastest. One cited

estimate holds that 80 percent of COVID-19 mortality reduction would come from vaccinating the first 20 percent of a population, which is why B's border is convex; they model it as a simple quadratic. - **Region C**: additional harm avoided by at-risk investment, building capacity in parallel with trials rather than after them. With three-month trials and investment starting soon after trials begin, availability comes three months sooner, so C is simply the $3 trillion gained by ending a $1 trillion-a-month pandemic three months earlier. - **Region D**: additional harm avoided if at-risk investment is made at double the scale. Instead of 21 months to 70 percent global coverage, doubling capacity cuts it to 10.5. Doubling does not double the harm avoided, mainly because the example already assumes enough capacity to finish vaccinating the target population before the pandemic ends, so doubling raises the speed of coverage rather than the eventual proportion covered.

Table 1 sets the example beside detailed estimates. In the numerical example, baseline capacity not at risk is 3.2 billion courses annually for a $14.0 trillion benefit; baseline capacity at risk is 3.2 for $17.0 trillion; double baseline capacity at risk is 6.4 for $20.5 trillion. In the Castillo and colleagues estimates, baseline capacity not at risk is 3.0 billion courses for $15.8 trillion; baseline at risk is 3.0 for $18.8 trillion; double at risk is 6.0 for $21.4 trillion. The baseline capacity figure in the example comes

from 7.9 billion global population times 70 percent coverage, divided by 21 months to reach that coverage, times 12 months per year.

The detailed estimates were built from announced bilateral deals with manufacturers, from which Castillo and colleagues derived a global capacity range centred on 3.0 billion courses annually, projected rollout in proportion to deals signed with high-, middle- and low-income countries, and accounted for nonlinear benefits across demographic groups using epidemiological evidence, with some allowance for externalities across groups.

The per-course figures are the ones to carry. Baseline at-risk capacity installed globally saved $18.8 trillion, about **$6,200 per course of annual capacity**. Investing at risk rather than waiting three months accounts for a further $3.0 trillion, about $1,000 additional benefit per annual course accelerated. Doubling at-risk capacity to 6.0 billion annual courses would have raised global benefit by $2.6 trillion, about $870 per additional annual course.

The authors draw a methodological lesson too: a well-chosen numerical example can come close enough to detailed analysis to inform decisions that must be made in hours during a crisis rather than days or weeks, since both routes show that baseline capacity generates tens of trillions in surplus and that accelerating and expanding it is worth trillions more.

But they caution that toy examples are not perfect substitutes, because the $1.8 trillion gap between the example's $17.0 trillion and the detailed $18.8 trillion is small in percentage terms and not a surplus to ignore in absolute terms.

Section III: commercial versus social incentives

The social values above are several orders of magnitude larger than the $6 to $40 per course manufacturers actually received, which the authors read as an enormous gulf between social and commercial incentives.

Why price overstates commercial incentive: a firm contracted to supply a billion courses at $40 earns the same $40 billion whether it delivers in one month or one year, but the faster rate needs **twelve times** the capacity investment, an expense the firm would rather save when it brings no extra revenue. Whatever the level of a fixed price, from the observed $6 to $40 up to figures approaching the thousands of dollars of social value, its fixity removes much of the incentive for speed. With several competing vaccines a firm might expand capacity to steal sales that would otherwise have gone to rivals, but such business-stealing incentives may be weak in a pandemic with limited industry-wide supply.

Why fixed prices also discourage at-risk investment: success probabilities from phase 1 to phase 3 are

notoriously low. In the Ahuja and colleagues sample of COVID-19 candidates, even the most promising was given less than a **29 percent** chance of success and others less than 10 percent. Rather than sink an investment likely to fail and hard to recoup because it may not be fungible, firms prefer to establish safety and efficacy first, and with the same price whether sold today or in three months there is little reason to invest early. Against that, the social value of at-risk investment is reported here as running as high as $1,200 per annual course.

The section then weighs three policy options.

Delivery deadlines. Without penalties these slip. The EU's suit against AstraZeneca followed delivery of only a third of the contracted 300 million doses by the specified date, with the company citing production delays. The EU sought damages of 10 euros per dose per day of delay; the judge instead relaxed the deadlines and capped the penalty at 10 euros per dose, not per day, far below the social cost of stalling an EU campaign heavily dependent on AstraZeneca. The authors note that even damages on the order of the EU's social loss would not work, because the firm would go bankrupt before paying in full, which dulls incentives, and that bonuses for early delivery would have to be huge and raise the risk of firms sacrificing quality for speed.

Contracting on capacity rather than doses. Contracting on doses may simply have a firm fill orders in a slow trickle from existing capacity, leaving the buyer at the end of a long queue, as the EU was behind the UK for AstraZeneca supply. Where dose contracts put countries in a Prisoner's Dilemma with negative externalities, capacity contracts can generate positive externalities, because once dedicated capacity has served the contracting country it can serve others. Instead of fighting over a position in the queue, capacity contracts speed the rate at which the queue moves. The caveat: if inputs to capacity expansion are severely limited, capacity contracts just shift the Prisoner's Dilemma back a stage, from fighting over scarce doses to fighting over scarce capacity, which the authors say only strengthens the case for expanding input capacity, stockpiling storable inputs and making supply chains more resilient in advance.

Push versus pull funding. Push funding defrays capacity expenditure whether or not a successful product results; pull funding is paid only on successful delivery. Push funding is particularly good at inducing at-risk investment, since a firm will invest before knowing whether the vaccine works if the buyer carries most of the risk. Its problems are that it needs accounting control, gives little incentive to economise on investment or improve the chance of success, which is moral hazard, and does little to deter firms

with no realistic chance from joining, which is adverse selection. The authors therefore favour a mix: covering around **85 percent** of investment cost with push funding so firms keep some stake, complemented by pull funding in the form of a purchase guarantee lucrative enough to make firms contribute their share.

Section IV: the optimal vaccine portfolio

The question is how much spending can be justified when at-risk investment must be made under considerable uncertainty about which vaccines will succeed, and when the opportunity to accelerate later capacity is greatest early in a pandemic.

The benchmark is Operation Warp Speed, launched after the authors' 2020 op-ed, which spent **$18 billion on six candidates**, an order of magnitude less than that op-ed's recommendation that the United States spend $70 billion to develop and procure 15 to 20 candidates. Elsewhere the paper states that Ahuja and colleagues imply over three times Warp Speed's amount should have been spent, on 27 candidates.

The programme they analyse works like this. Each selected candidate invests at risk, in parallel with trials, in enough capacity to fully vaccinate the eligible US population, say 250 million courses, over about three months. The programme covers 85 percent of that at-risk investment with push funding, leaving 15

percent as the firm's stake. Successful vaccines are promised a price per dose carrying a mark-up over production cost, calibrated to induce the optimal number of candidates to participate, which means the bonus must let the marginal participant expect to cover its 15 percent share given its possibly low success probability. If the marginal candidate's probability of success is 20 percent, the bonus price must be five times its per-unit contribution to at-risk investment. The bonus is the pull component, paying for outputs rather than inputs, which mitigates moral hazard and adverse selection.

Expected benefits equal the probability that at least one vaccine succeeds times the gain from successful at-risk investment, since each firm builds enough capacity to vaccinate the country and the country is assumed to be fully vaccinated eventually either way, just faster with the programme, so the gain is region C of Figure 1.

To model correlated failure they use an **o-ring structure**, in which a candidate fails if a problem occurs at any level: the disease level, since vaccines might not have worked for COVID-19 had it proved as hard as malaria; the platform level, a particular worry for mRNA and DNA, never before used in a production vaccine; and the subcategory level, where within viral vectors adenovirus might fail while measles-vector vaccines remain viable. Failure at one level knocks out

every candidate branching from it, creating positive correlation in failure. Just as negative correlation in returns is desirable across assets, negative correlation in failure is desirable across vaccines, because it raises the chance of at least one success.

Selection uses a **greedy algorithm** that sequentially adds the candidate contributing most to the probability of at least one success. Its ranking does not match a ranking by stand-alone probability: the algorithm picks the mRNA candidate before an inactivated-virus candidate with a higher individual probability, because an inactivated-virus candidate is already in the portfolio.

Table 2 lists the top 20 candidates in order of greedy selection, with the cumulative probability that at least one in the portfolio succeeds. The first six show the shape: an inactivated, standard, phase 3 candidate takes the portfolio to 28.8 percent on its own; a viral vector adenovirus phase 3 candidate adds 19.6 points to 48.4; an mRNA LNP-encapsulated phase 3 candidate adds 10.0 to 58.4; a second inactivated phase 3 candidate adds 7.4 to 65.8; a protein subunit recombinant phase 2 candidate adds 5.0 to 70.8; and a protein subunit S protein phase 2 candidate adds 3.7 to 74.5. By the twentieth candidate the cumulative probability is 87.3 percent and each addition is adding 0.2 points.

The model inputs behind those numbers, as published in the table notes: the full list considered by Ahuja and colleagues covered all 142 candidates under development by an August 2020 cut-off defining the ex ante period; a 10 percent chance of a problem at the overall level of COVID-19 vaccines; a 20 percent chance of a problem in each platform, raised to 40 percent for mRNA and 60 percent for DNA because of their untried status; and a 20 percent chance of a problem in each subcategory. Within a subcategory an individual vaccine fails with a probability that falls the further it has progressed: 86 percent preclinical, 77 percent in phase 1, 68 percent in phase 2, 50 percent in phase 3.

Figure 2 compares the greedy algorithm with a naive method that selects candidates in order of individual success probability, shown as the mean of 1,000 repetitions. The curves are concave and fairly close together, which the authors read as evidence that the overall shape does not depend on the algorithm's sophistication.

Costing a portfolio: expected cost equals push funding for the programme's 85 percent share of at-risk capacity investment plus expected bonus payments, and the bonus price is endogenous, since it must be lucrative enough for the marginal entrant, the one with the lowest success probability, who discounts the bonus most heavily. Their worked

comparison: for the optimal five-candidate portfolio the marginal candidate has an 18.4 percent success probability, and taking round numbers of $10 per annual course of capacity with 85 percent push-funded, leaving the firm $1.5 billion, assuming prices fall to production cost outside the bonus period and the bonus is paid on the first 250 million accelerated courses, the bonus price must be **$33** for the firm to break even. For the optimal 15-candidate portfolio the marginal candidate's probability is 9.2 percent, half the five-candidate figure, so the bonus must double to $66. Expected programme costs come to $53.2 billion for five candidates and $177.0 billion for fifteen.

Why the bigger portfolio is still worth it: the extra $123.8 billion is justified because the 15-candidate portfolio's probability of at least one success is 15 percentage points higher. As long as a three-month acceleration of vaccination is worth at least $825 billion, a 15 percent higher chance of that gain covers the extra cost, and the estimated gain is about double that threshold, $1.6 trillion in the detailed calculations and more in the numerical example.

Table 3 extends the analysis to countries and income categories. Across all countries: 30 total distinct candidates, a mean of 8.8 per country, 2,290.0 million courses per month installed at risk and 538.9 million realised, 0.3 courses per capita per month installed,

$1,033.7 billion in expected benefits ($137.4 per capita) against $274.8 billion in expected costs ($36.5 per capita). High-income countries: 30 distinct candidates, mean 18.3, 1,418.0 million installed, $828.7 billion benefits ($699.3 per capita) against $170.2 billion costs. Middle-income: 18 distinct, mean 6.7, 906.9 million installed, $216.9 billion benefits ($40.7 per capita). Low-income: 18 distinct, mean 1.3, just 2.3 million installed, $0.6 billion benefits, which is about 60 cents per capita against 30 cents of cost. Among individual countries, the United States shows 27 candidates, 462.3 million courses per month installed, $301.7 billion in benefits or $923.4 per capita, against $55.5 billion in costs; Germany, the United Kingdom and Australia all sit near 1.3 to 1.4 courses per capita with per-capita benefits between $763 and $880; Chile has 12 candidates, 0.6 courses per capita and $183.0 per capita in benefits.

The table notes matter: these benefits reflect the gains from accelerating capacity availability by three months, not the difference between having the portfolio and not having it, and the category entries aggregate individual country programmes rather than representing the optimal programme for a coalition.

Section V: model versus reality

Comparison is hard, because capacity investments in failed candidates are seldom compiled, so actual at-risk investment is hard to observe, and the realised-

capacity figure in Table 3 is a counterfactual expectation that may differ from actuals because of errors in forecasting success probabilities.

A bound can still be drawn. Comparing optimal with actual capacity for only the vaccines ultimately approved, globally AstraZeneca, Pfizer, Moderna, Sinovac and Sputnik, Ahuja and colleagues estimated that investing at the recommended levels would have given the world capacity for **7.12 billion annual courses** of approved vaccines instead of the roughly 3 billion estimated from actual production runs, adding $1.7 trillion in benefits. The authors call this an underestimate, since it ignores capacity from vaccines that would have succeeded had they been funded.

On low-income countries, the small optimal portfolio, averaging 1.3 candidates, and low net benefits of about 30 cents per capita rationalise why most did not secure bilateral at-risk deals early, but do not explain persistently low purchases later. The 30-cents figure reflects a tiny portfolio, any member of which may fail, and only a three-month acceleration; the gain from buying vaccines once proven, measured per course rather than per capita, is orders of magnitude higher. In new runs of the Castillo model for this paper, capacity sufficient to vaccinate 40 percent of the population in every low-income country within a year, allocated to the highest-value consumers in

each country, would have been worth **$350 per annual course** at the start of 2021.

On supply versus demand in low-income countries, the authors say it may never be settled which mattered more. Low-income countries seem to have suffered lower mortality losses throughout, perhaps because of younger populations. The lack of orders at Africa's first vaccine factory may suggest low demand rather than low capacity limits vaccination there now, but waning demand as the pandemic wanes does not prove demand was low at the outset, and one cited argument holds that greater demand would have materialised had loan programmes been available earlier.

On why high-income countries still fell short of recommended investment without facing financing constraints, they offer two explanations and a suspicion. First, the mathematics of optimisation: slightly under- or overshooting an optimum costs little, and halving the budget of the recommended US at-risk programme would still have delivered **91 percent** of the full-budget net surplus. Yet even that half budget would more than double both Operation Warp Speed's total vaccine spending and the number of candidates funded, which implies even the relatively generous programme substantially undershot the optimum. Second, the recommendations may themselves have been too

high, since the portfolio calculations assumed a dollar of expenditure costs a dollar, whereas if spending must be raised through distorting taxes the calculations may have underweighted saving money. Against that, they suggest officials may have given undue weight to factors with more personal political cost than true social cost: avoiding the appearance of handouts to big pharmaceutical firms or foreign suppliers, of backing projects that fail, or of wasting money on products left unused in a suddenly waning pandemic.

What was not read

Section VI, further details on programme design; section VII, dose stretching; and section VIII, the conclusion, were not read and are not summarized. From the paper's own introduction, section VI covers combining push and pull funding and contracting on capacity rather than doses; section VII covers a first-doses-first strategy and fractional dosing, with an option-value argument for testing or rolling them out experimentally and reversing if the benefits do not appear; and section VIII covers implications for COVID-19, lessons for future pandemics, and speculation on why actual country procurement fell short of the models. The fractional dosing argument is set out in full in reading 5.6 #3, by an overlapping group of authors.

Words to know

- **At-risk investment** Building capacity in parallel with clinical trials rather than after approval. The paper's central instrument, and the thing a fixed dose price gives firms no reason to do.
- **Push funding** Money that defrays investment whether or not a product succeeds. It induces at-risk investment but creates moral hazard and adverse selection, which is why they pair it with pull.
- **Pull funding** A reward paid only on successful delivery, such as a bonus price per dose. Here it is calibrated so the least promising candidate in the portfolio still expects to break even.
- **O-ring structure** A failure model in which a problem at any level, disease, platform or subcategory, kills everything below it. It makes failures positively correlated, which is what makes platform diversity valuable.
- **Greedy algorithm** A selection rule that repeatedly adds whichever candidate most increases the chance that at least one succeeds, which is not the same as ranking candidates by their own odds.
- **Moral hazard and adverse selection** Two problems with paying for inputs: firms have less reason to economise or to raise their chance of success, and firms with no realistic chance still want in.
- **Prisoner's Dilemma** The situation created by dose contracts, where each country's rational move

leaves everyone worse off. Capacity contracts are proposed as a way out of it.

- **Social value of capacity** What a course of annual production capacity is worth to the world, estimated here at about $6,200, against the $6 to $40 manufacturers received.

Numbers in this reading

- $4 to $60 per course in vaccine deals, against an estimated social value of capacity of about $6,200 per annual course. Section III gives the price manufacturers actually received as $6 to $40.
- Fewer than 5 percent of their populations vaccinated by the end of 2021 in many lower-income countries.
- Numerical example: $1 trillion of harm a month, two years, 70 percent coverage threshold, approval at date 0, three months to install capacity, 21 months after approval to reach 70 percent, cut to 10.5 months if capacity is doubled. Baseline capacity of 3.2 billion annual courses comes from 7.9 billion population times 70 percent divided by 21 months times 12.
- Table 1, numerical example: 3.2 billion courses not at risk, $14.0 trillion; 3.2 at risk, $17.0 trillion; 6.4 at risk, $20.5 trillion. Castillo estimates: 3.0 not at risk, $15.8 trillion; 3.0 at risk, $18.8 trillion; 6.0 at risk, $21.4 trillion.

- $6,200 per annual course for baseline at-risk capacity; about $1,000 extra per annual course from investing at risk rather than waiting three months; about $870 per additional annual course from doubling. Section III gives the social value of at-risk investment as running as high as $1,200 per annual course.
- $1.8 trillion: the gap between the example's $17.0 trillion and the detailed $18.8 trillion for the same policy.
- 80 percent of COVID-19 mortality reduction from vaccinating the first 20 percent of a population, in one cited estimate.
- Twelve times the capacity investment to deliver a billion courses in one month rather than one year, for the same revenue.
- Less than 29 percent for the most promising candidate and less than 10 percent for others: success probabilities in the Ahuja sample.
- 300 million doses contracted, a third delivered by the specified date; the EU sought 10 euros per dose per day of delay and the judge capped it at 10 euros per dose.
- 85 percent push funding, 15 percent firm stake. A marginal candidate with a 20 percent success probability needs a bonus price five times its per-unit contribution.
- $18 billion on six candidates for Operation Warp Speed, against the authors' recommendation of

$70 billion on 15 to 20 candidates, and Ahuja's implied figure of over three times Warp Speed on 27 candidates.

- 250 million courses over about three months: the capacity each candidate builds in the modelled programme.
- Greedy portfolio: first candidate 28.8 percent, then 48.4, 58.4, 65.8, 70.8, 74.5 cumulative, reaching 87.3 percent by the twentieth.
- O-ring inputs: 142 candidates by the August 2020 cut-off; 10 percent problem probability at the disease level; 20 percent per platform, 40 percent for mRNA and 60 percent for DNA; 20 percent per subcategory; individual failure of 86 percent preclinical, 77 percent phase 1, 68 percent phase 2, 50 percent phase 3.
- Five-candidate portfolio: marginal candidate 18.4 percent, bonus price $33, expected cost $53.2 billion. Fifteen-candidate portfolio: marginal candidate 9.2 percent, bonus price $66, expected cost $177.0 billion. The extra $123.8 billion buys 15 percentage points more probability, justified if a three-month acceleration is worth at least $825 billion, against an estimated $1.6 trillion.
- Table 3, all countries: 30 distinct candidates, mean 8.8 per country, 2,290.0 million courses per month installed at risk, 538.9 million realised, $1,033.7 billion benefits ($137.4 per capita) against $274.8 billion costs ($36.5 per capita).

- High income: mean 18.3 candidates, $699.3 per capita benefits. Middle income: mean 6.7, $40.7 per capita. Low income: mean 1.3, about 60 cents per capita in benefits against 30 cents in costs.
- United States: 27 candidates, $923.4 per capita benefits against $169.8 per capita costs. Chile: 12 candidates, $183.0 per capita.
- 7.12 billion annual courses of approved vaccines under recommended investment, against roughly 3 billion actually achieved, a difference worth $1.7 trillion.
- $350 per annual course: the value at the start of 2021 of capacity sufficient to vaccinate 40 percent of the population in every low-income country within a year.
- 91 percent: the share of the full-budget net surplus that halving the recommended US at-risk budget would still have delivered.

What the author is unsure about

They are careful about the status of their own numbers. Failure risks were estimated from academic sources plus discussion with medical and industry experts, and they stress that substantial uncertainty remains, that different probability estimates select different optimal portfolios, and that improving the probability estimates may add as much surplus as improving the selection mechanism. They note their toy example and their detailed model differ by $1.8

trillion on the same policy, which is small proportionally and large absolutely. They flag that capacity investments in failed candidates are rarely compiled, so actual at-risk investment cannot be observed directly and the realised-capacity column is a counterfactual expectation. They say it may never be settled whether supply or demand mattered more in low-income countries. And they offer two competing explanations for why rich countries underinvested, one of which concedes their own recommendations may have been too high because they assumed a dollar of public spending costs a dollar. They also revise the 70 percent herd immunity assumption themselves, noting the field moved away from it. Finally, this summary itself rests on sections I to V only; the paper's own conclusions in section VIII have not been read.

For the discussion

- The cleanest formulation of the market failure is that a firm paid the same price whether it delivers in a month or a year has to spend twelve times as much to go fast, so speed is pure cost to the manufacturer and pure benefit to everyone else.
- The portfolio section gives a defensible answer to the objection that funding fifteen candidates wastes money on fourteen failures: the extra $123.8 billion buys a 15 percentage point higher chance of a gain worth about $1.6 trillion.

- Capacity contracts versus dose contracts is the most transferable idea here, because it converts a scramble for queue position into something that makes the queue move, and it is the same insight the advance purchase commitment critique on page 5.5 arrives at from the other direction.

5.5-4 · CORE

COVAX explained

Seth Berkley · 2020

Read the original paper → · the course assigns: whole piece, read for financing and procurement (the same source is assigned again at 5.6 #2, read there for manufacturing and delivery)

> **COVAX was built as a shared purchasing mechanism: pool the money of rich and poor countries into one portfolio of vaccine candidates, so that nobody's access depends on whether their own bet comes in.**

Average summary

Berkley, writing in September 2020 as head of Gavi, explains the financing machinery of COVAX, one of three pillars of the Access to COVID-19 Tools Accelerator launched that April and coordinated by Gavi, CEPI and the WHO. Its initial aim was **2 billion doses available by the end of 2021**, enough to

protect high-risk and vulnerable people and frontline healthcare workers.

The logic is a portfolio one. There were then more than 170 candidate vaccines in development and most would fail: on previous experience, candidates at the preclinical stage have roughly a 7 percent chance of succeeding and those in clinical trials about 20 percent. COVAX assembled the largest and most diverse portfolio, nine candidates already in development and a further nine under evaluation, so that participants would gain access to whichever ones proved safe and effective.

Two separate pots sit inside it. The **COVAX Facility** is where self-financing economies buy in, and where pooled purchasing power is used to negotiate prices and to pay manufacturers to expand production before approval. The **Gavi COVAX Advance Market Commitment** is a wholly separate fund for the 92 middle- and lower-income countries that cannot pay, financed mainly through official development assistance, philanthropy and the private sector, and not cross-subsidised by the Facility's participants.

The purchasing terms are the detail worth knowing. Self-financing countries can request doses for 10 to 50 percent of their populations, and choose one of two arrangements. A Committed Purchase means guaranteeing to buy an agreed volume, in exchange for a lower upfront payment of US$1.60 per dose, or

15 percent of the total cost per dose, with the right to opt out if a vaccine's price turns out to be twice or more what was expected. An Optional Purchase lets a country decline any particular vaccine without losing its share of others, but costs more upfront: US$3.10 per dose plus a risk-sharing guarantee of US$0.40 per dose, to protect the Facility against liabilities if participants walk away after it has contracted with a manufacturer. The total cost ends up the same either way, since the Facility passes through the negotiated price plus a speed premium and a small operating fee.

Allocation is by rule rather than by bidding. Once a portfolio vaccine is approved, doses go to all participating countries at the same rate, proportional to total population, with about 5 percent held back as a buffer for acute outbreaks and humanitarian use. No country receives enough to vaccinate more than 20 percent of its population until every country in the financing group has been offered that amount.

At the time of writing, 78 higher-income countries had confirmed interest, with legally binding commitments due by 18 September and upfront payments by 9 October, and the AMC had raised about US$700 million of a US$2 billion seed target needed by the end of 2020.

Why it's assigned

Page 5.5 is the financing and procurement option, and this is the primary document for the one real-world attempt to buy pandemic vaccines collectively. Read next to the economics papers on this page it is the concrete instrument their arguments are about: a pooled advance purchase mechanism, with an explicit price for commitment and an explicit price for optionality. For the Combat layer it also answers a question the vaccine readings leave open, which is who pays, and in what order people get the thing we spent a year proving works.

In depth

The problem Berkley says COVAX was built for

His framing is that it became clear early in the pandemic that ending the crisis needs not just vaccines but access to them everywhere, which prompted global leaders to call for a solution that would accelerate development and manufacture of vaccines, diagnostics and treatments and guarantee rapid, fair and equitable access to them in all countries. COVAX is the vaccines answer, and he describes it as the result of an extraordinary global collaboration with more than two-thirds of the world engaged.

COVAX is one of three pillars of the **Access to COVID-19 Tools (ACT) Accelerator**, launched in April 2020, which brings together governments, global

health organisations, manufacturers, scientists, the private sector, civil society and philanthropy. COVAX is the vaccines pillar, coordinated by Gavi, the Coalition for Epidemic Preparedness Innovations and the WHO. It acts as a platform supporting research, development and manufacturing of a wide range of candidates, and negotiating their pricing, with all participating countries having equal access regardless of income level. The initial aim was 2 billion doses available by the end of 2021, enough to protect high-risk and vulnerable people and frontline healthcare workers.

Berkley sets out two different value propositions. For lower-income funded countries, and for a number of higher-income self-financing countries without bilateral deals, COVAX is a lifeline and the only viable way their citizens get access. For the wealthiest self-financing countries, some of which are also negotiating bilateral deals, it is an insurance policy that protects their citizens twice over: directly, by increasing their chances of securing doses, and indirectly, by reducing the chance of resurgence through making sure the rest of the world gets doses too.

The portfolio argument

COVAX exists, he writes, because without it there is a very real risk that the majority of people in the world would go unprotected, letting the virus and its impact

continue unabated. To prevent that you first need vaccines that are safe and effective, which is by no means a certainty.

The numbers behind that caution are the clearest financing argument in the piece. There were then more than **170 candidate vaccines** in development, and the vast majority of those efforts were likely to fail. Based on previous vaccine development, candidates at the preclinical trial stage have roughly a **7 percent** chance of succeeding, and those that make it to clinical trials about **20 percent**. To raise the odds, COVAX created the largest and most diverse portfolio of candidates, with nine already in development and a further nine under evaluation.

By joining, both self-financing and funded countries gain access to that portfolio as and when candidates prove safe and effective. Self-financing countries are guaranteed sufficient doses to protect a certain proportion of their population, depending on how much they buy in. Subject to funding availability, funded countries receive enough doses to vaccinate up to **20 percent** of their population in the longer term. Because demand is initially likely to exceed supply, allocation is spread across countries based on the number of doses available, and increases as availability increases.

The two mechanisms

Gavi created the **COVAX Facility** through which self-financing and funded economies participate, and inside it sits an entirely separate funding mechanism, the **Gavi COVAX Advance Market Commitment**, which supports access for lower-income economies. Together they make participation possible for all countries regardless of ability to pay.

The Facility's role is to maximise the chances of people in participating countries getting access as quickly, fairly and safely as possible. Participants get not just the largest and most diverse portfolio but an actively managed one: the Facility continually monitors the vaccine landscape to identify the most suitable candidates based on scientific merit and scalability, and works with manufacturers to incentivise them to expand production capacity in advance of regulatory approval.

Berkley states the economic reason this matters in a single sentence that connects directly to the other readings on this page. Normally manufacturers are reluctant to risk the significant investments needed to build or scale up vaccine manufacturing until they have approval for a vaccine. But in a pandemic costing the global economy **US$375 billion every month**, that reluctance would inevitably mean delay and initial shortages once vaccines are licensed. So the Facility provides investments and incentives to make manufacturers ready to produce the moment a

vaccine is approved. It also uses the collective purchasing power that comes from having many countries participate to negotiate competitive prices, which are passed on to participants.

Self-financing participants can request doses sufficient to vaccinate between **10 and 50 percent** of their populations, and pay in proportion to what they request. For these countries the Facility is a critical insurance policy that significantly increases their chances of securing vaccines even if their own bilateral deals fail, and by pooling resources they also help increase the world's chances of getting vaccines quickly and in the quantities needed. Berkley is careful to add that there are no guarantees any candidate will ultimately succeed, and that the point of a global approach is to share the risks so the world can share the rewards.

At the time of writing, **78 higher-income countries and economies** had confirmed their interest in participating, with more possibly to follow. Countries had until **18 September** to commit to legally binding agreements and until **9 October** to make their upfront payments.

Committed versus Optional Purchase

This is the procurement heart of the piece. Self-financing countries joining the Facility choose between two arrangements.

A **Committed Purchase Arrangement** requires a guarantee to procure an agreed volume of doses through the Facility. In exchange for that firm commitment, participants provide a lower upfront payment of **US$1.60 per dose, or 15 percent of the total cost per dose**. They are effectively committing to buy a set number of vaccines that, once available, will be fairly and equitably allocated among participants. They retain the ability to opt out of purchasing a vaccine if its price turns out to be twice or more what was expected.

An **Optional Purchase Arrangement** lets participants opt out of receiving any particular vaccine without jeopardising their full share of doses of other candidates, subject to supply. This may suit participants that already have bilateral agreements through which they have secured enough doses of that particular vaccine. The trade-off for that greater choice is a higher proportion paid upfront: a down payment of **US$3.10 per dose** and a risk-sharing guarantee of **US$0.40 per dose**, which protects the Facility against liabilities arising when participants decide not to purchase a candidate after the Facility has already contracted with the manufacturer. There is a second cost: by opting out of vaccines allocated to them and waiting for another, these countries may experience a delay in receiving their full committed volume.

In the end the total cost for the vaccines is the same for the two options. As a pass-through facility, participants pay the amount negotiated by the Facility, plus a speed premium invested in accelerating and scaling up manufacturing, plus a very small operating fee. Some manufacturers will provide vaccines at flat prices while others will tier prices based on income levels.

The Advance Market Commitment

The Gavi COVAX AMC's primary focus is ensuring that the **92 middle- and lower-income countries** that cannot fully afford to pay get equal access to vaccines as higher-income self-financing countries, and at the same time. Its funding is entirely separate from the Facility's, which means the AMC is in no way cross-subsidised by the funds of self-financing participants. Instead it is funded mainly through Official Development Assistance, together with contributions from the private sector and philanthropy.

At the time of writing, the AMC had raised about **US$700 million** of an initial seed capital target of **US$2 billion** needed by the end of 2020. Berkley's ask is therefore directed at sovereign donors, who need to decide the extent to which they will contribute or allocate ODA toward that target, and to do so in a timely manner.

How doses are allocated

Once a portfolio vaccine has successfully completed clinical trials, proved safe and effective, and received regulatory approval, available doses are allocated to all participating countries at the same rate, proportional to their total population size. A small buffer of about **5 percent** of the total number of available doses is kept aside to build a stockpile to help with acute outbreaks and to support humanitarian organisations, for example to vaccinate refugees who might not otherwise have access.

The ceiling is the rule that matters most. Even though self-financing participants can request enough doses to vaccinate between 10 and 50 percent of their population, no country receives enough doses to vaccinate more than **20 percent** of its population until all countries in the financing group have been offered this amount. The only exception is countries that have opted to receive fewer than 20 percent.

What Berkley asks for next

He calls the speed at which a comprehensive global solution was built remarkable, and says implementation now hinges on countries buying into the Facility so that it can make urgent investments. Having many self-financing economies sign up means work can begin on formal agreements with manufacturers and developers to secure doses, which both increases the chances of successful development and ensures the productive capacity

exists to manufacture the needed volumes the moment a vaccine is ready.

Two funding gaps remain. On research and development, CEPI leads the COVAX work with nine candidates supported, eight of them already in clinical trials; governments have committed **US$1.4 billion** and an additional **US$1 billion** is still needed to move the portfolio forward. And the AMC must meet its fundraising target of at least US$2 billion by the end of 2020, while continuing to discuss with AMC-eligible economies what participation will mean for them, because otherwise ability to pay becomes a barrier to access and the majority of the world is left unprotected.

Words to know

- **COVAX** The vaccines pillar of the ACT Accelerator, coordinated by Gavi, CEPI and the WHO, which pools country money into one portfolio of candidates and allocates the successful ones by rule.
- **ACT Accelerator** The April 2020 global collaboration covering diagnostics, treatments and vaccines. COVAX is its vaccines pillar.
- **COVAX Facility** The pooled purchasing mechanism self-financing economies pay into, which negotiates prices and pays manufacturers to scale up before approval.

- **Advance Market Commitment** Here, the separate Gavi fund for 92 middle- and lower-income countries, financed by development assistance and philanthropy rather than by other participants.
- **Committed Purchase Arrangement** Guaranteed volume in exchange for a lower upfront payment of US$1.60 per dose or 15 percent of cost per dose.
- **Optional Purchase Arrangement** The right to decline a particular vaccine, bought with a higher upfront payment of US$3.10 per dose and a US$0.40 per dose risk-sharing guarantee.
- **Speed premium** The additional amount built into the pass-through price, invested in accelerating and scaling up manufacturing.
- **Official Development Assistance** Government aid spending, the main intended source of AMC funding.

Numbers in this reading

- 2 billion doses available by the end of 2021: the initial aim, enough for high-risk and vulnerable people and frontline healthcare workers.
- More than 170 candidate vaccines in development; roughly a 7 percent chance of success at the preclinical stage and about 20 percent for those in clinical trials.
- Nine candidate vaccines already in the COVAX portfolio and a further nine under evaluation.

- 10 to 50 percent of their populations: the range self-financing countries may request doses for.
- Up to 20 percent of their population in the longer term: what funded countries receive, subject to funding availability.
- US$375 billion every month: what the pandemic was costing the global economy, the figure Berkley uses to justify paying for capacity before approval.
- 78 higher-income countries and economies had confirmed interest; legally binding agreements due 18 September, upfront payments due 9 October.
- Committed Purchase: US$1.60 per dose upfront, or 15 percent of the total cost per dose, with an opt-out if a vaccine's price is twice or more what was expected.
- Optional Purchase: US$3.10 per dose down payment plus a US$0.40 per dose risk-sharing guarantee.
- 92 middle- and lower-income countries covered by the Gavi COVAX AMC.
- About US$700 million raised of a US$2 billion seed capital target needed by the end of 2020.
- About 5 percent of available doses held back as a buffer stockpile for acute outbreaks and humanitarian use.
- 20 percent: the ceiling on any country's population coverage until every country in the financing group has been offered the same.

- CEPI's R&D work: nine candidates supported, eight in clinical trials, US$1.4 billion committed by governments and US$1 billion still needed.

What the author is unsure about

Berkley is the chief executive of Gavi, one of the three organisations running COVAX, so this is an advocacy document as well as an explanation, and the tone is a pitch to countries that had not yet signed. Within that he states real uncertainty. Safe and effective vaccines are by no means a certainty, and he gives the base rates to prove it. He says plainly that there are no guarantees any candidate will ultimately succeed, which is the reason for the portfolio. He describes the dose figures as subject to funding availability, and the deadlines, participation counts and fundraising totals are as of 3 September 2020 and were overtaken by events. What he does not discuss is what happens if wealthy countries buy heavily through bilateral deals while also participating, which is the failure mode later readings in this unit describe.

For the discussion

- The price list is the clearest statement in the unit of what optionality costs: US$1.60 per dose if you commit, US$3.10 plus a US$0.40 guarantee if you want the right to say no, for the same eventual total.

- The 20 percent ceiling is a real rationing rule rather than a sentiment, and it is worth asking whether a rule like that can survive contact with bilateral deals made outside the Facility.

5.6-1 · CORE

Essential Medicines Supply Chain and Manufacturing Resilience Assessment

Advanced Regenerative Manufacturing Institute (ARMI) / NextFAB, with ASPR · 2022

Read the original paper → · the course assigns: whole report, focused on its findings and recommendations for supply chain and manufacturing resilience. The appendices listing the 86 medicines and the public comments were not read in detail

> **The United States cannot reliably supply its own essential medicines because the generics that matter most are the least profitable to make, and the report's answer is to pay for resilience deliberately rather than expect a lowest-cost market to produce it.**

Average summary

This is an action plan commissioned by the US Department of Health and Human Services, through the Office of the Assistant Secretary for Preparedness and Response, and produced by ARMI with Nexight Group. It follows a February 2021 White House

Executive Order on supply chain resilience, which directed a 100-day supply chain review, and builds on an August 2020 order directing the FDA to identify essential medicines and critical inputs. The method was deliberately industry-led: about 40 stakeholder interviews across material suppliers, manufacturers, group purchasing organizations, wholesale distributors and healthcare providers; more than 80 completed surveys; two workshops each with a clinical steering committee of over 40 people and an industry steering committee of over 30; public comments from October 2020; and review of more than 100 sources.

It produces two things. First, a prioritised list: clinicians narrowed the FDA's essential medicines list to **86 critical medicines** needed for acute patient care, meaning rescue and lifesaving use in intensive care, cardiac care and emergency departments, stabilising patients for discharge, and urgent or emergency surgery.

Second, six challenge areas and four groups of solution strategies. The challenges are market structure, where most essential medicines are generics with low returns and just-in-time inventory magnifies any upstream interruption; global competition, where foreign government subsidies and lower operating costs have moved production offshore, with 87 percent of API manufacturing sites and 63 percent of finished dosage form sites located

overseas as of March 2021; labour and workforce; manufacturing processes, where at least 63 percent of US drug shortages between 2013 and 2017 traced back to quality issues; supply chain and distribution, where companies will not share data; and regulation, where setting up a new plant takes 5 to 7 years and adding a line to an existing plant takes 3 to 5.

The solution strategies are grouped as supply chain coordination, security and transparency; expanded onshore or nearshore production capacity; advanced manufacturing capabilities and innovative research and development; and purchasing, stockpiling and distribution approaches. Within each, the industry steering committee marked the first two as high priority, which gives a clear ranking: share data across the supply chain and strengthen government-industry coordination; guarantee stable long-term demand for generics at a sustainable price and offset capital investment; invest in innovative manufacturing paradigms and in workforce; and use federal buying power to reform procurement while revising purchasing models to reward quality and resilience rather than lowest cost.

Three case studies ground it in the ward: injectable dexamethasone, albuterol inhalers and fentanyl injection, each of which went short during COVID-19.

Why it's assigned

Page 5.6 is the manufacturing and delivery option, and this is the reading that takes the question off the vaccine and onto everything else a hospital needs to treat the people a pandemic produces. The Combat layer tends to be told as a vaccine story; this report is about the sedatives, steroids and inhalers that keep patients alive before any vaccine exists, and about why the US could not make them. It also supplies the mechanism the rest of the unit assumes: capacity that is cheap to run at full tilt and impossible to surge.

In depth

What the report is and how it was made

The White House issued an Executive Order on supply chain resilience in February 2021, directing key federal agencies to conduct a 100-day supply chain review. ASPR was tasked with identifying risks in the supply chain for pharmaceuticals and active pharmaceutical ingredients, and with policy recommendations to address them. That built on an August 2020 Executive Order directing the FDA, with federal partners, to identify a list of essential medicines, medical countermeasures and critical inputs medically necessary to have available at all times, in adequate amounts and appropriate dosage forms.

ASPR then asked ARMI and NextFAB for a deeper analysis, carried out through the National Forum to

Secure America's Supply Chain for Essential Medicines. The evidence base is industry and clinical testimony: about 40 interviews, more than 80 surveys, four workshops across clinical and industry steering committees, a public comment period, and review of more than 100 sources. The report was developed specifically for the Biden-Harris Administration to inform its investments in domestic manufacturing capacity.

The vision is stated in three verbs: **anticipate** issues quickly to ensure safety, quality and predictability, with a centralized control tower platform and a dashboard of integrated data, forecasting of supply and demand, and strategic alliances; **prepare** for risks, with supply redundancy and multi-sourcing, sufficient emergency inventory, clear distribution logistics and contingency plans; and **respond** to shifting demand, with source shifting, flexibility, and adequate workforce and distribution capacity.

Two notes on scope. The downselected list assessed small molecules and therapeutic biologics, so blood and blood products, fractionated plasma products, vaccines and volume expanders were excluded because their supply chains differ, as were categories not needed for routine acute care such as biological, chemical and radiologic threat countermeasures and COVID-19 and pandemic influenza countermeasures. And the report notes other challenges exist, such as

managing wastewater and industrial waste from pharmaceutical manufacturing, which were only discussed inside the six challenge themes.

The 86 medicines

Clinical stakeholders narrowed the FDA's list to the 50 to 100 medicines considered by consensus most critical for typical acute patient care, defined as rescue or lifesaving use in intensive care units, cardiac and coronary care units and emergency departments; stabilising patients in hospital continued care to enable discharge; and urgent or emergency surgery. The exercise produced a downselected list of **86 critical medicines**, listed in Appendix A, intended to let government target medicine-specific next steps, with strategies tailored to each medicine's own supply chain and vulnerabilities.

The six challenges

Market structure. Most essential medicines are generics, which offer low return on investment when not offset by external incentives. Originator brand manufacturers often decline to make essential medicines whose patents and exclusivities have expired. Low margins drive current manufacturers out of the market, and those who remain have more difficulty consistently meeting demand. The thin margins present a high-risk, low-reward outlook for

new entrants considering building new facilities or maintaining and upgrading existing ones. The same applies to drug accessories such as sterile injectable packaging, which is capital-intensive, low margin, and tends to be produced in facilities already running at near full capacity with little room for equipment downtime. Pressure to lower costs also makes manufacturers de-emphasize quality management maturity and continuous improvement, because thin margins inhibit investment in redundant supply and new equipment. And **just-in-time inventory** practices, widely used to lower storage costs, limit safety stocks and on-hand inventory, so any upstream problem becomes a shortage quickly. COVID-19 illustrated that directly, with sudden demand spikes, facility closures and import delays.

Global competition. Some countries, notably China and India, have invested heavily in subsidizing manufacture of medicines and active pharmaceutical ingredients, which makes it hard for US manufacturers to compete without similar support. Offshore facilities have lower labour rates and operating costs, and less stringent environmental and safety regulations also lower costs. In the case of China, the ability to manufacture APIs at lower cost through economies of scale has allowed it to force foreign competitors out of the market. As of March 2021, **87 percent of manufacturing sites making APIs and 63 percent producing finished dosage forms were located**

overseas, and a single country, China, controls most of the world's fermentation-based API production, having pursued strategies such as dumping products below production cost. Even for APIs with onshore or nearshore sources, the raw chemical components of most APIs are overwhelmingly produced in China or India, so a shortage in one precursor can hit seemingly unrelated medicines that share an ingredient. The report adds a realistic caveat: producing all precursor materials in the US may be unrealistic, since some require rare components only available where the natural deposits are.

Labour and workforce. The US faces a shortage of workers with specialized knowledge for pharmaceutical manufacturing. An estimated **60 percent of US manufacturing jobs will go unfilled between 2015 and 2025** because of talent shortages. The workforce lags international competitors in STEM degrees, many positions need additional specialized training in areas like biologics, formulation and advanced manufacturing, and training new employees is slow, which is exactly the wrong property in a surge. Competitive wages in other US STEM fields also push wages up beyond what the generics market can bear, which encourages manufacturers to move offshore.

Manufacturing processes. Traditional methods make quality problems harder to detect and slower to fix. At

least **63 percent of US drug shortages between 2013 and 2017** could be traced back to quality issues. When contamination is detected on a traditional aseptic processing line for sterile products, the line must be shut down and cleaned, which can take weeks to months and often causes a shortage. Generic medicines often run multiple products on one production line as a cost efficiency, so a quality or cross-contamination concern shuts the whole line, and the shared line also limits the ability to increase production for a surge. The generics market generally lacks the incentives and support for innovations such as **continuous manufacturing** and inline monitoring, which could address quality problems, because it cannot attract the investment cutting-edge technologies need, and there is a gap in government funding for carrying technologies from academic laboratories to commercial scale. Purchasers also usually cannot see where a medicine was made or how mature the maker's quality systems are, so manufacturers lack any market reward for improving them. Finally, surge capacity is structurally hard: manufacturers must operate close to full capacity to stay viable, equipment cannot easily be switched on and off as demand changes, with some equipment needing to run continuously and other equipment breaking down or becoming obsolete during long disuse, and increasing capacity also needs additional staff plus the time to hire and train them.

Supply chains and distribution. The chains are complex, with intermediaries such as wholesale distributors adding markup, and insurance and purchasing practices by insurers, pharmacy benefit managers, pharmacy services administrative organizations and group purchasing organizations affecting supply or availability. GPOs say they reduce drug costs for hospitals by consolidating operations, but their use can limit the number of suppliers hospitals can choose from and drive down manufacturer margins. Maintaining multiple redundant sources improves resilience but costs money in keeping suppliers qualified and audited. Data sharing is the deeper problem: companies are protective of supply chain data, volumes produced by different companies are often not public, so it is unclear whether capacity and supply are balanced; sources of raw materials are often not disclosed, creating the impression of more diversity than exists; and a lack of standardization in company names across sites and countries means one company under several names can make a supply chain look more resilient than it is. The consequences are concrete: an inability to predict and prevent shortages, purchasers unable to assess one manufacturer's resilience against another's, clinicians who do not know why a medicine is on backorder, and uncoordinated responses that produce hoarding, waste from expired products and the bullwhip effect of conflicting signals travelling up and down the chain. Stockpiling has its

own problems. There is uncertainty about what should be stockpiled federally versus regionally or by individual organizations, some components such as vials are hard to stockpile because sizes and equipment are not standardized, stockpiles are hard to maintain because they are rarely used and finished dosage forms have limited shelf life, rotation is impractical for products with no active market, and excess inventory needs space and staff.

Regulations. The report is careful to say the rigorous US regulatory environment promotes a high standard of quality and safety, and then names the unintended constraints. Manufacturers find it easier to keep using the suppliers they used at the time of approval, which reduces flexibility if a supplier has a shortage or is tied up in emergency production. Gaining approval for updated production processes is time consuming and expensive, which discourages adopting more efficient processes as technology advances. More stringent US environmental regulations can make overseas manufacture more cost effective, since some API and finished-dose antibiotic processes release high levels of pollutants that raise local health risks including antimicrobial resistance, and manufacturers often move abroad rather than seek technological solutions to minimise waste. And lead times are long: setting up a new plant can take **5 to 7 years**, and adding a single line in an existing plant **3 to 5 years**, including

meeting Current Good Manufacturing Practice requirements.

The case studies

Injectable dexamethasone. After the June 2020 UK study showed dexamethasone could reduce deaths in COVID-19 patients receiving respiratory support, demand rose sharply: a **610 percent increase in hospital orders in June 2020**, and a **137 percent increase in price in October 2020**. Most manufacturers were forced to backorder the drug, causing multi-month delays. The WHO advised providers to prioritise the most serious cases, and treatment guidelines allowed alternative glucocorticoids such as prednisone, methylprednisolone or hydrocortisone where dexamethasone was unavailable. The FDA worked with manufacturers and expedited applications related to increasing supply, and added dexamethasone sodium phosphate to the list of medications permitted to be temporarily compounded by outsourcing facilities and pharmacy compounders during the pandemic. This is the direct sequel to the RECOVERY readings on page 5.4: the trial found the drug, and the supply chain could not deliver it.

Albuterol inhalers. After nebulizers were identified as a potential source of airborne spread, inhalers became a common hospital alternative, producing shortages that affected the nearly **25 million** people

in the United States with asthma plus patients with other lung diseases. Demand for inhalers rose **400 percent** early in the pandemic, and CDC recommendations encouraging asthma sufferers to seek extended supplies made it worse. In an Asthma and Allergy Foundation of America survey, **46 percent** of respondents had trouble accessing prescription asthma or allergy medication during the pandemic, and sufferers were advised to use partially effective expired inhalers or ration their use. The FDA approved the use of generic albuterol inhalers for the first time in April 2020, which helped.

Fentanyl injection. Fentanyl became critical for patients on ventilators, who could be ventilated for up to two weeks. Demand rose in April 2020 as hospitals needed large 50-millilitre vials to treat ventilated patients humanely. Supply constrained its use. When one Colorado hospital system could not find vials large enough to fill a bedside drip bag in one go, pharmacy technicians cracked as many as **25 thumb-size glass bottles** to draw out enough liquid for a single 50 millilitre intravenous bag for one ventilator patient, who might need two or more bags a day. Before the pandemic the DEA had called for cutting authorized production of fentanyl by 31 percent; in early April 2020 it reversed that with an order for a 15 percent increase. A potential future mitigation named is allowing batches to ship pending sterility testing results, with final release once results meet criteria.

The solution strategies

A long-term government commitment, through budgetary, regulatory or policy action, is needed to change the current paradigm, and many solutions will have to be combined. In each of the four groups below, the first two strategies were marked by the industrial steering committee as high priority.

Increased supply chain coordination, security and transparency. 1) Improve data and information sharing, integration and standardization across the entire supply chain, through public-private partnerships creating shared data infrastructure, mechanisms for companies to share data while protecting intellectual property, consistent standards, use of existing data such as provider inventory and wholesale distributor data, required reporting of key information like inventory levels to government at higher frequency, and stronger oversight of foreign manufacturing through better data sharing with the FDA. 2) Strengthen collaboration and coordination between government and all supply chain participants, by clearly defining roles for FDA, FEMA and CDC and creating a National Command Center or Federal Health Authority, monitoring for over-consolidation of GPOs, reviewing the essential medicine list comprehensively every two years in line with the WHO cycle and with enough notice for industry to adjust, and establishing an essential medicine consortium including an open capacity

consortium that tracks which manufacturers have spare capacity for which product. 3) Develop predictive tools and risk assessments, including a global supply chain assessment covering source location, volume and number of facilities for APIs, finished dosage forms and components such as filters, stoppers and vials; an index of APIs with resilience scores based on quality management maturity, country of origin and business continuity; and mandated, anonymised, aggregated transparency on inventories to create a virtual stockpile. 4) Establish a more comprehensive centralized control tower platform, including a national critical drug tracking, monitoring and alert system with regional inputs, quarterly federal reporting and an early alert system for API shortages. 5) Increase physical and cyber security posture, including designating the pharmaceutical industry and its supply chain as a high-priority critical infrastructure sector.

Expanded onshore or nearshore production capacity. 1) Ensure stable, long-term demand for generic medicines at a sustainable price, through long-term price and volume guaranteed contracts that protect investment from low-priced foreign imports, and a pricing floor for generics on the essential list to prevent unsustainable price erosion. 2) Incentivize manufacturers to add or expand domestic production by offsetting capital investment, using loans, grants, direct investment and cost-sharing, long-term tax

incentives for relocating, building, refurbishing or repurposing production lines, risk analyses to prioritise the medicines with the greatest offshore dependence, and subsidies or tax incentives for maintaining extra supplies, operating with excess capacity, preparing contingency plans with standby qualified equipment, and designing processes that can scale rapidly. 3) Develop a complete supply chain model from precursor materials to finished products using both onshore and ally-shored resources, including strengthened partnerships with nearshore allies in Latin and Central America, cooperation with Canada, Europe, Mexico and Japan, and a Pharmaceutical Supply Chain Trade Agreement with allies who commit not to impose trade restrictions during public health emergencies. 4) Streamline implementation and communication around regulatory procedures, including earlier regulator input on process development and validation, stronger FDA-manufacturer communication about support for rapidly qualifying new suppliers, and developing shortage manufacturing establishments by pre-qualifying existing facilities so new lines can ramp quickly.

Advanced manufacturing capabilities and innovative research and development. 1) Identify, promote and invest in innovative and sustainable manufacturing paradigms, in particular augmenting a centralized model with decentralized, regionally

based, on-demand manufacturing such as distributed or modular production, which supports smaller-scale and targeted production, reduces reliance on complex logistics and stockpiling, allows point-of-care production and reduces geographic disparities in distribution; plus sustained investment in platform technologies and in distributed networks of flexible platform capability so that no single large facility is a point of failure. 2) Develop strategies for targeted workforce investment, including training centres offering short practical courses in engineering, data analytics and process development, and partnerships with educational institutions covering equipment donation, hands-on learning, programme advisory roles and internships. 3) Support and invest in innovative drug research and development, including a comprehensive database of essential medicine substitutions and similar formulations that rely less on hard-to-get APIs, investment in novel platforms such as cell-free systems and miniaturization, and regional development hubs or incubators on the research park model. 4) Promote smart manufacturing through increased automation and other digital solutions. The published report prints the same three implementation actions under this strategy as under the previous one, so its specific actions are not distinguishable from strategy 3's. 5) Encourage widespread adoption of new manufacturing technologies through standardization and greater regulatory certainty while ensuring quality, including

more efficient review of manufacturing process changes, a standardized approach for applications using continuous manufacturing, and standards preventing cross-contamination.

Purchasing, stockpiling and distribution approaches. 1) Leverage federal buying power to reform procurement protocols, with more predictable purchasing including domestic-production targets in purchase agreement programmes, more than one contract per essential medicine, a preference for domestically produced essential medicines procured through the VA, Department of Defense or CMS on the model of the Berry Amendment, and a non-profit consortium negotiating reimbursement rates directly with CMS, Medicare and Tricare, with higher rates for products made domestically in facilities with demonstrated quality maturity. 2) Revise purchasing models to emphasize product quality and supply chain resilience rather than lowest cost, with a rating system encouraging investment in quality management maturity, non-proprietary information disseminated through the FDA or an independent nonprofit, and CMS reimbursement or federal contracts made contingent on quality and resilience practices including metrics on manufacturing line readiness, production capacity and risk management. 3) Regionalize the Strategic National Stockpile, with local or regional stockpiles managed by neutral parties reporting to a consortium, partnerships

between large academic medical centres and smaller community hospitals for medicine sharing, and limits on how many centres stockpile in one area. 4) Develop a distributed stockpile of APIs and precursor materials, with increased safety stock of high-risk materials and a preference for domestically produced APIs. 5) Incentivize manufacturers and wholesale distributors to hold buffer stock, rotate stock and manage distribution, including a vendor-managed inventory model run by manufacturers and distributors rather than hospitals, and connecting Defense Production Act rated orders to total vendor capacity. 6) Develop mandates or explore statutory authority for maintaining stock reserves, including requiring producers to keep a reserve rather than shipping everything off the line, investigating FDA authority to require stability data enabling expiration date extensions, using the CDC's CHEMPACK model for guidelines on using supplies beyond listed expiry with validation in extreme shortage, and buy-back programmes for stockpiled medicines nearing expiry. 7) Proactively develop comprehensive distribution logistics plans, including mandatory crisis reporting of stock levels, a federal allocation mechanism to channel medicines to hot zones and underserved areas, clear prioritisation criteria for fair share allocation of shortage medicines, and logistics contingency plans for disruptions with nearshore and offshore producers.

Next steps

By analysing manufacturing and supply data for the prioritised medicines, the government can identify those whose vulnerabilities need the most targeted action. The report presents its strategies as moving toward a more coordinated and prepared supply chain, in which the reliable availability of medicines when and where they are needed strengthens the country's ability to respond to crisis- or market-based demand fluctuations and to ensure economic and national security. It names assessing feasibility and impact, and identifying specific action plans with roles and responsibilities, as the work that follows.

Words to know

- **Active pharmaceutical ingredient (API)** The component of a medicine that does the therapeutic work, as distinct from the finished dosage form. 87 percent of the sites making them are overseas.
- **Finished dosage form** The final product as administered, the vial, tablet or inhaler. 63 percent of those sites are overseas.
- **Just-in-time inventory** Producing small batches and holding little stock to lower costs. Efficient in calm conditions and brittle in a surge.
- **Quality management maturity** How developed a manufacturer's quality systems are. The report's recurring point is that purchasers cannot see it, so nobody is paid for having it.

- **Continuous manufacturing** Making product in a continuous flow with inline monitoring rather than in discrete batches. It would help with quality and surge, and the generics market cannot fund it.
- **Current Good Manufacturing Practice (CGMP)** The regulatory standard a facility must meet, and part of why a new plant takes 5 to 7 years.
- **Bullwhip effect** Small demand changes amplifying into large swings as they travel up a supply chain, worsened here by hoarding and poor information.
- **Ally-shoring** Locating production with trusted partner countries rather than purely domestically, which the report treats as a realistic middle path.
- **Strategic National Stockpile** The US federal reserve of medicines and supplies, which the report recommends regionalising.
- **Virtual stockpile** Not a warehouse but a mandated, anonymised, aggregated picture of what inventory exists across industry.

Numbers in this reading

- 86 critical medicines in the downselected list, drawn from the FDA's essential medicines list, for acute patient care.
- Method: about 40 interviews, more than 80 surveys, two workshops with a clinical steering committee of over 40 people and two with an industry steering committee of over 30, a public

comment period from October 2020, and review of more than 100 sources.

- 87 percent of API manufacturing sites and 63 percent of finished dosage form sites located overseas as of March 2021.
- An estimated 60 percent of US manufacturing jobs going unfilled between 2015 and 2025 because of talent shortages.
- At least 63 percent of US drug shortages between 2013 and 2017 traceable to quality issues.
- 5 to 7 years to set up a new plant; 3 to 5 years to add a single line in an existing plant.
- Injectable dexamethasone: a 610 percent increase in hospital orders in June 2020 and a 137 percent price increase in October 2020, with multi-month backorders.
- Albuterol: nearly 25 million people with asthma in the US; a 400 percent increase in inhaler demand early in the pandemic; 46 percent of respondents to an AAFA survey had trouble accessing asthma or allergy medication; first generic albuterol inhaler approved April 2020.
- Fentanyl: ventilated patients treated for up to two weeks; large 50-millilitre vials needed; as many as 25 thumb-size glass bottles cracked to fill one 50 millilitre bag for one patient, who might need two or more bags a day; fentanyl is 50 to 100 times more potent than morphine; the DEA had proposed cutting authorized production by 31 percent and in

early April 2020 ordered a 15 percent increase instead.

- 2017: Hurricane Maria damaged Puerto Rican facilities producing injectable saline for the United States, and supplies remained critically impacted for several years.
- Every two years: the recommended review cycle for the essential medicines list, aligned with the WHO cycle.

What the author is unsure about

The report is explicit that its challenges and solutions are based on industry input obtained through interviews, surveys and workshops, which means it reports what stakeholders said rather than independently measured findings, and the regulatory challenges in particular are framed as aspects of US regulation that are perceived as limiting flexibility. It says the solutions will need to be combined into an integrated strategy and that next steps may include assessing their feasibility and impact and identifying who does what, so no strategy here has been costed or tested. It notes that producing all precursor materials domestically may be unrealistic because some require components only available where the natural deposits are. It excludes several categories from the medicines exercise, including vaccines, blood products and threat countermeasures, so the list is about acute care rather than pandemic

countermeasures as such. It acknowledges other challenges, such as pharmaceutical wastewater, that it only treats inside its six themes. One oddity worth knowing: the published report prints the same three implementation actions under the smart manufacturing strategy as under the preceding research and development strategy, so that strategy's own actions cannot be read from the document.

For the discussion

- The sharpest structural point is that the medicines most needed to keep people alive are the least profitable to make, so a market optimised for lowest unit cost will keep producing exactly the fragility this report describes, and no amount of visibility fixes that on its own.
- The dexamethasone case study is the other half of her deep dive: RECOVERY found the drug in June 2020, and hospital orders rose 610 percent while the price rose 137 percent and manufacturers went to backorder, which is a reminder that proving a treatment works is only the first bottleneck.

5.6-2 · CORE

COVAX explained

Seth Berkley · 2020

Read the original paper → · the course assigns: whole piece, read for manufacturing and delivery (the same source is

assigned at 5.5 #4, read there for financing and procurement)

> **COVAX was designed to buy manufacturing capacity before anyone knew which vaccine worked, and then to release the doses by a fixed rule rather than by who shouted loudest.**

Average summary

Read for page 5.6, this is a document about factories and queues rather than about money. Berkley's starting point is a manufacturing fact: normally manufacturers will not risk the significant investment needed to build or scale up vaccine production until a vaccine is approved. In a pandemic costing the global economy **US$375 billion every month**, that caution guarantees delay and shortage in the weeks after licensing, which are the weeks that matter most.

So the COVAX Facility does two things at once. It monitors the vaccine landscape continuously and picks candidates on **scientific merit and scalability**, which makes manufacturability a selection criterion rather than an afterthought. And it works with manufacturers to provide investments and incentives so they are ready to produce the doses the world needs the moment a vaccine is approved. The cost of that readiness is visible in the price every participant pays: the negotiated price per dose, plus a speed premium invested in accelerating and scaling up manufacturing, plus a very small operating fee.

The target was 2 billion doses available by the end of 2021, enough to protect high-risk and vulnerable people and frontline healthcare workers. The portfolio was built on the assumption that most candidates fail, with more than 170 in development, roughly a 7 percent success rate at the preclinical stage and about 20 percent for those in clinical trials; COVAX held nine candidates in development and nine more under evaluation. On the research side, CEPI led the work with nine candidates supported, eight already in clinical trials, against US$1.4 billion committed by governments and a further US$1 billion still needed.

Delivery is governed by rules rather than by negotiation. Once a portfolio vaccine is approved, doses go to all participating countries at the same rate, proportional to population. About 5 percent of available doses is held back as a buffer stockpile for acute outbreaks and for humanitarian organisations, for example to vaccinate refugees who would otherwise have no access. And no country receives enough to vaccinate more than 20 percent of its population until every country in the financing group has been offered the same, with the only exception being countries that asked for less.

Berkley also notes a delivery consequence of the purchasing choices: a country that opts out of a vaccine allocated to it and waits for another may be

delayed in receiving its full committed volume, so flexibility costs time as well as money.

Why it's assigned

Page 5.6 is the manufacturing and delivery option, and this is the same source as 5.5 #4 read from the other end. The financing page asks who pays; this page asks how the doses get made at scale and in what order they arrive. For the Combat layer it is the only reading that gives a concrete worldwide allocation rule, which is where the unit's abstract claims about equitable deployment either hold or collapse. Read beside the ARMI supply chain report, it also shows the same idea at two scales: pay in advance for capacity you may not need, because capacity cannot be summoned once demand arrives.

In depth

The manufacturing problem COVAX was built around

COVAX is the vaccines pillar of the **Access to COVID-19 Tools (ACT) Accelerator**, launched in April 2020, coordinated by Gavi, the Coalition for Epidemic Preparedness Innovations and the WHO. Berkley describes it as a platform supporting the research, development and manufacturing of a wide range of COVID-19 vaccine candidates and negotiating their pricing, with all participating countries having equal access once vaccines exist.

The reason the manufacturing half of that sentence is there is stated plainly. Normally, manufacturers are reluctant to risk the significant investments needed to build or scale up vaccine manufacturing facilities until they have received approval for a vaccine. In the context of a pandemic costing the global economy US$375 billion every month, that reluctance would inevitably lead to significant delay and initial vaccine shortages once vaccines are licensed. To avoid that, the Facility works with manufacturers to provide investments and incentives to ensure they are ready to produce the doses needed as soon as a vaccine is approved.

This is the same at-risk logic the economics readings on page 5.5 argue for, expressed as an operating practice rather than a model. The Facility is not only buying doses; it is buying the readiness to make them.

Scalability as a selection criterion

What makes the portfolio a manufacturing instrument rather than only a financial one is how candidates are chosen. The Facility continually monitors the COVID-19 vaccine landscape to identify the most suitable candidates based on **scientific merit and scalability**, and participants get not just the largest and most diverse portfolio but an actively managed one.

The scale of the filtering problem explains why a portfolio is necessary at all. There were more than **170 candidate vaccines** in development, and the vast majority of those efforts were likely to fail. Based on previous vaccine development, those at the preclinical trial stage have roughly a **7 percent** chance of succeeding, while those that make it to clinical trials have about **20 percent**. COVAX's answer was the world's largest and most diverse portfolio, with **nine candidate vaccines already in development and a further nine under evaluation**.

On the upstream research side, CEPI leads the COVAX vaccines research and development work, with nine vaccine candidates supported, eight of which were already in clinical trials. Governments had committed **US$1.4 billion** toward that effort, and an additional **US$1 billion** was still needed to continue moving the portfolio forward.

Paying for speed

The price structure makes the cost of readiness explicit. As a pass-through facility, participants pay the amount negotiated by the Facility for the doses, plus a **speed premium invested in accelerating and scaling up manufacturing**, plus a very small fee for the facility's operation. Some manufacturers will provide vaccines at flat prices while others will tier prices based on income levels.

The Facility's other manufacturing lever is bargaining power. It uses the collective purchasing power that comes from having so many countries participate to negotiate highly competitive prices from manufacturers, which are then passed on to participants.

Berkley's framing of what countries buy by joining is worth noting, because it is about production rather than allocation: by pooling resources through the Facility, participating countries and economies are essentially helping to increase the world's chances of bringing about COVID-19 vaccines as quickly as possible, and in the quantities that we need. Having many self-financing economies sign up means COVAX can begin signing formal agreements with manufacturers and developers, which both increases the chance of successful development and ensures the necessary productive capacity is in place to manufacture the needed volumes the moment a vaccine is ready.

How many doses, and for whom

The initial aim was to have **2 billion doses** available by the end of 2021, which should be enough to protect high-risk and vulnerable people as well as frontline healthcare workers. That is a deliberate statement about what the first production run is for: not population coverage, but the people whose protection does most.

Self-financing countries and economies can request vaccine doses sufficient to vaccinate between **10 and 50 percent** of their populations, paying in proportion to what they request. Funded countries, subject to funding availability, receive enough doses to vaccinate up to **20 percent** of their population in the longer term. Because demand is initially likely to exceed supply once vaccines become available, allocation is spread across countries based on the number of doses available, increasing as availability increases.

The allocation rule

This is the delivery mechanism itself, and it is short enough to remember.

Once any of the COVAX portfolio vaccines has successfully undergone clinical trials, proved safe and effective and received regulatory approval, available doses are allocated to all participating countries **at the same rate, proportional to their total population size**.

A small buffer of about **5 percent** of the total number of available doses is kept aside to build a stockpile, to help with acute outbreaks and to support humanitarian organisations, for example to vaccinate refugees who may not otherwise have access. That is the only part of the allocation not driven by

population, and it is reserved for the people a population-based rule would otherwise miss entirely.

Then the ceiling. Even though self-financing participants can request enough doses to vaccinate between 10 and 50 percent of their population, **no country receives enough doses to vaccinate more than 20 percent of its population until all countries in the financing group have been offered this amount**. The only exception is countries that have opted to receive fewer than 20 percent. In delivery terms, that converts a race into two rounds: everyone to 20 percent first, then the rest.

Where delivery and purchasing terms interact

The two purchase arrangements have a delivery consequence as well as a financial one. Self-financing countries choosing a Committed Purchase guarantee to procure an agreed volume, in exchange for a lower upfront payment of US$1.60 per dose, or 15 percent of the total cost per dose, and can opt out only if a vaccine's price turns out to be twice or more what was expected. Those choosing an Optional Purchase can decline any particular vaccine without losing their share of other candidates, paying a higher down payment of US$3.10 per dose plus a risk-sharing guarantee of US$0.40 per dose, which exists to protect the Facility against liabilities if a participant walks away after the Facility has already contracted with a manufacturer.

The sentence that matters for this page is the one about timing: by opting out of vaccines allocated to them and waiting for another to become available, these countries may inevitably experience a delay in receiving their full committed volume. Optionality is paid for twice, once in cash and once in queue position.

Alongside the Facility sits the entirely separate Gavi COVAX Advance Market Commitment, which supports access for the **92 middle- and lower-income countries** that cannot fully afford to pay, funded mainly through Official Development Assistance plus private sector and philanthropic contributions, and not cross-subsidised by the Facility's participants. At the time of writing it had raised about US$700 million of an initial seed capital target of US$2 billion needed by the end of 2020. Its purpose, in delivery terms, is that those countries get access at the same time as higher-income self-financing countries, rather than afterwards.

What Berkley asks for

He calls building a comprehensive global solution this quickly a remarkable accomplishment, and says implementation now depends on countries buying into the Facility so it can make urgent investments now. At the time of writing, **78 higher-income countries and economies** had confirmed interest, with legally binding agreements due by 18 September

and upfront payments by 9 October. The two outstanding needs are the research and development gap of US$1 billion on top of the US$1.4 billion governments had committed, and the AMC's US$2 billion target by the end of 2020, without which ability to pay becomes a barrier to access and most of the world is left unprotected.

Words to know

- **Scalability** Whether a candidate can actually be manufactured in the volumes needed. COVAX uses it alongside scientific merit when choosing what to back, which makes it a selection criterion rather than a later problem.
- **Speed premium** The extra amount built into every participant's price, invested in accelerating and scaling up manufacturing. It is the line item where buying readiness becomes visible.
- **Pass-through facility** A buyer that charges participants what it paid, plus defined extras, rather than marking up. It is why the two purchase arrangements cost the same in the end.
- **Buffer stockpile** The roughly 5 percent of doses held back from the population-based allocation for acute outbreaks and humanitarian use, including refugees.
- **Financing group** The set of participating countries across which the 20 percent ceiling applies before anyone goes further.

- **Advance Market Commitment** The separate fund covering 92 middle- and lower-income countries, designed so they receive doses at the same time as self-financing countries rather than later.

Numbers in this reading

- US$375 billion every month: the cost of the pandemic to the global economy, the figure Berkley uses to justify paying manufacturers to scale up before approval.
- 2 billion doses available by the end of 2021: the initial target, aimed at high-risk and vulnerable people and frontline healthcare workers.
- More than 170 candidate vaccines in development; about a 7 percent chance of success from the preclinical stage and about 20 percent from clinical trials.
- Nine candidates in the portfolio already in development, nine more under evaluation.
- CEPI's R&D work: nine candidates supported, eight in clinical trials, US$1.4 billion committed by governments, US$1 billion still needed.
- 10 to 50 percent: the share of their population self-financing countries may request doses for. Up to 20 percent in the longer term for funded countries.
- About 5 percent of available doses held back as a buffer stockpile.

- 20 percent: the ceiling on any country's coverage until all countries in the financing group have been offered the same.
- Committed Purchase: US$1.60 per dose upfront, or 15 percent of total cost per dose. Optional Purchase: US$3.10 per dose plus US$0.40 per dose risk-sharing guarantee, and a possible delay in receiving the full committed volume.
- 92 middle- and lower-income countries under the Gavi COVAX AMC, which had raised about US$700 million of a US$2 billion target.
- 78 higher-income countries and economies had confirmed interest, with agreements due 18 September and payments due 9 October.

What the author is unsure about

Berkley heads Gavi, one of the three organisations running COVAX, and this piece is written to persuade countries to join before a deadline, so read the confidence in it accordingly. He is straightforward about scientific uncertainty, giving failure base rates and stating that there are no guarantees any candidate will ultimately succeed, which is the stated reason for a portfolio. Dose availability for funded countries is explicitly subject to funding availability. On manufacturing he asserts that investments and incentives will make producers ready the moment a vaccine is approved, without giving numbers for how much capacity was being paid for or by when, and

without discussing the inputs, fill-finish capacity or cold chain that later readings treat as the binding constraints. He also does not address what happens to the allocation rule when participating countries hold large bilateral contracts outside the Facility. All figures are as of 3 September 2020.

For the discussion

- Picking candidates on scientific merit and scalability is the most transferable idea in the piece, because it treats manufacturability as part of what makes a vaccine good rather than as a problem for later.
- The allocation rule is unusually concrete for a global agreement, and worth testing out loud: proportional to population, about 5 percent held back for outbreaks and refugees, and nobody past 20 percent until everyone has been offered 20 percent.
- The delay attached to the Optional Purchase is the piece's honest admission that flexibility and speed trade against each other, which is the same trade the manufacturing readings on this page describe when they ask factories to stay warm for a product nobody is buying yet.

5.6-3 · CORE

Testing fractional doses of COVID-19 vaccines

Witold Więcek, Amrita Ahuja, Esha Chaudhuri et al. · 2022

Read the original paper → · the course assigns: whole paper, excluding the SI Appendix

> **Half or quarter doses of some COVID-19 vaccines might protect almost as well as full ones and would stretch supply by hundreds of millions of doses a month, and nobody tested this because no manufacturer had any reason to.**

Average summary

Więcek and colleagues make an argument about dose size that is really an argument about speed. Where vaccine supply is the binding constraint, using less of each dose means vaccinating more people sooner, and vaccinating more people sooner saves lives even if each person is protected a little less well.

The clinical case rests on two links. First, neutralizing antibody titers turn out to be remarkably predictive of efficacy against symptomatic infection across vaccines. Second, phase 1 and 2 dose-ranging trials measured those titers at several doses. Putting them together: for Moderna's mRNA-1273, 50 and 25 microgram doses produced antibody titers not significantly different from the standard 100 micrograms, with model-predicted efficacy for the

fractional doses and measured efficacy for the standard dose all in the **90 to 95 percent** range. For Pfizer's BNT162b2, 10 and 20 microgram doses have model-predicted efficacy of 70 to 85 percent, against roughly 95 percent for the standard 30 micrograms.

The public health case comes from an epidemiological model. If half doses are as efficacious as full doses, switching would reduce deaths by 22 to 47 percent for a baseline 95 percent efficacious vaccine and 20 to 35 percent for a 70 percent one. Even if a half dose produced a fivefold reduction in antibody titers relative to a 95 percent efficacious full dose, the lower dose would still reduce total mortality; for a 70 percent effective vaccine the threshold is a 2.5-fold reduction.

The supply case: implementing fractional dosing globally for the most promising vaccines could increase supply by **450 million to 1.55 billion doses per month**, on 2021 production targets. Fractional dosing is not new, either. Several countries used one-fifth doses of yellow fever vaccine between 2016 and 2018 on WHO advice.

The authors are honest about risks: shorter duration of immunity, possible selection pressure for variants, equity objections and vaccine hesitancy. They answer each, and note the whole policy is reversible, since a stretched dose can be topped up later.

Their central institutional claim is that this research will not happen without public money. A course of vaccine was worth $500 to $1,000 socially against the $6 to $40 manufacturers received, firms would lose revenue if countries drew more doses from each vial, and they face reputational risk if something goes wrong. None of the fractional dosing studies under way were sponsored by vaccine makers.

Why it's assigned

Page 5.6 is the manufacturing and delivery option, and this is its most counterintuitive reading: a way to expand effective supply without building a single factory. The Combat layer keeps asking how to get protection to people faster, and this paper answers by changing the denominator rather than the numerator. It also closes a loop with her deep dive on page 5.4, because the reason nobody knew whether fractional doses work is a clinical trial incentive problem, the same one Więcek describes in his Asterisk piece on the same page of the course as this paper's lead author.

In depth

The value of speed, restated

The paper opens with the scale that justifies everything after it. Early in the pandemic the IMF estimated COVID-19 would cost the world $12 trillion in short-run GDP losses alone over a two-year period,

and subsequent estimates including health costs and long-run impacts are much larger. On those estimates, accelerating mass vaccination by even a month would be worth at least **$500 billion**.

That is why investing in multiple candidates and installing manufacturing capacity in parallel with research, rather than in sequence, had high expected social value despite the risk that individual investments would fail or prove redundant. Their rough estimate is that Operation Warp Speed would have paid for itself if it advanced vaccination in the United States by less than a day, which the footnote supports with the figures: between March and September 2020 the pandemic cost the US an estimated $16 trillion in long-term losses, approximately $75 billion per day, and OWS spent $18 billion. Because the social value of accelerating vaccine availability so greatly exceeds the commercial value to manufacturers, capacity investments on that scale would likely not have occurred without public financing.

OWS generated enough vaccine for the United States, but many countries still faced supply constraints, and the authors argue for pursuing in parallel any option with even a modest chance of accelerating vaccination: improving delivery systems, expanding supply, and using existing supply more efficiently, as

insurance against new variants requiring boosters or shocks to the supply chain.

Using lower doses is one such option, and it has precedent. **Fractional dosing** has been employed successfully for multiple diseases, including in 2016 to 2018 when several countries used one-fifth doses of yellow fever vaccine to combat epidemics, on WHO advice.

Why fractional doses might work

Efficacy of fractional doses of COVID-19 vaccines had not been tested, with one exception the authors flag: for AstraZeneca's ChAdOx1, a low-dose then full-dose regimen appears to have worked well. But phase 1 and 2 trials of various vaccines measured immune response in the form of neutralizing antibody titers at different doses, with four doses tested for Moderna's mRNA-1273.

The bridge is the Khoury model, which found that **neutralizing antibody titers for standard doses were remarkably predictive of efficacy** against symptomatic infection in phase 2 and 3 trials across eight vaccines and in convalescent individuals. The authors combined that modelled relationship with antibody data from dose-ranging studies to derive predicted efficacy for fractional doses.

The figure

Figure 1 plots mean neutralization level on the horizontal axis against protection from symptomatic infection on the vertical axis, following the Khoury curve, with a shaded 95 percent confidence band. Lighter points show the mean normalized immune response and measured clinical efficacy of specific vaccines at standard doses, plus the response in convalescent individuals; antibody levels are normalized to convalescent levels using each vaccine's trial data. The authors then calculate the ratio of mean antibody response for fractional versus standard doses from dose-ranging trials and plot the fractional doses on the same curve as darker shapes, with diamonds for doses in the elderly and circles for non-elderly adults. What to notice is the shape of the curve: it is steep at low antibody levels and flattens at high ones, which is why a large proportional drop in antibodies costs very little efficacy for a highly effective vaccine and a great deal for a weaker one. The authors note the small sample sizes typical of early-stage trials and do not include measures of uncertainty on the fractional points.

The results: at 14 days after the second dose, 50 and 25 microgram doses of mRNA-1273 produced antibody titers not significantly different from the standard 100 microgram dose, and the model-predicted efficacy of the two fractional doses and the measured efficacy of the standard dose all fall within the 90 to 95 percent range for symptomatic disease.

For BNT162b2, 10 and 20 microgram doses have model-predicted efficacy of 70 to 85 percent, against roughly 95 percent for the standard 30 micrograms.

Variants, severe disease, and who is left to vaccinate

Many phase 2 and 3 trials did not measure efficacy against variants of concern, and protection from symptomatic disease may be lower against them. For BNT162b2 and ChAdOx1, real-world data suggest effectiveness of the standard dose fell to **88 percent and 67 percent** respectively against the Delta variant, decreases also predicted by drops in antibody titers.

But all vaccines with available data are highly effective against severe outcomes even with variants present. In the United Kingdom, BNT162b2 and ChAdOx1 were found to be **96 percent and 92 percent** effective against hospitalization; data from Chile suggest that even CoronaVac, the least effective vaccine in their figure, is 86 to 88 percent effective against hospitalization and death. Cell-mediated immunity rather than neutralizing antibodies may underlie that durable protection from severe disease. Although no clinical data exist, the authors say it is reasonable to expect the decrease in protection against severe disease and death from fractional doses to be much smaller than for symptomatic infection.

They add two more reasons fractional dosing may fit the remaining unvaccinated population. Large portions of the unvaccinated have already been exposed and acquired some immunity, and recent evidence suggests previously infected individuals may need only one dose to be highly protected against reinfection. And because of age prioritisation and demographic patterns in low- and middle-income countries, the global unvaccinated population is younger on average than the world population and perhaps younger than the populations vaccines were tested in, and immune responses in younger people are stronger. Regulators have already acted on that logic: the US approved 10 micrograms of Pfizer, a one-third dose, for 5 to 11 year olds, with 3 micrograms, a one-tenth dose, under test, and Moderna is testing 50 micrograms, a half dose, for children. Some clinical data also suggest fractional doses produce fewer side effects, so if efficacy is comparable they might be superior on an individual benefit-risk basis. In November 2021 the FDA amended the emergency use authorization for the Moderna vaccine to include a 50 microgram booster for everyone over 18.

The epidemiological model

The traditional development process optimises outcomes for the individual taking the vaccine, trading efficacy against side effects. The authors' point is that

when vaccines are short, switching to a lower dose and vaccinating more people can improve overall public health even if efficacy for each vaccinated individual falls significantly, because the alternative is leaving more people completely unprotected for longer.

They simulate vaccination across a range of epidemic scenarios in a modified susceptible-exposed-infected-recovered model with a single epidemic peak, accounting for the age-varying effect of vaccination on infections and deaths. The base case is a vaccine with 95 percent efficacy against infection, a vaccination rate of **0.25 percent of the population per day**, approximately the recent global median, older individuals vaccinated first, and vaccination rates constrained by supply so that they are inversely proportional to dose size. They also run a case with 70 percent efficacy against infection, comparable to reported effectiveness of ChAdOx1 against symptomatic Delta infection, and a third analysis varying efficacy against infection while holding efficacy against death fixed at 95 percent. The base case assumes 80 percent of the eligible population takes the vaccine, and lowering that to 60 percent attenuates the benefits without changing the qualitative result.

Table 1 reports deaths averted as a percentage, with ranges corresponding to different epidemic scenarios

from R = 0.99 to R = 2.

Switching from a 95 percent effective full dose, if the fractional dose loses no immune response at all: - half dose: 22 to 47 percent fewer deaths - one-third dose: 32 to 69 percent - one-quarter dose: 37 to 80 percent

As the antibody ratio falls, the gains shrink but stay positive for fractional doses. At a ratio of 0.8, which the model maps to 94 percent efficacy, a half dose still averts 21 to 45 percent. At 0.4, mapping to 87 percent efficacy, 18 to 34 percent. At 0.2, mapping to 76 percent efficacy, 13 to 22 percent. The first column, for a full dose at reduced response, is negative throughout, which is the comparison showing the cost of simply having a weaker vaccine without the speed gain.

Switching from a 70 percent effective full dose, the same pattern holds but the thresholds bite sooner. At no loss, a half dose averts 20 to 35 percent of deaths. At a ratio of 0.4, mapping to 49 percent efficacy, the half dose ranges from -1 to 15 percent, meaning it can go either way. At 0.2, mapping to 34 percent efficacy, half, third and quarter doses are all negative or near zero in at least part of the range.

The headline thresholds: even if a half dose leads to a **fivefold reduction** in antibody titers relative to a 95 percent efficacious full dose, the lower dose would still reduce total mortality; for a baseline 70 percent

effective vaccine, the threshold is a **2.5-fold reduction**. The footnote explains why the two differ so much, and it is the key to the whole figure: a 2.5-fold reduction in mean neutralization level drops a 95 percent effective vaccine only to 87 percent efficacy, but drops a 70 percent effective vaccine to 49 percent.

A separate finding from the third model run: at vaccination rates typical in many low- and middle-income countries, even a vaccine with high efficacy against infection does not prevent large outbreaks, simply because not enough people are vaccinated in time. In those settings accelerating vaccination is still beneficial, because it confers direct protection against hospitalization and death to more people.

The risks, and the answers

The authors list four and respond to each.

Duration of immunity. Their modelling does not consider the rate of immunity loss, which will likely be affected by changing dose size. Protection from severe outcomes may be longer lasting. Even assuming duration is proportional to dose, they argue it will likely be optimal to improve coverage in the short term with lower doses and then use future supply as boosters, and shortages should ease over time as production comes online.

Variant emergence. Fractional dosing might increase the probability of more harmful variants emerging through selection pressure during a prolonged period of partial immunity. They suggest this may be sufficiently offset by reductions in the total number of infected individuals, and that the overall risk of new variants emerging among vaccinated individuals may be low.

Equity. Programmes using lower doses could be criticized as inequitable. Their response has three parts: if there is little efficacy loss and fewer side effects, lower doses may actually be better for the individuals who would be vaccinated either way; vaccinating more people with a somewhat less efficacious vaccine is still more equitable than the status quo; and reduced doses of some vaccines are likely more efficacious than the standard of care in many low- and middle-income countries, since a half dose of mRNA-1273 may be more efficacious than a standard dose of ChAdOx1, and a half dose of ChAdOx1 could be more efficacious than a standard dose of CoronaVac. They add a point about queues that is worth keeping: doubling the speed of a year-long vaccination programme cuts the wait by a week for someone two weeks from the front of the queue, but by six months for the person at the end.

Hesitancy. Switching could contribute to vaccine hesitancy, but if lower doses are safer, reduce side

effects and are more effective than the low-efficacy vaccines currently used in short-supply countries, fractional dosing might even help, since surveys suggest side effects remain the most common concern among the undecided. Their precedent is the 2016 yellow fever outbreak in the Democratic Republic of the Congo and Angola, where the WHO concluded the rationale for switching to lower doses was well understood by the target population: questions were raised, no significant resistance or misinformation specific to fractional dosing was observed, and uptake reached **98 percent** of the target population.

Why nobody has tested it

As of the paper, no regulatory agency or immunization advisory group had recommended fractional dosing for COVID-19 vaccines, and the WHO Strategic Advisory Group of Experts interim statement of 10 August 2021 encouraged more research. Despite the global shortage, the high expected value of testing, and promising trial data available since autumn 2020, very few studies had been conducted. The exception is mRNA-1273, where three results suggest safety, strong immune response to primary and booster doses and durability at lower doses: intradermal administration of a 10 microgram dose induces antibody responses comparable to convalescent individuals; a 50 microgram booster induces antibody

levels similar to or higher than those after a standard primary series; and a primary series of two 25 microgram doses elicits antibody and cellular protection comparable to convalescent individuals, with cellular protection persisting after six months.

As of November 2021 the authors knew of an observational study of half-dose ChAdOx1 primary series in Brazil, two randomized immunogenicity trials comparing low with full doses for BNT162b2 and mRNA-1273, a randomized evaluation of several low-dose boosters in the United Kingdom, and a platform trial on fractional doses in previously primed individuals announced by CEPI in October 2021. Preliminary results from one show that 88 percent of subjects with no previous exposure who received half a dose developed neutralizing antibodies, with titers similar between half and full doses. The sentence that matters: **none of these studies are sponsored by vaccine makers.**

How it could be tested, and the incentive gap

Large efficacy trials were needed for the initial approvals, but given advances in establishing correlates of protection, immunogenicity data on fractional doses may be enough for some policy makers, particularly national immunization advisory groups, to recommend their use, especially for younger adults or lower-risk groups. Immunogenicity trials can be conducted in a matter of months, at low

cost, and even when infection risk is low, with sample sizes on the order of 100 subjects per study arm sufficient to measure immune responses precisely, which makes it easy to study many dosages and subpopulations. Randomized efficacy trials would give more useful data in the longer term but take more time and much larger samples.

Alternatively, where supply is limited and infection risk high, policy makers could roll out the dose that offers the best expected public health outcome on current data and use rollout data to assess effectiveness and safety, limiting the rollout to certain areas or lower-risk age groups and collecting observational data in prospective cohort or case-control studies. The United Kingdom did something similar when it extended the gap between doses to 12 weeks in December 2020 on limited data, and the key property is reversibility: if evidence suggests the change is not effective or supply increases rapidly, the approach can be adjusted or reversed, which for fractional dosing means increasing the second dose size or providing boosters.

Logistics also need validation studies, but the answers look easy: advisory groups must assess how many times a vial can safely be punctured and whether smaller volumes can be administered consistently and with which syringes. In the FDA recommendation for 50 microgram Moderna

boosters, and in all studies so far, lower doses were obtained from standard preparations by drawing less volume into syringes, which suggests other vaccines may not need changes to fill and finish processes.

The incentive analysis is the paper's institutional core. The authors previously estimated the social value of an additional course of vaccine at **$500 to $1,000**, dwarfing the **$6 to $40** price manufacturers receive in current contracts, a gap that leads to significant private underinvestment in manufacturing capacity and probably in research on other ways to accelerate vaccination. They give three reasons manufacturers may actively avoid testing fractional doses for primary series, and say their outreach to selected manufacturers seems to confirm it. First, if evidence convinced countries to use lower doses off-label by drawing more doses from each vial, and public relations concerns prevented raising the vial price, companies would lose revenue. Second, manufacturing vials with specific doses involves substantial sunk costs, and seeking regulatory approval for a new dose is costly and slow, so manufacturers may prefer to focus on products sold at much higher prices in high-income countries, such as variant-modified boosters or combined COVID-19 and flu vaccines, and may see testing lower doses of existing products as interfering with that strategy. Third, firms face reputational risk if something goes wrong with a lower dose.

They add that no single country internalizes the total global benefit of testing, because the information generated is a global public good: after the UK's successful experiment with a longer gap between doses, many other countries adopted the same schedule. That suggests a role for global institutions to invest in and coordinate studies.

Value late in the pandemic, and the Omicron coda

Even with supply increasing, the authors argue the research retains high expected value as an option against shocks that might delay supply: new variants or other pathogens requiring repurposed manufacturing capacity, manufacturing or supply chain disruptions or quality issues, such as those Novavax had experienced. Their conservative estimate is that supply constraints would continue to bind in many countries at least until March 2022, predicted from manufacturers' production estimates and simple demand assumptions, accounting for boosters and assuming only a fraction of excess supply in high-income countries is donated.

They then undercut their own projection with a real example. Nigeria received 30.2 million doses between March and November 2021, an eight-month interval, 28 million of which came through donations from COVAX or other countries. To meet the supply model's predictions the country would have to receive 160

million doses in four months, five times the amount received so far in half the time.

Even with enough doses, they argue fractional dosing has other benefits: comparable protection with fewer side effects, especially for younger or previously infected people; increased supply of the most effective vaccines, since lower doses of some are likely more effective than full doses of others; and the precedent itself, because COVID-19 policies will influence future pandemic response and setting a precedent for dose optimisation is likely to influence future decision makers.

The coda, added in late 2021, reports that the emergence of Omicron led many countries to accelerate and expand booster programmes because of lower effectiveness of available vaccines, with some administering fourth doses. That put substantial pressure on supply, especially of mRNA vaccines, so despite increased production the share going to low-income countries had not grown in recent months. And if vaccines have to be reformulated against the new variant, the repurposing of manufacturing will disproportionately affect countries where few people have had a primary series. Their closing point is that fractional dosing research is insurance against exactly this kind of shock.

Witold Więcek declares that he provides scientific consultancy for Certara and for 1Day Sooner, with no

material conflicts of interest regarding COVID-19 vaccine development. The work was supported by the Wellspring Philanthropic Fund and Open Philanthropy.

Words to know

- **Fractional dosing** Giving less than the standard amount per dose, so the same supply covers more people. Already used for yellow fever at one-fifth doses in 2016 to 2018.
- **Neutralizing antibody titer** A measure of how much virus-blocking antibody a person has. The paper's whole clinical argument rests on this predicting efficacy.
- **Immunogenicity trial** A study measuring immune response rather than disease outcomes. Small, quick and cheap, about 100 subjects per arm, which is why the authors favour it here.
- **Efficacy versus effectiveness** Efficacy is the effect measured in a trial or predicted by a model; effectiveness is what is observed in the real world. The authors keep the two words distinct throughout.
- **Variant of concern** A viral variant with properties that make it more dangerous, such as Delta or Omicron, against which standard-dose effectiveness fell.
- **Dose stretching** The family of strategies for covering more people with the same supply,

including fractional doses and delaying second doses.

- **Option value** The worth of having information ready in case a bad scenario happens, even if it never does. It is the authors' main argument for testing late in a pandemic.
- **Global public good** Something whose benefits cannot be confined to the country that pays for it, like the evidence from a dosing trial, which is why no single country funds enough of it.

Numbers in this reading

- $12 trillion: the IMF's early estimate of short-run global GDP losses over two years. Accelerating mass vaccination by even a month would be worth at least $500 billion.
- Operation Warp Speed spent $18 billion and would have paid for itself by advancing US vaccination by less than a day; the US was losing roughly $75 billion a day, about $16 trillion between March and September 2020.
- 450 million to 1.55 billion doses per month: the potential increase in supply from global fractional dosing of the most promising vaccines, on 2021 production targets.
- One-fifth doses of yellow fever vaccine used by several countries from 2016 to 2018 on WHO advice, with 98 percent uptake among the target population in the DRC and Angola in 2016.

- mRNA-1273: 50 and 25 microgram doses produced antibody titers not significantly different from the standard 100 micrograms, with all three in the 90 to 95 percent efficacy range.
- BNT162b2: 10 and 20 microgram doses have model-predicted efficacy of 70 to 85 percent, against roughly 95 percent for the standard 30 micrograms.
- Delta: standard-dose effectiveness fell to 88 percent for BNT162b2 and 67 percent for ChAdOx1.
- Against hospitalization: 96 percent for BNT162b2 and 92 percent for ChAdOx1 in the UK; 86 to 88 percent for CoronaVac against hospitalization and death in Chile.
- Paediatric and booster precedents: 10 micrograms of Pfizer (one-third) approved for 5 to 11 year olds with 3 micrograms (one-tenth) under test; Moderna testing 50 micrograms (one-half) for children; FDA amended Moderna's authorization in November 2021 to include a 50 microgram booster for all over 18.
- Model assumptions: 95 percent and 70 percent baseline efficacy against infection; vaccination rate of 0.25 percent of the population per day, approximately the recent global median; 80 percent uptake in the base case, tested down to 60 percent; epidemic scenarios from R = 0.99 to R = 2.

- Deaths averted switching from a 95 percent effective full dose with no loss of response: 22 to 47 percent for half doses, 32 to 69 percent for thirds, 37 to 80 percent for quarters. At an antibody ratio of 0.4 (87 percent efficacy): 18 to 34, 28 to 59 and 34 to 73 percent.
- Deaths averted switching from a 70 percent effective full dose with no loss: 20 to 35 percent for half doses, 30 to 52 for thirds, 35 to 64 for quarters. At a ratio of 0.2 (34 percent efficacy) the numbers turn negative across most scenarios.
- Thresholds: a fivefold reduction in antibody titers still leaves half doses of a 95 percent effective vaccine reducing mortality; for a 70 percent effective vaccine the threshold is 2.5-fold. A 2.5-fold reduction drops 95 percent efficacy to 87 percent but drops 70 percent efficacy to 49 percent.
- 88 percent: the share of previously unexposed subjects given half a dose who developed neutralizing antibodies, in preliminary results from one ongoing trial.
- About 100 subjects per study arm: enough for an immunogenicity trial to measure immune responses precisely.
- $500 to $1,000 social value per additional course, against the $6 to $40 price manufacturers receive.
- Nigeria received 30.2 million doses between March and November 2021, 28 million of them donated,

and would need 160 million doses in four months to match the supply model: five times the amount in half the time.

- March 2022: when the authors conservatively estimate supply constraints would still bind in many countries. Airfinity projected global supply sufficient to vaccinate 70 percent of the world by May 2022 under different assumptions.
- 10 August 2021: the date of the WHO Strategic Advisory Group of Experts interim statement on fractional dosing, which encouraged more research.

What the author is unsure about

This is a paper arguing for testing rather than for switching, and it says so. Efficacy of fractional doses had not been tested for COVID-19 vaccines except the ChAdOx1 low-then-full regimen. The efficacy numbers for fractional doses are model predictions derived from antibody data, not measurements, built on an exploratory approach with small sample sizes typical of early-stage trials, and the authors do not include measures of uncertainty on those points. There are no clinical data on protection against severe disease from fractional doses, only a reasonable expectation. The model does not consider the rate of immunity loss, which dose size will probably affect. The variant-selection risk is addressed with a may rather than a finding. The supply projections are described as

needing caution and are immediately contradicted by the Nigeria example. And the lead author declares scientific consultancy for Certara and for 1Day Sooner. Their overall position is that even a modest chance fractional doses work makes testing worthwhile, because the policy is reversible and the downside is bounded.

For the discussion

- The shape of the antibody-efficacy curve is the heart of this paper: a 2.5-fold drop in antibodies costs a 95 percent vaccine only eight points of efficacy and costs a 70 percent vaccine twenty-one, which is why fractional dosing is a strategy for the strongest vaccines and not a general one.
- The most striking fact is institutional rather than clinical: none of the fractional dosing studies under way were sponsored by vaccine makers, and the authors give three concrete reasons a manufacturer would rather not know the answer.
- The queue argument is the clearest equity point in the unit: doubling the speed of a year-long campaign saves someone near the front a week and someone at the back six months, which makes speed and fairness the same policy rather than competing ones.