

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Brad Power
January 10, 2024

“The goals of what we're trying to do are to (1) do some biological discovery..., (2) make better ways of integrating this into clinical care, and then (3) doing it in a way that everyone can access it.” – Eli Van Allen

“The patient is the model system that we want to study, instead of cell lines in a petri dish, not mice, but rather human experiences... Applying algorithms to the high dimensional data we can generate from our patients to learn as much as we can from them.” – Eli Van Allen

“For me, it can be mentally jarring flipping between my clinical and research brains. In my research brain, I'm looking at spatial transcriptomics and single cell data of like, millions, billions, or one day trillions of data points from all these high dimensional sources. My clinical brain is like, “Alright, I'm doing Lupron and docetaxel, or Lupron or Abi (abiraterone).” It's like, this or that, and what do I have to guide us on those things? It's a completely different scope. So the goal, where the data comes in, is you build the bridge, and you also build new bridges.” - Eli Van Allen

Meeting Summary

In an ideal world, patients, caregivers, and physicians (and AI working on their behalf) would be able to search a database of the experiences of patients similar to them and help them in their decision-making process about treatment. And researchers would be able to develop more complex biomarkers and processes to predict disease progression and drug response.

To achieve these insights we can harness “bioinformatics” – tons of cancer patient data and technological advances, especially artificial intelligence.

- **Why/Biological Discovery:** We can discover new drug targets, understand why a disease occurs, why it becomes resistant to the drugs we give it, then how we can intervene.
- **For Whom/Clinical Guidance:** We can personalize treatment plans for cancer patients: who should get what drug, where should that decision-making be done, and how algorithms can guide them.
- **How/Equitable Implementation:** We can do this so that everyone, near and far, can experience it, and not just a select few patients who happen to lock into a clinical trial or happen to be living near a big quaternary care cancer center. We can bridge the gap between data/AI/computational tools and clinical applications in the long term.

Medical oncologist and computational biologist Dr. Eli Van Allen is uniquely qualified to describe the landscape of patient data repositories and translating that data into clinical use. He is an institute member at the Broad Institute of MIT and Harvard, Scientific Champion for Count Me In, an associate professor of medicine at Harvard Medical School, and chief of the Division of

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Population Sciences at the Dana-Farber Cancer Institute. His areas of focus are computational cancer genomics, the application of new molecular technologies to advance precision cancer medicine, and studying resistance to cancer therapeutics using biologically guided artificial intelligence. Van Allen’s current research includes integrative studies of genitourinary (prostate, bladder, kidney, and testicular) cancers to determine the appropriate treatment for individual patients. He earned a B.S. from the Symbolic Systems Program at Stanford University and his M.D. from the David Geffen School of Medicine at UCLA. He completed his residency in internal medicine at the University of California, San Francisco, and served as a fellow in medical oncology at Dana-Farber.

Why/Biological Discovery: How are cancer patient data and AI guiding clinical innovations in cancer care and uncovering new drug targets?

- *Case Example: AI was used to predict which tumors are lethal.* To understand which genes are lethal and which are not in prostate cancer, 1000 prostate cancer patient whole genome sequences (20,000 genes) were fed into an AI model (a biologically-informed interpretable neural network) with genes mapped to molecular pathways mapped to processes. This enabled stratification, prediction, and interpretation – you can look at the answers from the model and understand why it predicted what it predicted.

For Whom/Clinical Guidance: How are cancer patient data and AI personalizing treatment plans for cancer patients?

- Computer vision algorithms can analyze medical images and predict which cancer patients are most likely to benefit from immunotherapy.
- AI can match patients to clinical trials based on genetics and generate longitudinal data on disease evolution.
- An algorithm was used to figure out which patients with prostate cancer or melanoma have inherited genetic events that are actionable.

How/Equitable Implementation: What changes are needed to bridge the gap between data/AI/computational tools and clinical applications in the long term?

- Gather training data that represents humanity, where any patient from anywhere can participate in research and contribute their data.
- Simplify data access rules, e.g., HIPAA regulations, to facilitate collaboration among researchers for more data sharing and overcome challenges in sharing clinical trial data.
- Develop better tools for interpreting data.
- Develop better tools for inclusion of patient-reported outcomes.
- Evaluate personalized medicine algorithms to avoid bias and ensure equity, understanding both "for whom" and "why" a drug will be effective.
- Ease access for patients to their genomic data.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

- Ease access for patients to experimental therapies. For example, implement the Promising Pathway Act, a new law that would change the clinical trial system to allow everybody to access experimental treatments through a conditional approval, within an observational clinical trial where experience is tracked to get enough proof to see whether it should graduate to full approval.
- Manage the exposure risk of a provider in one place guiding the clinical care of local patients, which could then be impacting patients around the world who are following those treatments and outcomes, creating possible problems of access, equity, and bias, and whether these models are generalized and be translatable.
- Find new drug targets that already have existing drugs or clinical trials to accelerate time to patient access.
- Inform patients of their rights under HIPAA to demand all of your raw data from every hospital you've ever been to or every commercial sequencing site – not just a little .pdf from a portal. (Then patients can share their data with whomever they want.)

The information and opinions expressed on this website or platform, or during discussions and presentations (both verbal and written) are not intended as health care recommendations or medical advice by Cancer Patient Lab, its principals, presenters, participants, or representatives for any medical treatment, product, or course of action. You should always consult a doctor about your specific situation before pursuing any health care program, treatment, product or other course of action that might affect your health.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Meeting Notes

SUMMARY KEYWORDS

patients, data, work, question, ellie, algorithms, tumor, cancer, point, called, ai, prostate cancer, information, clinical, drug, access, biology, models, snowballed, clinical trials

SPEAKERS

Eli Van Allen (45:31), Brad Power (3:29), Brian McCloskey (2:55), Frank Nothhaft (2:20), Rick Stanton (1:40), Al Musella (1:36), Jeff Krolick (1:30), David Plunkett (0:14), Mike Donohoo (via chat), Eric Hall (via chat)

OUTLINE

1. Using patient data for medical research and personalized insights. (0:00)
2. Patient cancer data and its use in clinical innovations. (2:15)
3. Personal background, cancer research, and data-driven approaches in oncology. (4:12)
4. Using AI in cancer research and treatment. (9:14)
5. Using AI for cancer diagnosis and treatment. (14:27)
6. Using AI to personalize cancer treatment. (19:35)
7. AI modeling in cancer research and liquid biopsy results. (24:29)
8. Analyzing cancer DNA in blood samples. (28:24)
9. Personalized cancer treatment and patient data. (34:08)
10. Using AI for personalized medicine. (39:04)
11. AI for drug discovery and regulatory barriers. (44:06)
12. Personalized cancer treatment options and the challenges of interpreting data. (52:49)

SUMMARY

- **Using patient data for medical research and personalized insights.** [0:00](#)
 - Eli Van Allen, a medical oncologist and computational biologist, discusses how Count Me In can help patients with personalized clinical insights.
 - He discusses patient cancer data and its use in guiding clinical innovations.
- **Personal background, cancer research, and data-driven approaches in oncology.** [4:12](#)
 - Eli Van Allen shares his origin story, from boredom in school to discovering computers and eventually starting a nonprofit for kids whose parents have cancer.
 - He is now an M.D. clinician at Dana Farber, seeing mostly prostate cancer patients, and previously worked in technology companies before pivoting to cancer medicine.
 - He shares a personal story of a patient with metastatic kidney cancer who was able to go into remission through a clinical trial, highlighting the potential of data-driven cancer care.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

- He emphasizes the importance of harnessing large amounts of cancer patient data to find new drug targets and understand disease resistance, with the goal of developing personalized treatment plans.
- **Using AI in cancer research and treatment. [9:14](#)**
 - Eli Van Allen discusses the potential of AI in oncology, particularly in prostate cancer, to predict which tumors are lethal and identify genetic lesions for drug development.
 - He highlights the limitations of current AI approaches, including the lack of understanding of why certain predictions are made, and the need for more research to integrate patient data and advance equitable cancer treatment.
 - He discusses using a biologically informed neural network to analyze cancer patient data and identify new drug targets.
 - He also highlights the importance of understanding "for whom" the drug will be effective, in addition to "why" it will be effective.
- **Using AI for cancer diagnosis and treatment. [14:27](#)**
 - Eli Van Allen discusses using computer vision algorithms to analyze medical images and predict which cancer patients are most likely to benefit from immunotherapy.
 - The study aims to develop ethical frameworks for using these algorithms in clinical care, while avoiding potential harm to patients.
 - He discusses the potential of using AI models like ChatGPT to analyze cancer cells and generate hypotheses for new discoveries.
 - He demonstrates how ChatGPT can be used to write a song in the style of Britney Spears about the top five men's singles tennis players, showing its ability to generate novel content.
- **Using AI to personalize cancer treatment. [19:35](#)**
 - Dana Farber is using AI to match patients to clinical trials based on their genetics, with algorithms already in deployment.
 - Eli Van Allen is involved in the Metastatic Prostate Cancer Project, which invites patients from anywhere to participate in research by clicking a "Count me in" button and completing a survey and consent form.
 - He discusses a research project aimed at generating longitudinal data on the course of disease, with a focus on understanding tumor evolution and developing new biology that could help everyone.
 - He thanks his team and acknowledges the unique opportunity to work with a talented group of individuals, including those on the Zoom call, to make a meaningful impact in the field of oncology.
- **AI modeling in cancer research and liquid biopsy results. [24:29](#)**
 - Eli Van Allen acknowledges limitations of toy model that snowballed into Nature paper.
 - He questions whether AI models can accurately analyze liquid biopsy data due to limitations in prior knowledge.
- **Analyzing cancer DNA in blood samples. [28:24](#)**
 - Rick Stanton is curious about a high VAF (variant allele fraction) in a tumor sample, suggesting a possible tumor suppressor mutation.
 - Eli Van Allen explains that high VAF can occur if the mutation is needed to turn off a gene, and provides an example of men with de novo metastatic prostate cancer having a high shed tumor burden in their blood.
 - Eli Van Allen discusses challenges in sharing data for AI research, citing regulatory issues and a lack of access to data from the Promise Program.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

- Eli Van Allen advocates for simplifying data access rules to make it easier for researchers to collaborate and share data.
- **Personalized cancer treatment and patient data.** [34:08](#)
 - Sequencing vendors struggle to provide personalized treatment recommendations due to limited data and lack of patient context.
 - Eli Van Allen discusses the challenges of integrating computational tools into clinical care, including the need for better data sharing and implementation.
 - He also highlights the importance of including patient-reported outcomes in cancer research, such as supportive care and other factors not captured in medical notes.
- **Using AI for personalized medicine.** [39:04](#)
 - Brad Power expresses interest in using AI for personalized medicine, specifically finding patients with similar profiles and understanding treatment outcomes.
 - Eli Van Allen explains that within their organization, there is a project focused on returning results to patients as part of a national cancer institute consortia, but acknowledges the challenges of implementing personalized medicine, including uncertainty and complexity for providers.
 - He discusses the challenges of developing patient similarity algorithms that can generalize across different populations and medical environments.
 - He expresses concerns about bias and equity in the development and implementation of these algorithms, and the need for careful consideration and evaluation to ensure they are effective and ethical.
- **AI for drug discovery and regulatory barriers.** [44:06](#)
 - Eli Van Allen highlights the potential for short-term patient gains from using a data asset like the one being developed, such as the work on TRL prediction from slides.
 - He notes that there are other examples of research that could be more quickly translated to the clinic, such as digital pathology and precision oncology.
 - Frank Nothaft discusses the challenges of sharing clinical trial data, including regulatory issues and concerns about patient privacy.
 - Eli Van Allen highlights the complexity of HIPAA regulations and the challenges of managing identifiable information in clinical trials.
 - He advocates for patients' right to access their genomic data and encourages them to demand their data through HIPAA right of access.
 - His call to action has faced resistance from some companies, which have intentionally violated HIPAA or refused to provide data, highlighting the challenges in scaling up this approach.
- **Personalized cancer treatment options and the challenges of interpreting data.** [52:49](#)
 - Al Musella discusses a project to provide experimental treatments to patients outside of clinical trials, including the Promising Pathway Act, which would give access to experimental therapies through a distributed clinical trial model.
 - Eli Van Allen shares a personal experience with a patient with metastatic bladder cancer who faced difficulties in accessing early access to pembrolizumab through Merck, highlighting the need for more efficient and equitable access to treatments.
 - Eli Van Allen shares a personal story of a patient who was cured of a metastatic solid tumor.
 - Brian McCloskey raises a concern about the lack of concordance in treatment options provided by different service providers, highlighting the need for more predictive and targeted approaches in cancer treatment.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

- Eli Van Allen expresses skepticism about the immediate practicality of some cutting-edge biotech tools, citing the need for more data to guide drug selection.
- Eli Van Allen and Brian McCloskey discuss the challenges of bridging the gap between research and clinical applications, with a focus on building new bridges in the long term.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

TRANSCRIPT

Brad Power

I'm the host of this session today along with Brian who's going to moderate. This is the Cancer Patient Lab. This is a community of advanced cancer patients who have the pleasure of talking today with Eli Van Allen. We've been in conversation with Eli for a long time because Count Me In is doing great work to allow patients to contribute their data to medical research.

Eli's also an expert in prostate cancer, and we started in prostate cancer. Many of our patients have been involved with Eli. Eli was very helpful with our good friend Bryce Olson when we had a virtual tumor board. And Bryce has been an inspiration for this whole community and for everyone here, and he's deeply missed. That's how we got to know Eli, which was probably three or four years ago now.

We've been wanting to have this conversation about how Count Me In can help patients. In particular, we're interested in any feedback that can be done clinically, because in the past, Eli has explained to us that it's Research Use Only, and it has to be kept for that research purpose. But of course, as patients, we'd like to find some insights that can be personalized and useful to us in clinical guidance. That's the territory that Eli will be discussing today, I hope.

Eli Van Allen 2:15

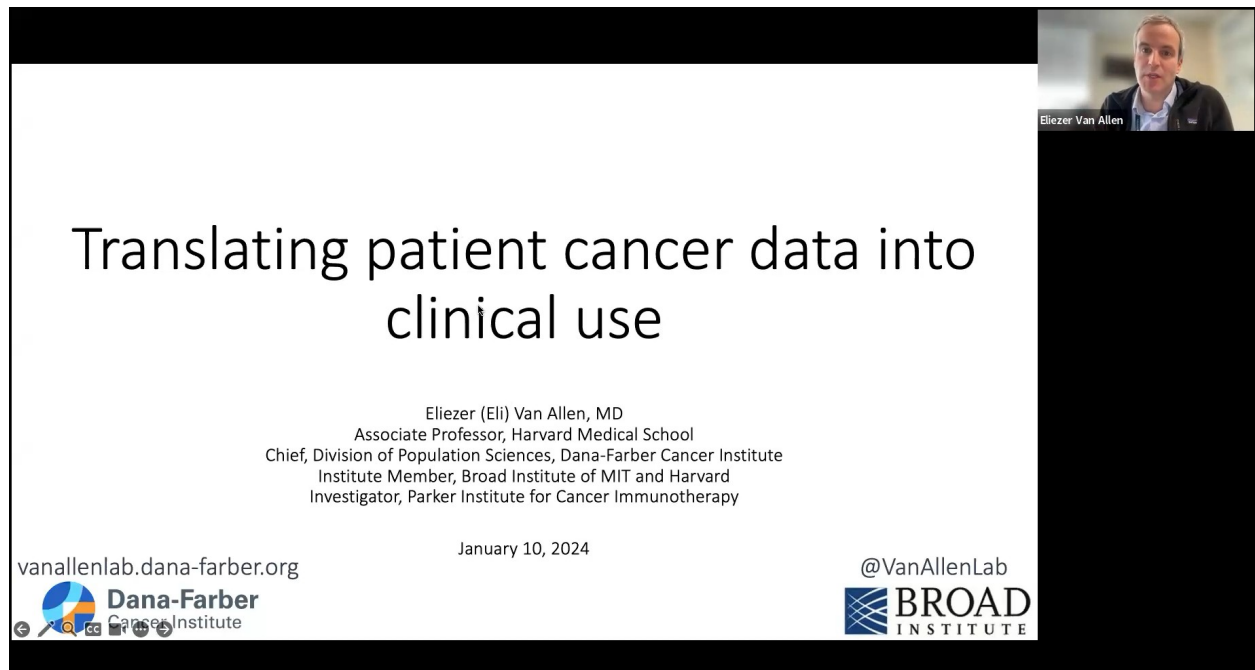
Thank you, Brad, for the very kind intro.

I took a slightly broader interpretation of the goals of this session based off of the way the invite was structured. I will certainly be touching on Count Me In as part of a larger conversation around why your data matters, and how we're trying to make the most of it.

Four years into this, we should all have it figured out by now.

It's a privilege to virtually be here and to present to this illustrious group. I'm privileged to have this invitation to do so. I'm a medical oncologist at Dana Farber, with a focus in genitourinary cancers like prostate cancer. I'm also a computational biologist and run a research lab at Dana Farber and at the Broad Institute.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]




Translating patient cancer data into clinical use


Eliezer (Eli) Van Allen, MD
Associate Professor, Harvard Medical School
Chief, Division of Population Sciences, Dana-Farber Cancer Institute
Institute Member, Broad Institute of MIT and Harvard
Investigator, Parker Institute for Cancer Immunotherapy

January 10, 2024

vanallenlab.dana-farber.org

 Dana-Farber
Cancer Institute


@VanAllenLab

 BROAD
INSTITUTE

Here's our lab website (vanallenlab.dana-farber.org) in case you're curious to poke around about what we do. My increasingly dwindling social media footprint (@VanAllenLab) is on the right. It's semi-retired at this point, but we'll see how that goes. Here are all my various titles.

I'll try to address what's happening in patient cancer data, and how it is being used to guide clinical innovations, which is a massive topic that I could never possibly do justice with 20 hours, much less 20 minutes. I'll try to touch on a few points and then leave the rest of our conversation.


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



Eli Van Allen

Disclosures

- Advisory/Consulting: Tango Therapeutics, Genome Medical, Genomic Life, Enara Bio, Janssen, Manifold Bio, Monte Rosa, Novartis Institute for Biomedical Research, Riva Therapeutics, Serinus Bio
- Research support: Novartis, BMS, Sanofi
- Equity: Tango Therapeutics, Genome Medical, Genomic Life, Syapse, Enara Bio, Manifold Bio, Microsoft, Monte Rosa, Riva Therapeutics, Serinus Bio
- Patents: Institutional patents filed on chromatin mutations and immunotherapy response, and methods for clinical interpretation; intermittent legal consulting on patents for Foaley & Hoag
- Editorial Boards: *Science Advances*



Here are my disclosures.



Eli Van Allen

Introductions, cont.



TEMPLE EMANUEL COMMUNITY DAY SCHOOL
BIVKA GIVONI
MILKA SANSON
HAYA LANIR
4A
1990 - 91

I thought I'd also give a little bit more background as an introduction of who I am, and how I got here. I trace back my origin story to how I ended up in this field that I'm in to fourth grade. I'm from Los Angeles. I went to Temple Emanuel Community Day School. Here's a picture of me in fourth grade with a ridiculous set of glasses. This was taken from our yearbook that my mom dug up for me for purposes like this. I remember being extremely bored growing up.

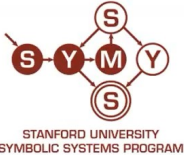


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Introductions, cont.




Most of my memories were being bored in school, until one day when they moved this Apple IIe into our classroom. For me, this was an intellectual awakening of sorts, and I got to basically start playing with computers, and setting up this course that I'm on today.

Introductions, cont.



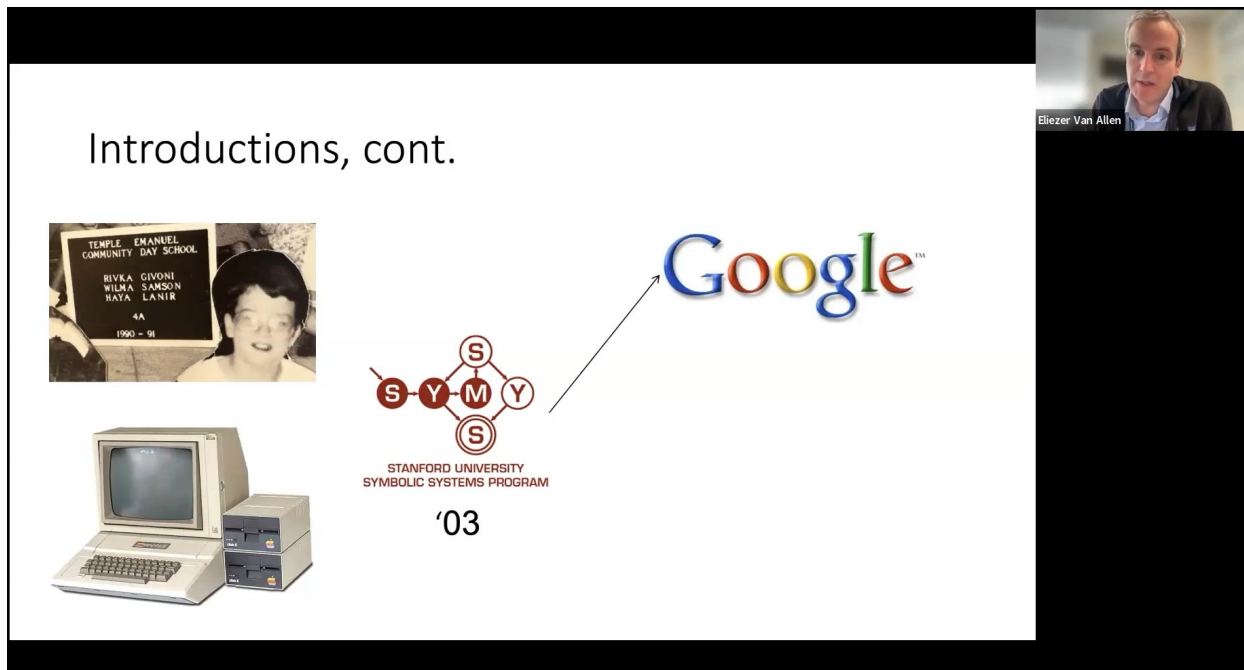
'03



I went to college at Stanford. I studied something called symbolic systems, which is a mix of computer science with philosophy and linguistics. It's basically a computer science degree.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Introductions, cont.



TEMPLE EMANUEL COMMUNITY DAY SCHOOL
RIVKA GIVONI
WILMA SAMSON
RAYA LANIR
4A
1990 - 91

STANFORD UNIVERSITY
SYMBOLIC SYSTEMS PROGRAM
'03

Google™

Eliezer Van Allen

One very tried and true career path for people who pursue this line of academic inquiry is to go work in technology companies. As you may have heard, in Silicon Valley, especially around that time in the world, there were a lot of them exploding. It was a very exciting time. That's what I thought I was going to do.

Introductions, cont.



TEMPLE EMANUEL COMMUNITY DAY SCHOOL
RIVKA GIVONI
WILMA SAMSON
RAYA LANIR
4A
1990 - 91

STANFORD UNIVERSITY
SYMBOLIC SYSTEMS PROGRAM
'03

Google™

Camp Kesem

Eliezer Van Allen

But as they say, a funny thing happened along the way. Some friends approached me interested in starting a nonprofit for supporting kids whose parents have cancer. That turned into

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

something called Camp Kesem. It's for kids whose parents have cancer. We started the first camp at Stanford in 2001. I'm very proud to say that there are now about 150 of them around the country. Certainly, I'm happy to point folks offline, if anyone is interested in learning more about this program, which for me, was a life changing experience and has snowballed into why I pivoted into going into cancer medicine, and why I'm at it today.

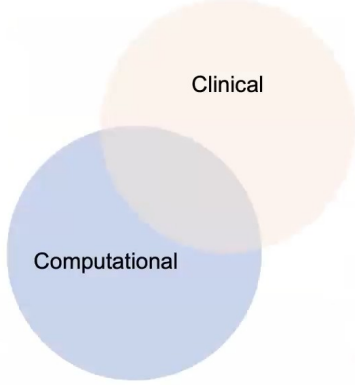


The image shows a video conference interface. The main content area is a white slide with the text "Introductions, cont." at the top left. In the center of the slide is a large, light orange circle containing the word "Clinical". In the top right corner of the video frame, there is a small inset video of a man, identified by a name tag as "Eli Van Allen".

So I was and am a clinician. I see patients on Monday mornings over at Dana Farber, mostly prostate cancer patients at this point.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Introductions, cont.

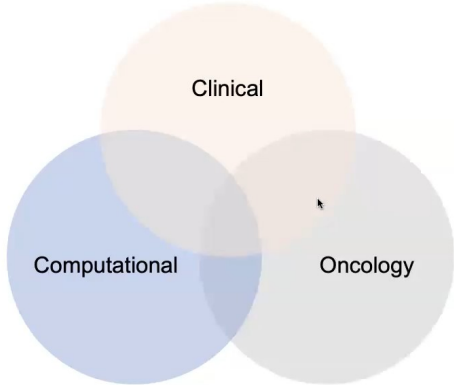


A Venn diagram consisting of two overlapping circles. The top circle is orange and labeled "Clinical". The bottom circle is blue and labeled "Computational". The overlapping area is a lighter shade of the colors. The diagram is centered on a white background.

Eliezer Van Allen

I was and am a computer nerd.

Introductions, cont.



A Venn diagram consisting of three overlapping circles. The top circle is orange and labeled "Clinical". The bottom-left circle is blue and labeled "Computational". The bottom-right circle is grey and labeled "Oncology". The overlapping areas are lighter shades of the colors. The diagram is centered on a white background.

Eliezer Van Allen

I figured out how to steer that energy into the world of cancer, and increasingly immersed myself in not only the biology of cancer, but also the clinical opportunities in cancer. This interface of where effectively **the patient is the model system that we want to study, instead of cell lines in a petri dish, not mice, but rather human experiences.**

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Introductions, cont.

Studying patients directly

A Venn diagram with three overlapping circles. The top circle is orange and labeled 'Clinical'. The bottom-left circle is blue and labeled 'Computational'. The bottom-right circle is grey and labeled 'Oncology'. The text 'Studying patients directly' is positioned above the orange circle. Below the blue circle is the text 'Novel algorithms'. The name 'Eli Van Allen' is visible in a small video inset in the top right corner.

Applying algorithms to the high dimensional data we can generate from our patients to learn as much as we can from them.

Introductions, cont.

Studying patients directly

A Venn diagram with three overlapping circles. The top circle is orange and labeled 'Clinical'. The bottom-left circle is blue and labeled 'Computational'. The bottom-right circle is grey and labeled 'Cancer biology'. The text 'Studying patients directly' is positioned above the orange circle. Below the blue circle is the text 'Novel algorithms'. The name 'Eli Van Allen' is visible in a small video inset in the top right corner.

And use it in combination with biology to maximize the potential of how we can then act on these data to come up with new drug targets, new ways of matching the right patients to the right drugs, and hopefully changing the way we do cancer care. This idea of clinical computational oncology is in a lot of respects a field I made up for myself, because it didn't


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

really exist when I was trying to pursue something like this 10 plus years ago, and now with a lot of excitement it's really snowballed as well and into a field. That is a bit of my intellectual background.

I thought I'd give a background of why we're doing what we're doing.

The best way for me to articulate this is from some of my own experiences as an oncologist.

The Goal



8.98 cm

Toni Choueiri


Elezer Van Allen

The image is a screenshot of a video presentation. It features a slide titled "The Goal" with a CT scan of a patient's abdomen. A yellow vertical line on the scan indicates a measurement of 8.98 cm for a lesion. The video interface includes a small inset of the speaker, Elezer Van Allen, in the top right corner and a name tag for Toni Choueiri in the bottom right corner. A standard video player control bar is visible at the bottom left.

This is not a prostate patient, but rather a kidney cancer patient that I remember seeing as a first year fellow in 2011. This patient had metastatic kidney cancer; the disease had spread from beyond the kidney to multiple other organs. This is a CT scan.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

The Goal




Toni Choueiri

Eliezer Van Allen

This is not a test. So I will just help you by drawing a big red circle around a huge metastatic deposit that this patient experienced.

The Goal



Life expectancy: 0-3 months


Toni Choueiri

Eliezer Van Allen

The life expectancy for his patient was zero to three months after having progressed, on the therapeutics we had at the time.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]


The Goal



Cancer immunotherapy

Life expectancy: 0-3 months


Toni Choueiri



Eli Van Allen


We got this patient onto a clinical trial, largely in an ad hoc, not informed way. It was just because it was what we were able to find. This patient happened to get on an immunotherapy trial, which we didn't know back then.

The Goal




Cancer immunotherapy

Five years later...



Life expectancy: 0-3 months

Toni Choueiri

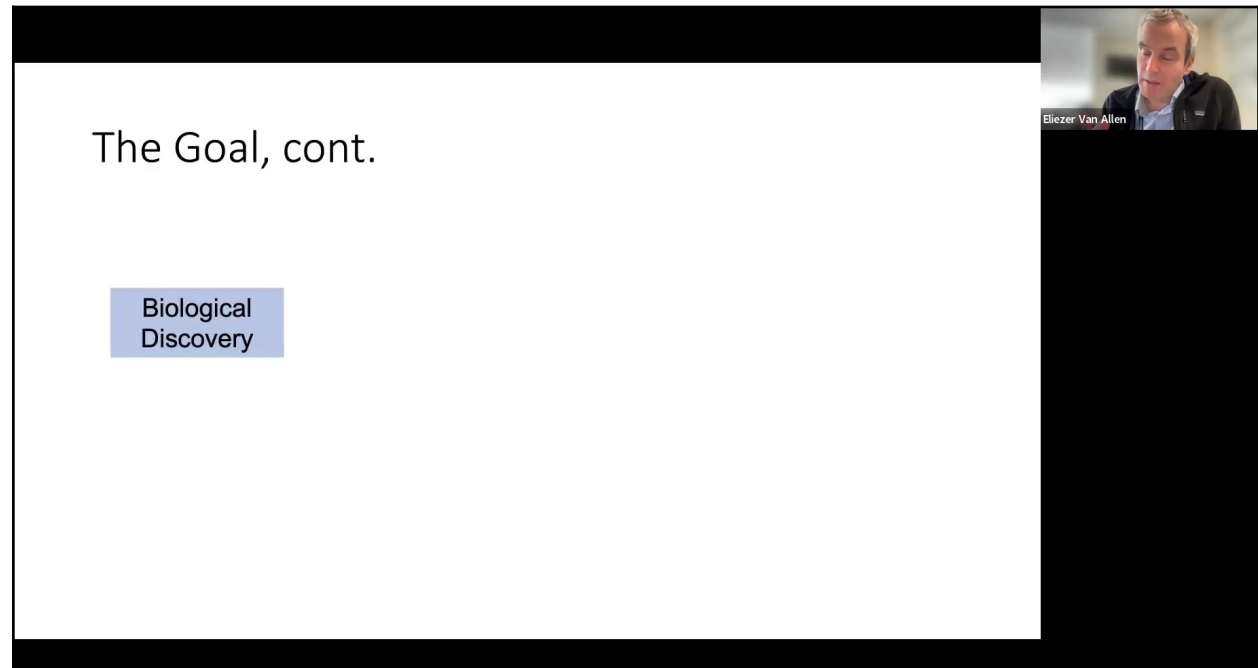


Eli Van Allen

But now we know it works like gangbusters for some patients, but not all. And many years later, actually, this is a five year scan, this patient is effectively cured now of a metastatic solid tumor.


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

This is what we're really trying to aspire to do across the board. This is the real overarching goal. But where does data come in? And how does data potentially drive or accelerate opportunities in these spaces?



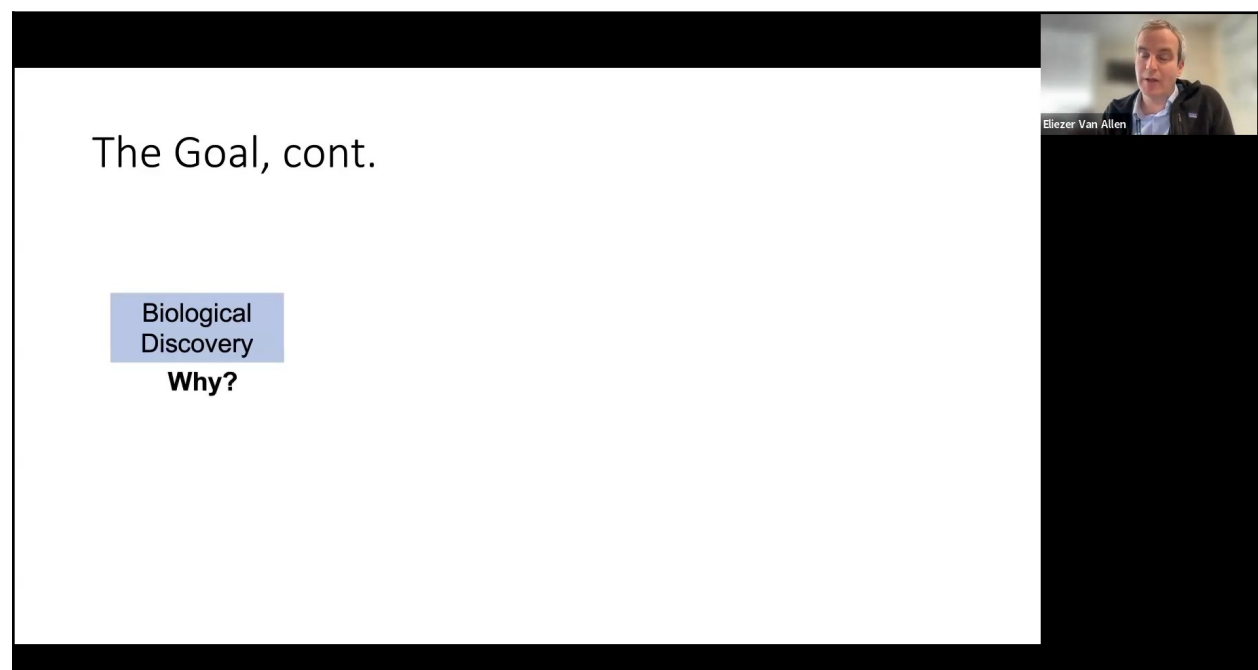
The Goal, cont.

Biological
Discovery



Eli Van Allen


One space is really around biological discovery. How can you harness tons of data and make new algorithms to find the patterns in huge amounts of cancer patient data to find new drug targets?



The Goal, cont.

Biological
Discovery

Why?



Eli Van Allen


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

And understand why a disease occurs? Why does it become resistant to the drugs we give it?
How can we then intervene? That's really at the core on the really pure research side.

The Goal, cont.

Biological Discovery
Why?

Clinical Evaluation




Eli Van Allen

But it's also potentially helpful for clinical evaluation.

The Goal, cont.

Biological Discovery
Why?

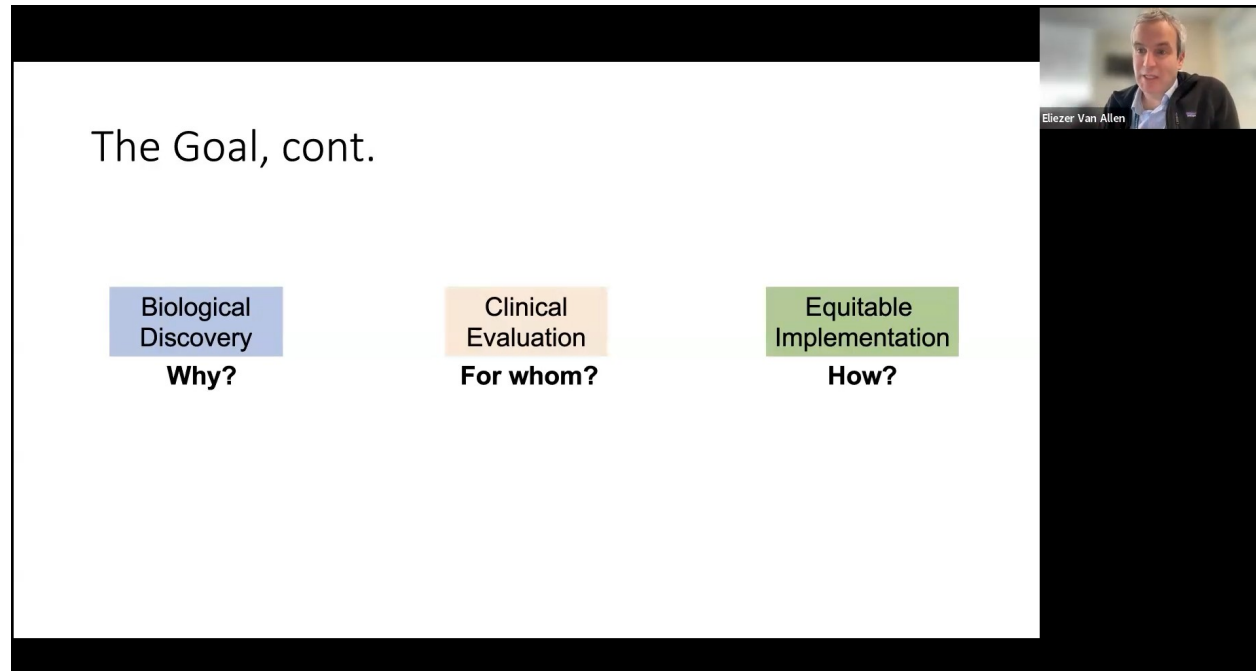
Clinical Evaluation
For whom?



Eli Van Allen

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

How do we figure out who should get what drug? Where should that decision-making be done?
What kinds of ways can algorithms help guide them?

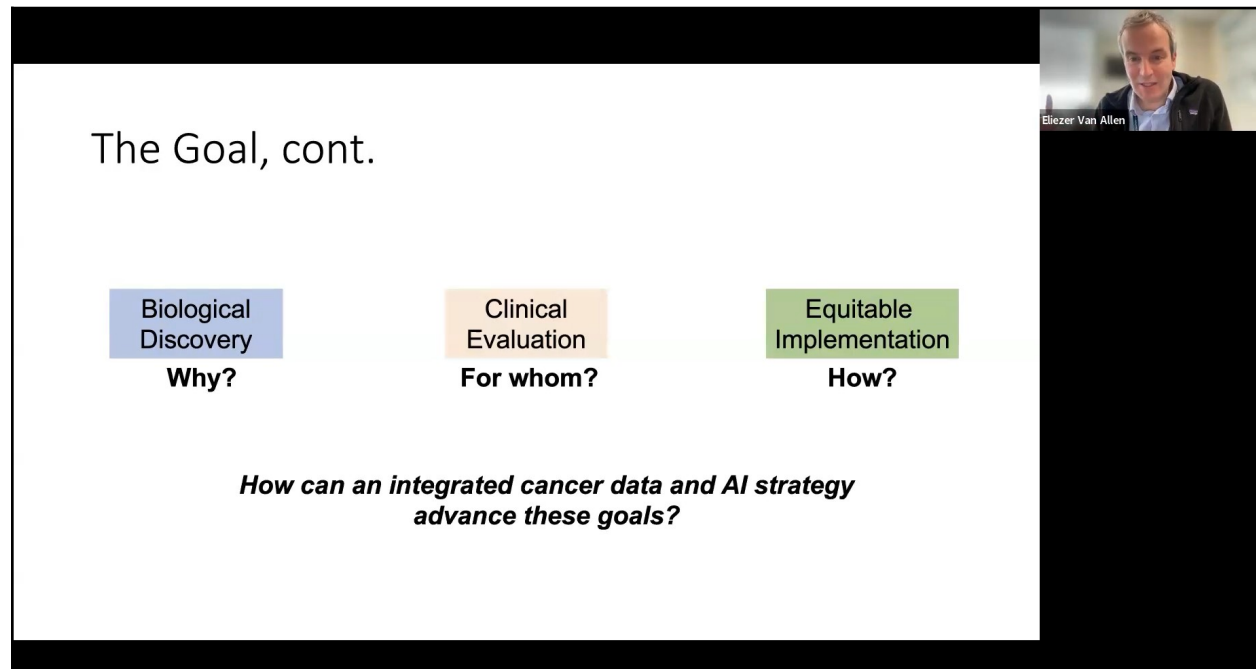


The Goal, cont.

Biological Discovery	Clinical Evaluation	Equitable Implementation
Why?	For whom?	How?

Elezer Van Allen

Finally, how can we do this in a way that's equitable, so that everyone can actually experience this, and not just a select few patients who happen to lock into a clinical trial or happen to be living near a big quaternary care cancer center, but really, for everyone near and far?



The Goal, cont.

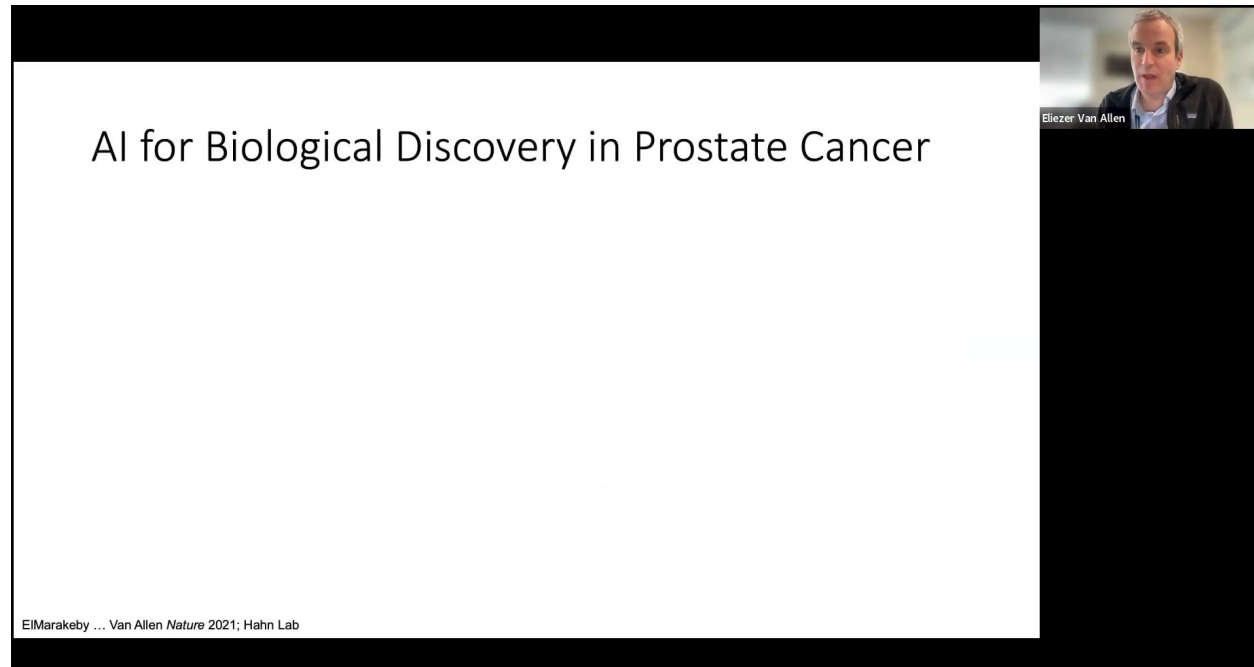
Biological Discovery	Clinical Evaluation	Equitable Implementation
Why?	For whom?	How?

How can an integrated cancer data and AI strategy advance these goals?

Elezer Van Allen

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

How do we do this? How can we integrate patient data and technological advances, especially artificial intelligence? I know a topic du jour that was requested is that I touch on here to actually advance these three goals: (1) why, (2) for whom, and (3) how. These are really at the heart of everything we're doing in oncology. But this is, again, a data-centric way into thinking about this. We've been thinking about this a ton for many, many years.



AI for Biological Discovery in Prostate Cancer

ElMarakeby ... Van Allen *Nature* 2021; Hahn Lab

One of the spaces we've really pursued aggressively is in prostate cancer. I've had the privilege of helping to contribute to some of the largest prostate cancer genomic studies, where we do sequencing of all 20,000 genes, and we generate massive quantities of data from every patient that we can – both the genomics of their tumor, but also the genomes of the patient as they were born – and integrate all these data for many wonderful discoveries. But one thing we realize is that we were pretty bad at understanding the genomics of the differences between tumors that were lethal and tumors that were not lethal.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



AI for Biological Discovery in Prostate Cancer

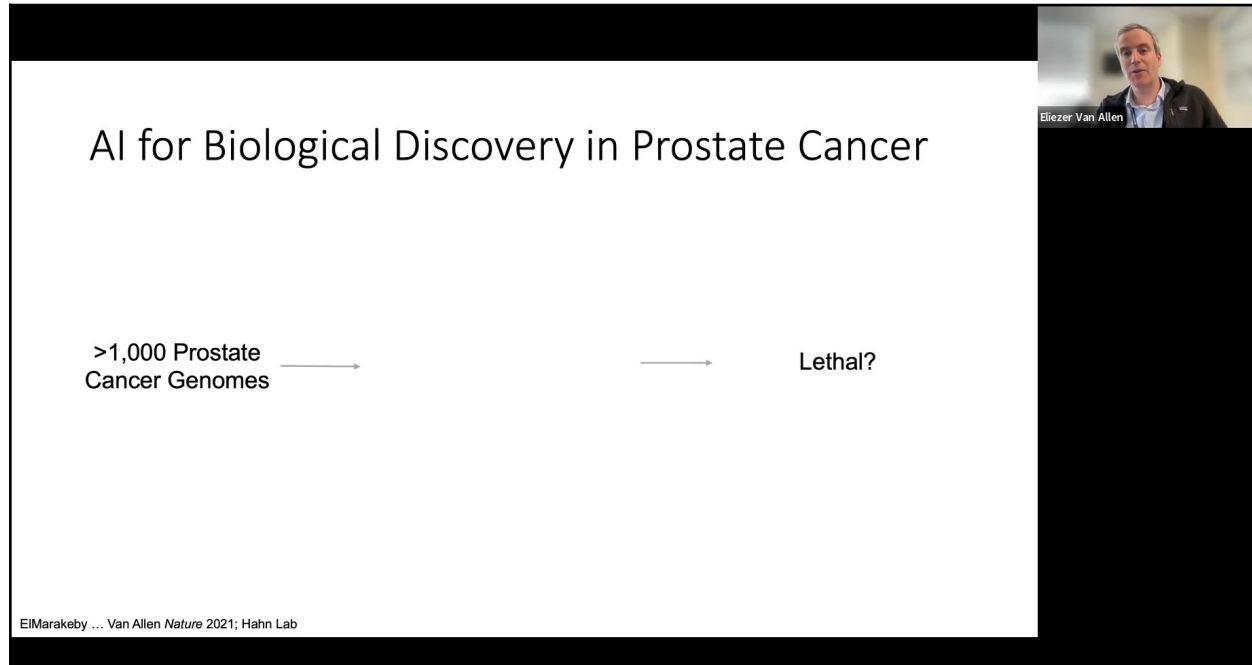
>1,000 Prostate
Cancer Genomes

EliMarakeby ... Van Allen *Nature* 2021; Hahn Lab

Eli Van Allen

This slide features a white background with a black border. The title "AI for Biological Discovery in Prostate Cancer" is centered at the top. Below it, the text ">1,000 Prostate Cancer Genomes" is displayed. A small video inset in the top right corner shows Eli Van Allen speaking. A small citation "EliMarakeby ... Van Allen Nature 2021; Hahn Lab" is located in the bottom left corner.

To frame the question in a 30,000 foot view, we're trying to tackle that “why” question, that biological discovery question. We had organized over 1000 prostate cancer genomes, and we wanted to figure out:



AI for Biological Discovery in Prostate Cancer

>1,000 Prostate
Cancer Genomes → → Lethal?

EliMarakeby ... Van Allen *Nature* 2021; Hahn Lab

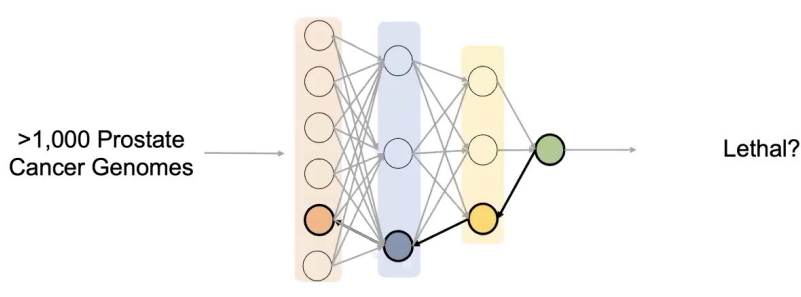
Eli Van Allen

This slide is similar to the first one but includes a flow diagram. The text ">1,000 Prostate Cancer Genomes" is followed by two right-pointing arrows, leading to the text "Lethal?". The rest of the slide content, including the title, video inset, and citation, remains the same.

“Is this going to kill you or not?” And hopefully, use that information, not only for prediction, but also to figure out what the genetic lesions actually were, and make drugs for those things.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI for Biological Discovery in Prostate Cancer



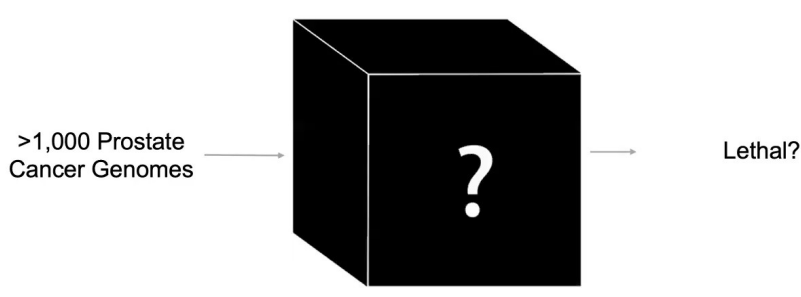
>1,000 Prostate Cancer Genomes → Lethal?

EIMarakeby ... Van Allen Nature 2021; Hahn Lab

Eliezer Van Allen

So increasingly, many people have been interested in solving this problem using AI. This is a generic cartoon of an AI neural network. The details are not important for the purposes of this talk. A lot of people are doing this.

AI for Biological Discovery in Prostate Cancer



>1,000 Prostate Cancer Genomes → Lethal?

EIMarakeby ... Van Allen Nature 2021; Hahn Lab

Eliezer Van Allen

The problem with this kind of approach is that these tend to be like black boxes. You might be able to make an algorithm to tell you which ones are bad or not bad. But you're not going to

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

understand why. If what you want to do is not only stratify, but also come up with new drug targets and new ways of intervening, you need to know why.

AI for Biological Discovery in Prostate Cancer

>1,000 Prostate Cancer Genomes →

→ Lethal?

Eli Van Allen

EIMarakeby ... Van Allen Nature 2021; Hahn Lab

We had this idea that we could merge biology, the whole world of biology, everything that has been over the last 100 200 500 years of everything, everyone is known about, the way our bodies are put together, and engineer that into a neural network. We combined the two things up front. Instead of having these neural networks being in a completely connected but completely uninterpretable form,

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI for Biological Discovery in Prostate Cancer

The diagram illustrates a biological network. On the left, an arrow labeled ">1,000 Prostate Cancer Genomes" points to a vertical column of six circles representing genes, labeled G1 through G6. G1, G2, G3, and G4 are light blue, while G5 and G6 are dark blue. To the right of the genes is a vertical column of three circles representing pathways, labeled P1, P2, and P3. P1 and P2 are light blue, while P3 is dark blue. Arrows connect G1 to P1, G2 to P1, G3 to P1, G4 to P2, G5 to P3, and G6 to P3. To the right of the pathways is a vertical column of three circles representing biological processes, labeled BP1, BP2, and BP3. BP1 and BP2 are light blue, while BP3 is dark blue. Arrows connect P1 to BP1, P2 to BP2, and P3 to BP3. A final arrow points from BP3 to a single green circle, which is labeled "Lethal?".

EIMarakeby ... Van Allen *Nature* 2021; Hahn Lab

Eliezer Van Allen

we layered them such that one layer was the genes, and the genes connected into pathways,

AI for Biological Discovery in Prostate Cancer

The diagram illustrates a biological network. On the left, an arrow labeled ">1,000 Prostate Cancer Genomes" points to a vertical column of six circles representing genes, labeled G1 through G6. G1, G2, G3, and G4 are light blue, while G5 and G6 are dark blue. To the right of the genes is a vertical column of three circles representing pathways, labeled P1, P2, and P3. P1 and P2 are light blue, while P3 is dark blue. Arrows connect G1 to P1, G2 to P1, G3 to P1, G4 to P2, G5 to P3, and G6 to P3. To the right of the pathways is a vertical column of three circles representing biological processes, labeled BP1, BP2, and BP3. BP1 and BP2 are light blue, while BP3 is dark blue. Arrows connect P1 to BP1, P2 to BP2, and P3 to BP3. A final arrow points from BP3 to a single green circle, which is labeled "Lethal?".


EIMarakeby ... Van Allen *Nature* 2021; Hahn Lab

Eliezer Van Allen

and the pathways connected in the processes. In this way, you would create a fully interpretable neural network, that you could then not only steer at the exact same data set that I talked about earlier, and use it for stratification and prediction, but also open that box and understand why. We actually did that, in this context, using what we call “a biologically-informed neural network”

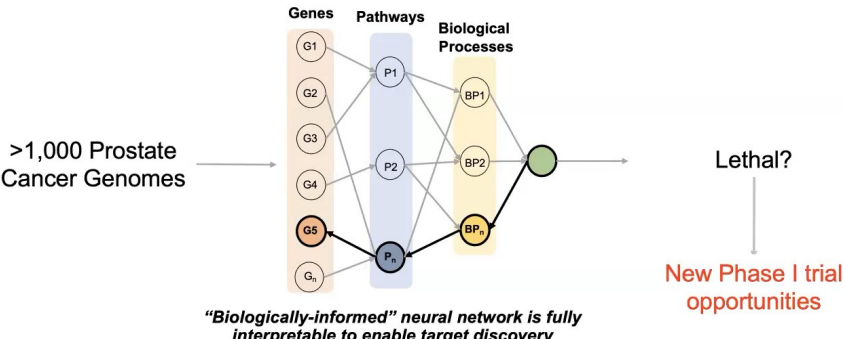
“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

based entirely off of data generated from cancer patients treated at a multitude of sites across the United States.



Eli Van Allen

AI for Biological Discovery in Prostate Cancer



The diagram illustrates a biologically-informed neural network. It starts with an input of '>1,000 Prostate Cancer Genomes' on the left. This input feeds into a network of three layers: Genes (G1, G2, G3, G4, G5, G_n), Pathways (P1, P2, P_n), and Biological Processes (BP1, BP2, BP_n). The nodes are interconnected, with arrows showing the flow of information from Genes to Pathways and then to Biological Processes. A final green node on the right represents the output of the network. Below the diagram, the text reads: "Biologically-informed" neural network is fully interpretable to enable target discovery. To the right of the diagram, the text asks "Lethal?" and points to "New Phase I trial opportunities".

>1,000 Prostate Cancer Genomes

Lethal?

New Phase I trial opportunities


"Biologically-informed" neural network is fully interpretable to enable target discovery

ElMarakeby ... Van Allen *Nature* 2021; Hahn Lab

Dr. Alok Tewari (DFCI Phase I)

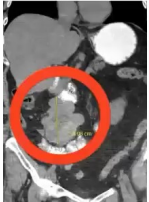
We found something interesting, did some experiments, and that snowballed into phase one clinical trials, repurposing drugs that we knew were not relevant for prostate cancer, but now because we just simply missed it, think they are relevant for a subset of prostate cancer. This, in a lot of respects, is how we can actually steer patient data and advances in data analytics, to come up with new drug targets, and new hypotheses that might help patients down the road.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



Eli Van Allen


AI for Clinical Evaluation



Nyman ... Van Allen, *Cell Rep Med* 2023

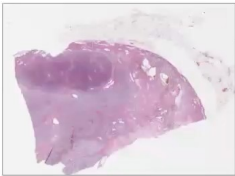

Eli Van Allen 13:48

In addition to the “why,” we also want to figure out the “for whom”. There's a lot of really exciting data science innovations happening in that space as well. One of them is around pictures. I showed you a picture earlier from a CT scan from a patient.



Eli Van Allen

AI for Clinical Evaluation




Nyman ... Van Allen, *Cell Rep Med* 2023

Another kind of picture that is very commonly seen in cancer care is pathology pictures. When the tumor gets taken out, the way the diagnosis is made, is the pathologist looks under the

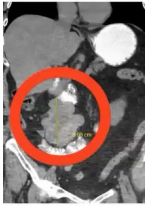
“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

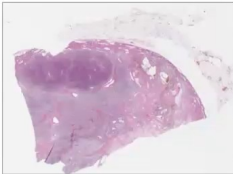
microscope at the tissue that's been removed from your body and tells you what they see. That is effectively how the diagnosis is made. But these are all pictures. In the data science world, especially with machine learning, and AI, these kinds of pictures are already being used for all sorts of other cool things.

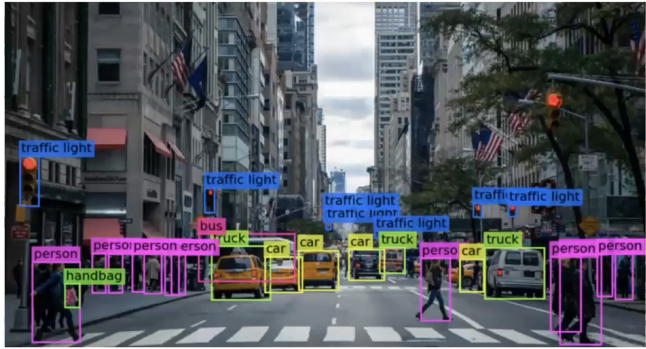


Eli Van Allen

AI for Clinical Evaluation





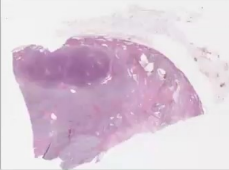


Nyman ... Van Allen, *Cell Rep Med* 2023


For example, those of you who live in certain suburban or urban areas have already seen self-driving cars out and about. I was just home in Los Angeles for a family trip. I was struck by how many Waymo cars I was seeing driving around the city – not actually struck, but metaphorically, not literally. These are image processing problems too. A lot of this stuff has already been solved from a technology perspective, but it has not been steered into a cancer perspective.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI for Clinical Evaluation, cont.


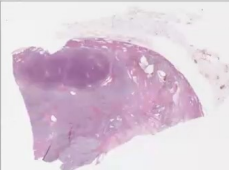


Nyman ... Van Allen, *Cell Rep Med* 2023
(TIL: Tumor Infiltrating Lymphocyte)




So what we did was we took, again, a different kind of data from patients, in this **case cancer patients who have kidney cancer, treated with cancer immunotherapies.**

AI for Clinical Evaluation, cont.



Graph Representation

Nyman ... Van Allen, *Cell Rep Med* 2023
(TIL: Tumor Infiltrating Lymphocyte)



We applied these computer vision algorithms to parse the data into interesting representations of tumor biology,

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI for Clinical Evaluation, cont.

The diagram illustrates the process of AI for clinical evaluation. It starts with a 2D histology image of a tumor section. This image is processed into two different representations: a Graph Representation, which shows the tumor structure with colored regions (red and blue) and a network of nodes and edges, and a TIL Density Map, which shows the density of Tumor Infiltrating Lymphocytes (TILs) in the tumor region.

Nyman ... Van Allen, *Cell Rep Med* 2023
(TIL: Tumor Infiltrating Lymphocyte)

Eliezer Van Allen

as well as the biology of immune cells and the microenvironment of these tumors, and used these two pieces of information

AI for Clinical Evaluation, cont.

The diagram illustrates the process of AI for clinical evaluation. It starts with a 2D histology image of a tumor section. This image is processed into two different representations: a Graph Representation, which shows the tumor structure with colored regions (red and blue) and a network of nodes and edges, and a TIL Density Map, which shows the density of Tumor Infiltrating Lymphocytes (TILs) in the tumor region. Both representations lead to a prediction of immunotherapy response.

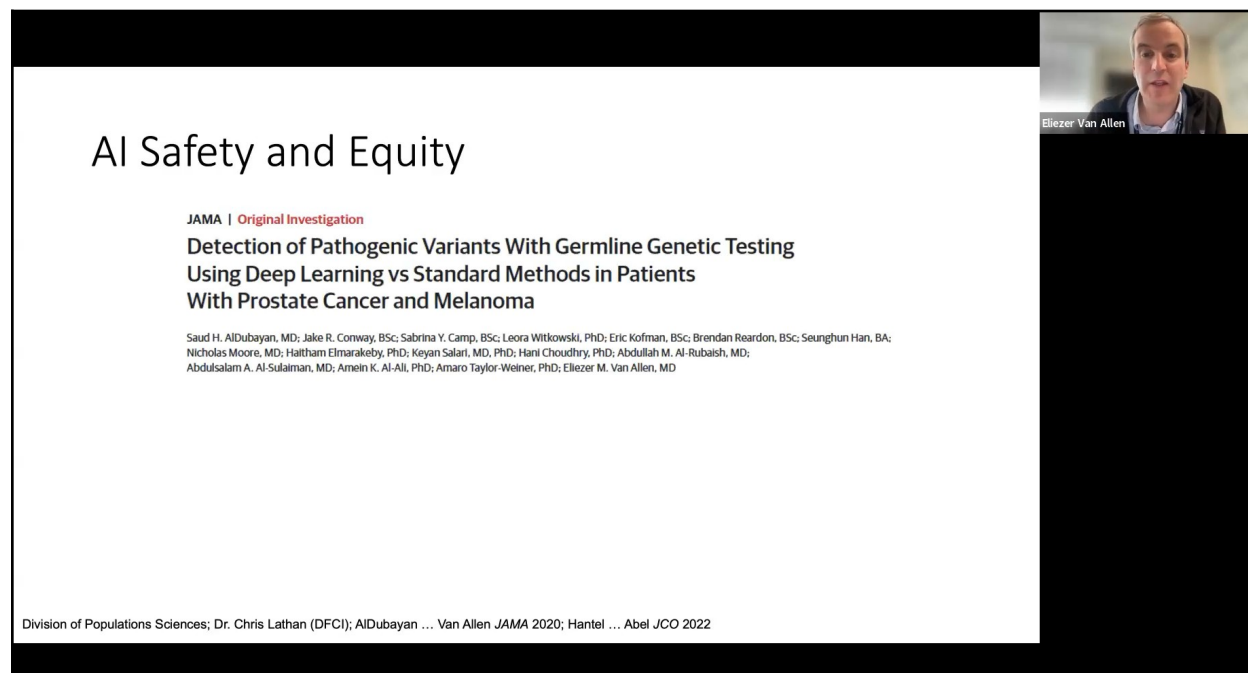
Successfully predicting immunotherapy response directly from 2D images

Nyman ... Van Allen, *Cell Rep Med* 2023
(TIL: Tumor Infiltrating Lymphocyte)

Eliezer Van Allen

to predict which patients are most likely to benefit from immunotherapy, simply from applying algorithms to two dimensional pictures. There are other nuances of the study that actually relate these pictures to the genetics. I won't have time to go into this, but I'm just providing a menu of the things that are out there.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



The slide displays the title "AI Safety and Equity" and the text of a JAMA article. The article title is "Detection of Pathogenic Variants With Germline Genetic Testing Using Deep Learning vs Standard Methods in Patients With Prostate Cancer and Melanoma". The authors listed are Saud H. Aldubayan, MD; Jake R. Conway, BSc; Sabrina Y. Camp, BSc; Leora Wittkowski, PhD; Eric Kofman, BSc; Brendan Reardon, BSc; Seunghun Han, BA; Nicholas Moore, MD; Haitham Elmarakeby, PhD; Keyan Salari, MD, PhD; Hani Choudhry, PhD; Abdullah M. Al-Rubaish, MD; Abdulsalam A. Al-Sulaiman, MD; Amein K. Al-Ali, PhD; Amaro Taylor-Weiner, PhD; and Eliezer M. Van Allen, MD. At the bottom of the slide, it says "Division of Populations Sciences; Dr. Chris Lathan (DFCI); AIDubayan ... Van Allen JAMA 2020; Hantel ... Abel JCO 2022".


AI Safety and Equity

JAMA | Original Investigation

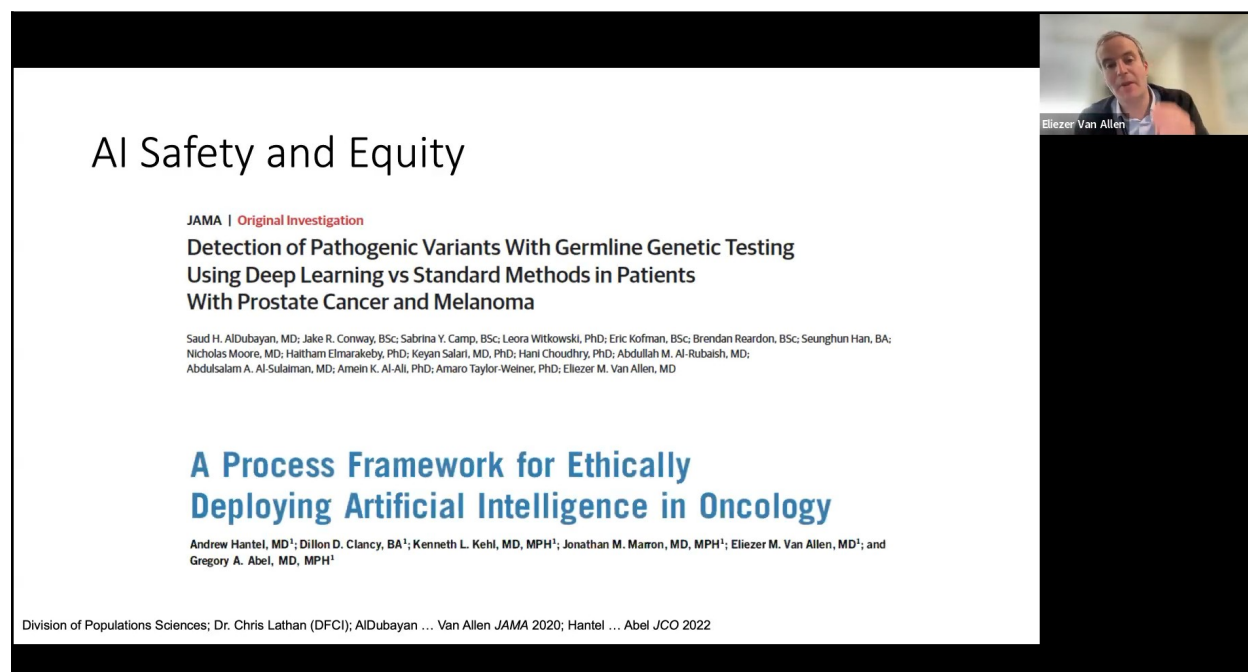
Detection of Pathogenic Variants With Germline Genetic Testing Using Deep Learning vs Standard Methods in Patients With Prostate Cancer and Melanoma

Saud H. Aldubayan, MD; Jake R. Conway, BSc; Sabrina Y. Camp, BSc; Leora Wittkowski, PhD; Eric Kofman, BSc; Brendan Reardon, BSc; Seunghun Han, BA; Nicholas Moore, MD; Haitham Elmarakeby, PhD; Keyan Salari, MD, PhD; Hani Choudhry, PhD; Abdullah M. Al-Rubaish, MD; Abdulsalam A. Al-Sulaiman, MD; Amein K. Al-Ali, PhD; Amaro Taylor-Weiner, PhD; Eliezer M. Van Allen, MD

Division of Populations Sciences; Dr. Chris Lathan (DFCI); AIDubayan ... Van Allen JAMA 2020; Hantel ... Abel JCO 2022



We talked about “why”. We talked about “for whom”. What about “how”? How do we do this in a way that's safe and equitable? Without going into the details, I'll just say, we have been developing paradigms for testing these algorithms in clinical care, and figuring out where some of these newer data science strategies are useful. Specifically, in this case, **using Google's algorithm for figuring out which patients with prostate cancer or melanoma have what's called germline or inherited genetic events that are actionable.**



The slide displays the title "AI Safety and Equity" and the text of a JAMA article. The article title is "Detection of Pathogenic Variants With Germline Genetic Testing Using Deep Learning vs Standard Methods in Patients With Prostate Cancer and Melanoma". The authors listed are Saud H. Aldubayan, MD; Jake R. Conway, BSc; Sabrina Y. Camp, BSc; Leora Wittkowski, PhD; Eric Kofman, BSc; Brendan Reardon, BSc; Seunghun Han, BA; Nicholas Moore, MD; Haitham Elmarakeby, PhD; Keyan Salari, MD, PhD; Hani Choudhry, PhD; Abdullah M. Al-Rubaish, MD; Abdulsalam A. Al-Sulaiman, MD; Amein K. Al-Ali, PhD; Amaro Taylor-Weiner, PhD; and Eliezer M. Van Allen, MD. Below this, there is a new section titled "A Process Framework for Ethically Deploying Artificial Intelligence in Oncology" by Andrew Hantel, MD¹; Dillon D. Clancy, BA¹; Kenneth L. Kehl, MD, MPH¹; Jonathan M. Marron, MD, MPH¹; Eliezer M. Van Allen, MD¹; and Gregory A. Abel, MD, MPH¹. At the bottom of the slide, it says "Division of Populations Sciences; Dr. Chris Lathan (DFCI); AIDubayan ... Van Allen JAMA 2020; Hantel ... Abel JCO 2022".

AI Safety and Equity

JAMA | Original Investigation


Detection of Pathogenic Variants With Germline Genetic Testing Using Deep Learning vs Standard Methods in Patients With Prostate Cancer and Melanoma

Saud H. Aldubayan, MD; Jake R. Conway, BSc; Sabrina Y. Camp, BSc; Leora Wittkowski, PhD; Eric Kofman, BSc; Brendan Reardon, BSc; Seunghun Han, BA; Nicholas Moore, MD; Haitham Elmarakeby, PhD; Keyan Salari, MD, PhD; Hani Choudhry, PhD; Abdullah M. Al-Rubaish, MD; Abdulsalam A. Al-Sulaiman, MD; Amein K. Al-Ali, PhD; Amaro Taylor-Weiner, PhD; Eliezer M. Van Allen, MD

A Process Framework for Ethically Deploying Artificial Intelligence in Oncology

Andrew Hantel, MD¹; Dillon D. Clancy, BA¹; Kenneth L. Kehl, MD, MPH¹; Jonathan M. Marron, MD, MPH¹; Eliezer M. Van Allen, MD¹; and Gregory A. Abel, MD, MPH¹

Division of Populations Sciences; Dr. Chris Lathan (DFCI); AIDubayan ... Van Allen JAMA 2020; Hantel ... Abel JCO 2022



“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

But it also means coming up with ways to engage with our ethicists. Because this is going to open up a huge can of worms that we can go into in the discussion in terms of what this is going to mean for cancer care. We need to think very carefully about how we do this in a way that's thoughtful and doesn't harm people as we're moving this new technology into the clinic.

The Goal, cont.

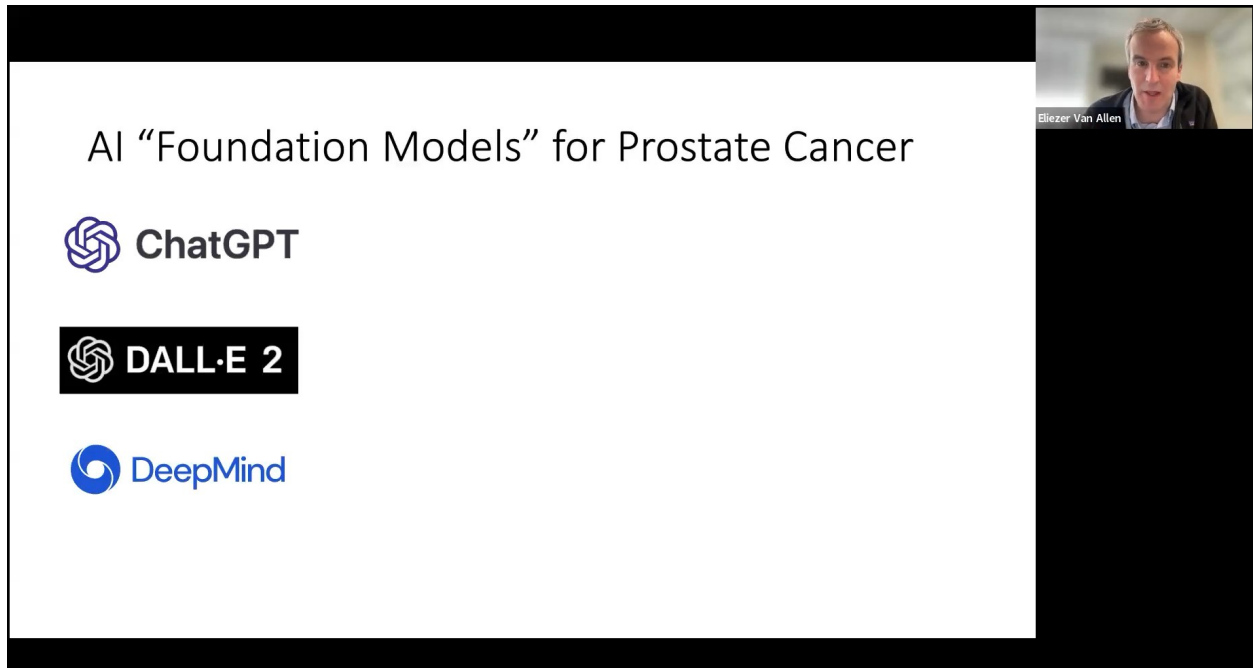
Biological Discovery	Clinical Evaluation	Equitable Implementation
Why?	For whom?	How?

These efforts hint at the larger potential...


Elezer Van Allen


We talked about “why”, we talked about “for whom”, and we're talking about “how”, but if you take nothing else from this presentation, I hope you get the sense that we're just barely scratching the surface of what's possible here.


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



AI “Foundation Models” for Prostate Cancer

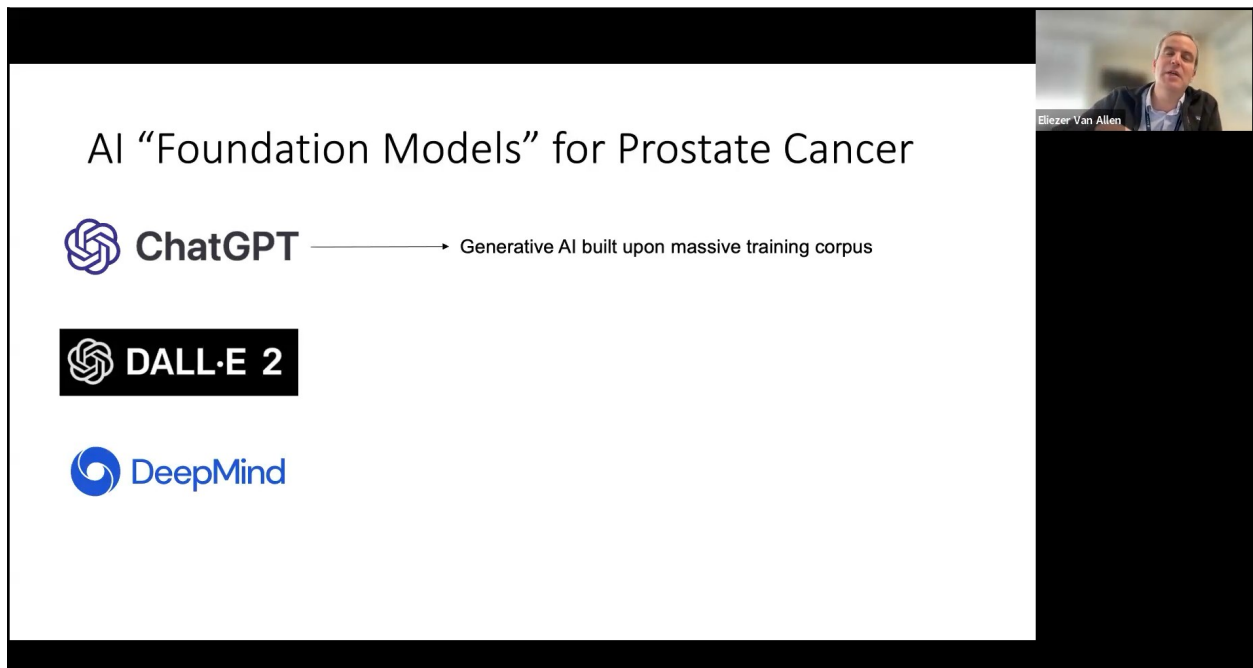
 ChatGPT

 DALL·E 2


 DeepMind


Eliazer Van Allen


What do I mean? Well, you may have started to hear in the press these things called foundation models. We wonder, is there such a thing for cancer, generally, or prostate cancer in particular? What do I mean by these foundation models? Things like ChatGPT, or DALL-E 2. If you have not played with these things, I strongly encourage you to. It's a lot of fun.



AI “Foundation Models” for Prostate Cancer

 ChatGPT → Generative AI built upon massive training corpus

 DALL·E 2


 DeepMind

Eliazer Van Allen


It's a kind of AI that's built off of tons of data. There is so much that maybe there are some legal issues. But that's a different story for another time.


“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI “Foundation Models” for Prostate Cancer

 **ChatGPT** → Generative AI built upon massive training corpus

↓
Validating existing concepts


 **DALL-E 2**

 **DeepMind**


Eliizer Van Allen


What you can do with these things is two kinds of fun tasks. One is easy: it can validate existing concepts.

AI “Foundation Models” for Prostate Cancer

 **ChatGPT** → Generative AI built upon massive training corpus

↓
Validating existing concepts

 **DALL-E 2**

 **DeepMind**

E what is the initial treatment management for grade 3b follicular lymphoma with avid lymph nodes above and below the diaphragm

Eliizer Van Allen

I can type a question about a disease that I have not thought about much because I do mostly prostate cancer in ChatGPT,

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

The slide features the title "AI “Foundation Models” for Prostate Cancer" at the top. Below the title are three logos: OpenAI ChatGPT, DALL-E 2, and DeepMind. To the right of the logos is a diagram. A horizontal arrow points from the ChatGPT logo to the text "Generative AI built upon massive training corpus". A vertical arrow points down from this text to the text "Validating existing concepts". Below this, there are two text boxes. The first box contains a question: "what is the initial treatment management for grade 3b follicular lymphoma with avid lymph nodes above and below the diaphragm". The second box contains an answer: "Grade 3B follicular lymphoma is considered a high-grade variant of follicular lymphoma with more aggressive behavior. The initial treatment management for grade 3B follicular lymphoma with avid lymph nodes above and below the diaphragm typically involves a combination of chemotherapy and immunotherapy. Here's a general outline of the treatment approach:". In the top right corner of the slide, there is a small video inset of Eli Van Allen.

and it will give me an answer that sometimes is correct. There are some huge caveats there.

This slide is identical to the one above, but with an additional red arrow pointing from the "Generative AI built upon massive training corpus" text to the text "Deriving new (unseen) concepts". This text is positioned above the second text box in the diagram. In the top right corner of the slide, there is a small video inset of Eli Van Allen.

What's perhaps more provocative is that you can use it to derive novel hypotheses and novel insights that you could then test in a way that would be impossible to scale in any other dimension. So what's the ChatGPT equivalent?

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

As an example, my wife, who's a former professional tennis player and Britney Spears fan, asked me about ChatGPT. She was like, “What's the big deal?” And I said, “Let me show you.” And I typed into ChatGPT, “Can you write a song in the style of Britney Spears about the top five men's singles tennis players?” She's like, “Yeah, sure. There's no way this is going to work.”

AI “Foundation Models” for Prostate Cancer

The diagram illustrates the application of AI foundation models. On the left, logos for ChatGPT, DALL-E 2, and DeepMind are shown. An arrow points from these models to the text “Generative AI built upon massive training corpus”. From this text, two arrows branch out to two columns of examples:

- Validating existing concepts:**
 - Q: what is the initial treatment management for grade 3b follicular lymphoma with avid lymph nodes above and below the diaphragm
 - A: Grade 3B follicular lymphoma is considered a high-grade variant of follicular lymphoma with more aggressive behavior. The initial treatment management for grade 3B follicular lymphoma with avid lymph nodes above and below the diaphragm typically involves a combination of chemotherapy and immunotherapy. Here's a general outline of the treatment approach:
- Deriving new (unseen) concepts:**
 - Q: can you now write a song in the style of britney spears about the top 5 men's singles tennis players
 - A: Sure, here's a fun and playful pop song in the style of Britney Spears about the top 5 men's singles tennis players:
Verse 1:
I'm checking out the courtside view,
Watching tennis stars shine so bright and true,
Got my shades on, feeling so cool,
Gonna sing about the top 5, that's the rule.
Chorus:
Djokovic, Medvedev, and Nadal,
Tsitsipas and Zverev, they're standing tall,
On the court, they're the kings of the game,
Serving up a storm, they're on their way to fame.

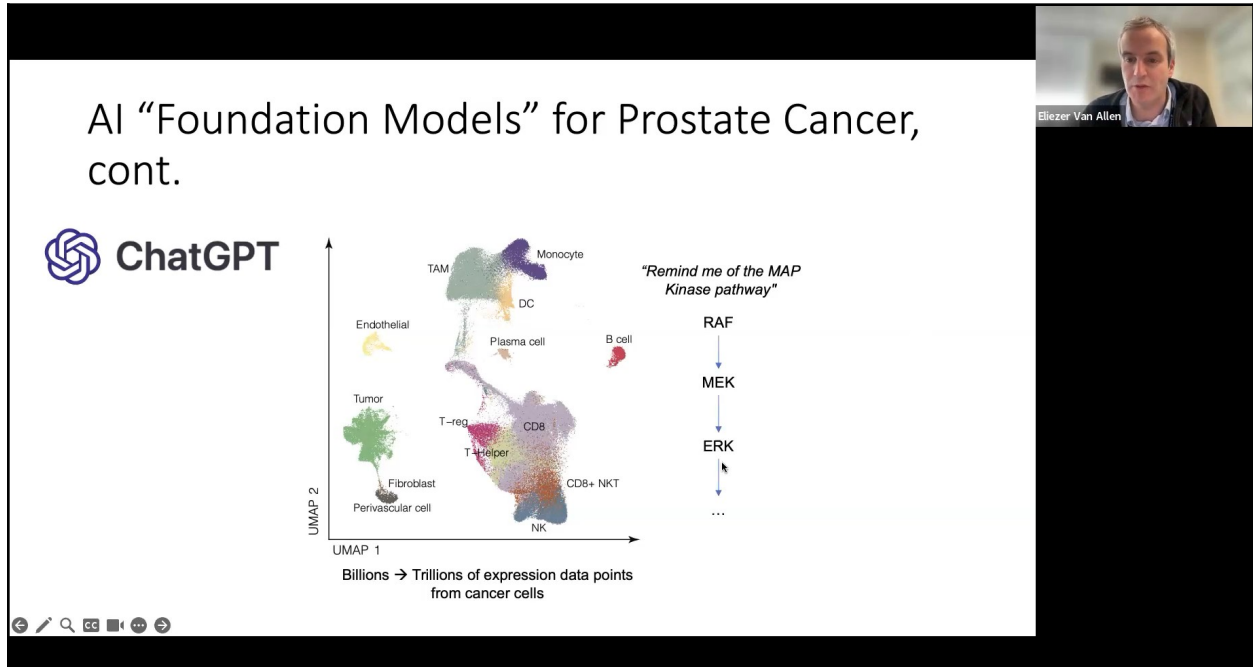
And of course, as you guys probably know by now, in 10 seconds it spit out this thing. But it was pretty right.

AI “Foundation Models” for Prostate Cancer, cont.

The UMAP plot displays a complex arrangement of cell clusters. The x-axis is labeled “UMAP 1” and the y-axis is labeled “UMAP 2”. Below the plot, it states “Billions → Trillions of expression data points from cancer cells”. The clusters are labeled with various cell types: TAM, Monocyte, DC, Plasma cell, B cell, Tumor, T-reg, CD8, T-helper, CD8+ NKT, NK, Fibroblast, Pervascular cell, and Endothelial.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

So why am I telling you this? Well, what if instead of being trained off of words and sentences and books, we were training off of billions or trillions of measurements from cancer cells, which we're now scaling up to generate and could conceivably have in the not too distant future?



You could then query such a model to ask it to validate things that you know to be true, and make sure that it actually is representing what you think it's representing. This is a niche biology question.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

AI “Foundation Models” for Prostate Cancer, cont.

ChatGPT

UMAP 2

UMAP 1

Billions → Trillions of expression data points from cancer cells

“Remind me of the MAP Kinase pathway”

RAF
↓
MEK
↓
ERK
↓
...

“Propose Gene_x signaling in prostate cancer”

Gene_x
↓
Gene_y
↙ ↘
Gene_A ERK
↓ ↓
... ...

Unprecedented opportunities for biological and drug discovery...

More provocatively, you could ask things like, for the 15,000 or so genes that are not represented in the kinds of knowledge bases that are at the core of some of the algorithms I showed earlier, have it generate some hypotheses for you that you could then have tested, that could yield some unprecedented opportunities for discovery.

Translating AI Guidance into Clinical Practice


Match to Genomic Trial
(MatchMiner)

Drs. Ken Kehl, Ethan Cerami, Mike Hassett, Charlotta Lindvall, Bill Lotter, many others

What about translating this kind of AI into clinical practice? Well, there are a million examples. One of the ones that's near and dear to our hearts at Dana Farber is using this information to

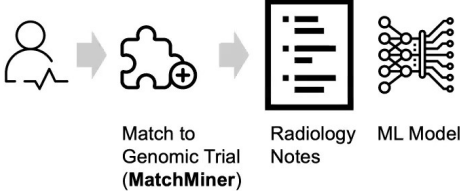
“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

match patients to clinical trials based on their genetics, for which we have algorithms that are already in deployment.



Eli Van Allen


Translating AI Guidance into Clinical Practice



```
graph LR; Patient[Patient] --> Match[Match to Genomic Trial (MatchMiner)]; Match --> Radiology[Radiology Notes]; Radiology --> ML[ML Model]
```

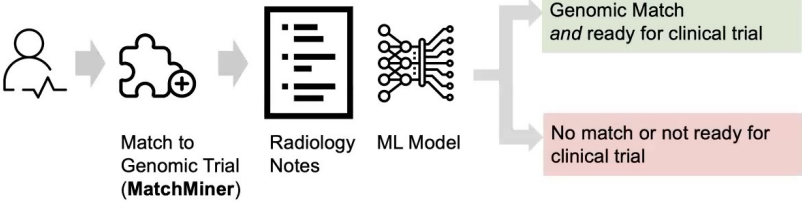
Drs. Ken Kehl, Ethan Cerami, Mike Hassett, Charlotta Lindvall, Bill Lotter, many others

Pairing those with algorithms that go into electronic health records and



Eli Van Allen

Translating AI Guidance into Clinical Practice

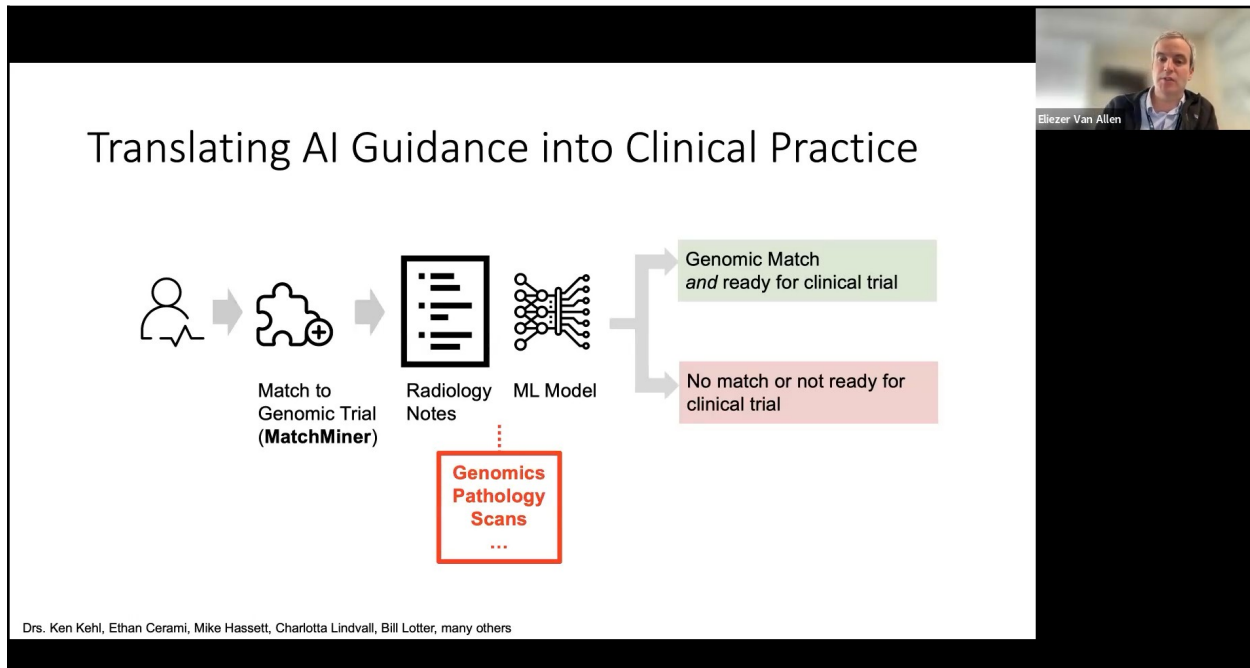


```
graph LR; Patient[Patient] --> Match[Match to Genomic Trial (MatchMiner)]; Match --> Radiology[Radiology Notes]; Radiology --> ML[ML Model]; ML --> MatchOut[Genomic Match and ready for clinical trial]; ML --> NoMatch[No match or not ready for clinical trial]
```

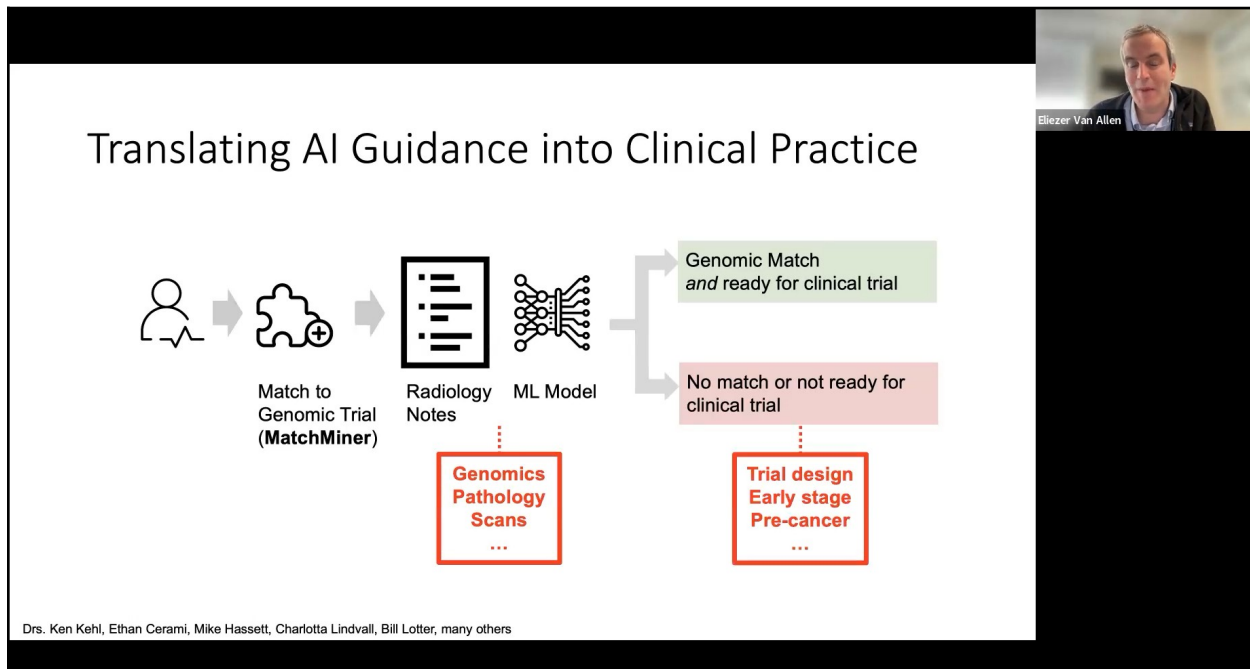
Drs. Ken Kehl, Ethan Cerami, Mike Hassett, Charlotta Lindvall, Bill Lotter, many others

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

finds patients most likely to not only be a match on the genomics for a trial, but also be at a point in their clinical care based on the CT scans and other radiology information that they need a clinical trial. This schematic that I'm showing here is already happening. It's ongoing at Farber.

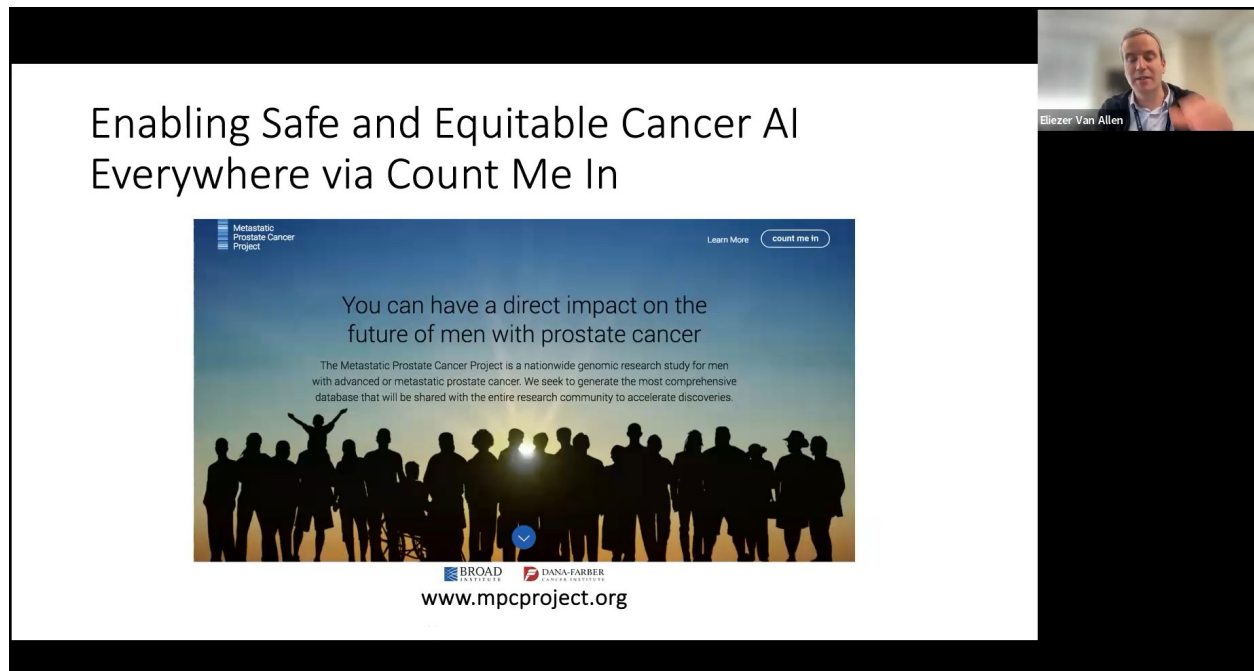


But it's again, just scratching the surface, because there are lots of other data modalities that one could insert into these kinds of models like I hinted at earlier.



“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

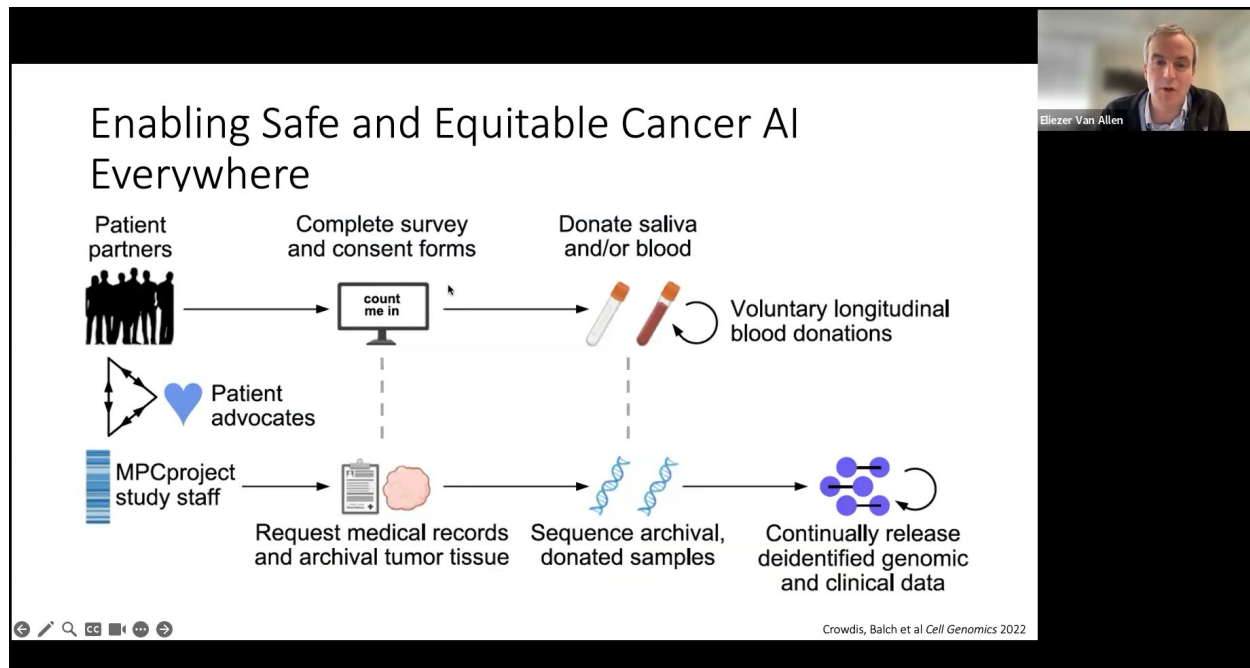
There are lots of opportunities for making new clinical trials that can actually represent this stuff more, which is very exciting.



The image shows a presentation slide. On the right side, there is a small video inset of a man, Eli Van Allen, speaking. The main part of the slide is a screenshot of the website for the Metastatic Prostate Cancer Project. The website header includes the text "Metastatic Prostate Cancer Project" and a "count me in" button. The main text on the website reads: "You can have a direct impact on the future of men with prostate cancer". Below this, it states: "The Metastatic Prostate Cancer Project is a nationwide genomic research study for men with advanced or metastatic prostate cancer. We seek to generate the most comprehensive database that will be shared with the entire research community to accelerate discoveries." The website also features logos for BROAD INSTITUTE and DANA-FARBER CANCER CENTER, and the URL www.mpcproject.org. The background of the website screenshot shows silhouettes of a diverse group of people.

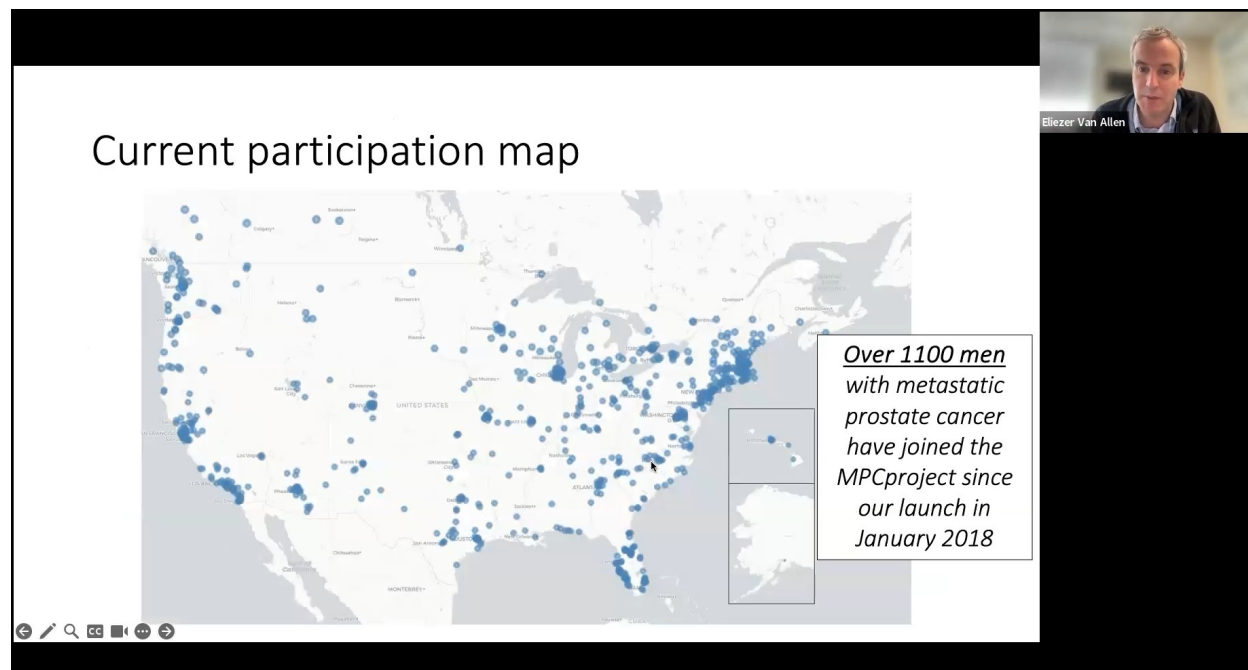
Snowballing into where I'm trying to end is really about safe and equitable cancer AI everywhere. To do that, you need to not only have a lot of the downstream technology capabilities accessible to everyone in a way that is representative of the patient's experiences in the hospital systems and everything on one end. But **upstream at the earliest possible point, you need to have training data that represents humanity that isn't biased by historical norms in our field, whereby we expect patients to come to us at one of the few select cancer centers to participate in this kind of research in the first place, but rather flip the whole paradigm on its head and invite any patients from anywhere across the country to participate in research.** I've had the privilege of being a part of Count Me In through the Metastatic Prostate Cancer Project going all the way back to 2018, including some of the core initial advocates like Bryce (Olson), and others to stand up a study called the Metastatic Prostate Cancer Project or MPCproject.org. Here's the front page of the website.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



Here is the way it works. Patients click the Count Me In button. They're asked to complete a survey and a consent form that gives us permission to access their information. It also gives us permission to ship patients saliva kits and blood biopsy kits. The saliva kit helps us to understand the genetics that they were born with. The blood biopsy kits help us to understand the most current status of their tumor genetics, as detected by the tumor genetics that has shed into the blood. We can have folks send multiple of those kinds of kits, so that we can get longitudinal information about their disease course. We generate this data. We make it available as quickly as possible. As was alluded to, this is a research project. We are, from a regulatory perspective, unable to return individualized results. We can come back to that in the conversational piece why that is and what we're trying to do about it. But for now it is a research project that still ideally helps to feed this broader goal of helping us discover new biology that might help everyone everywhere.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

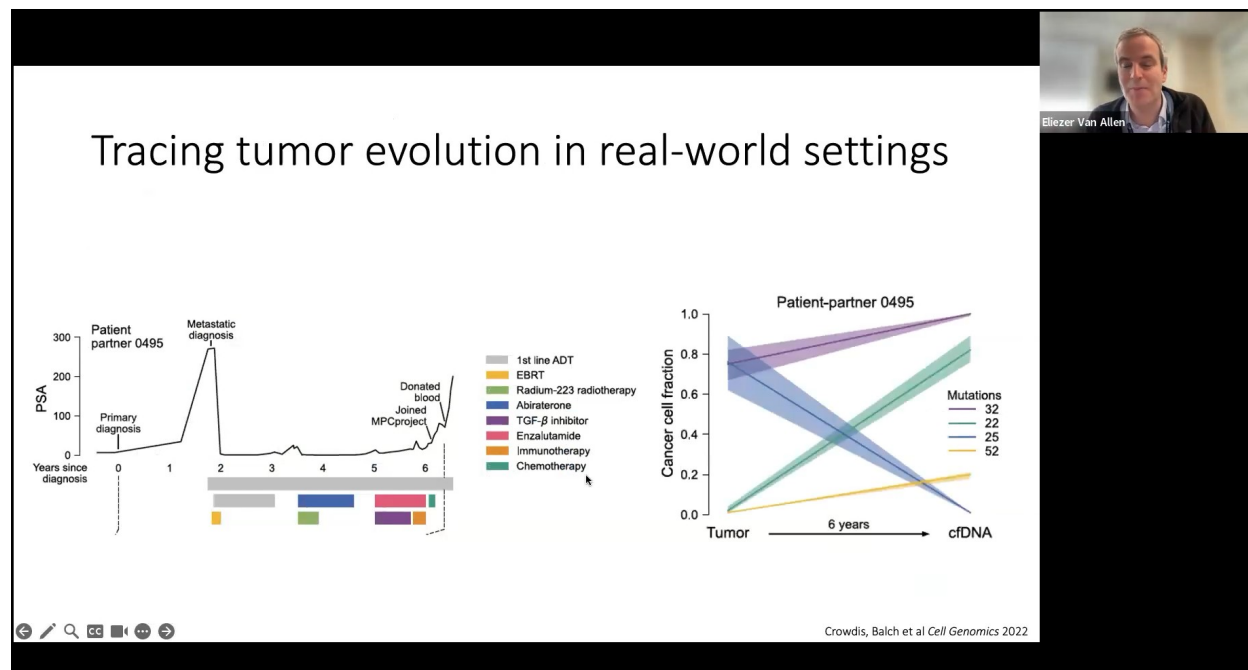


Here's our current map of participants. Roughly currently, we've had over 1100 men with representation across the United States and Canada. I think we have all 50 states at this point. I should double check that. Don't quote me on that.

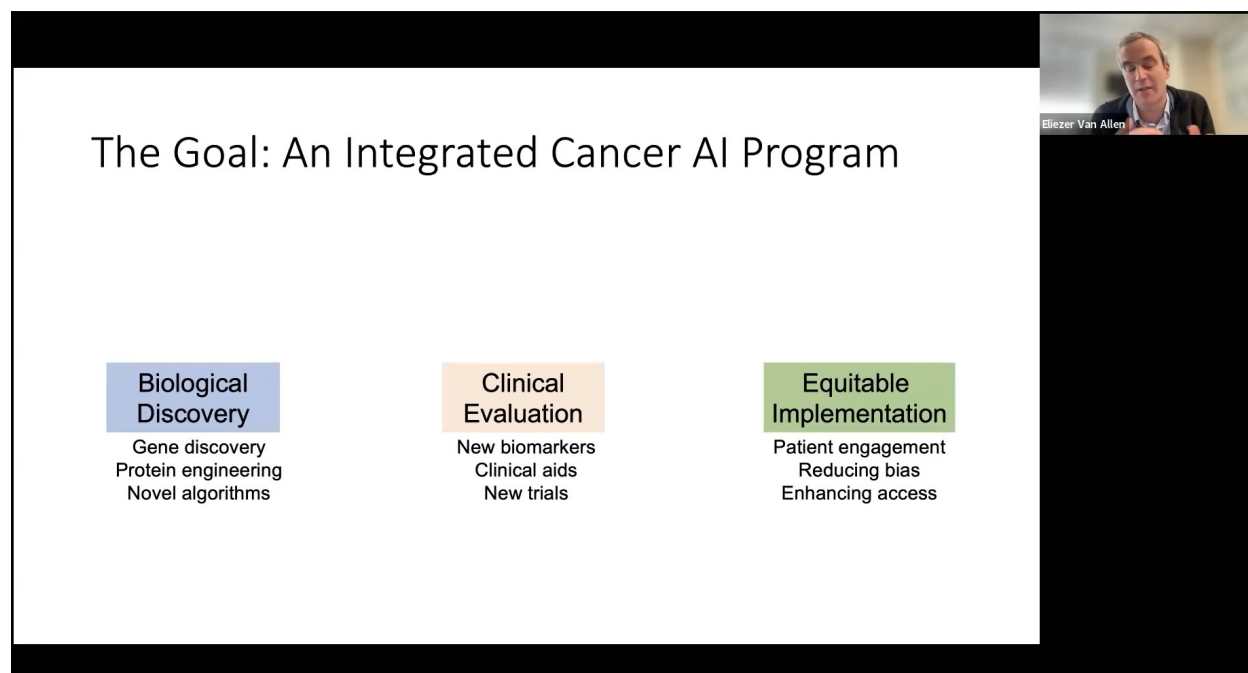


We can do some really interesting things. Like not only understanding real world data at one time point, but really study patients' experiences with different therapeutics,

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]



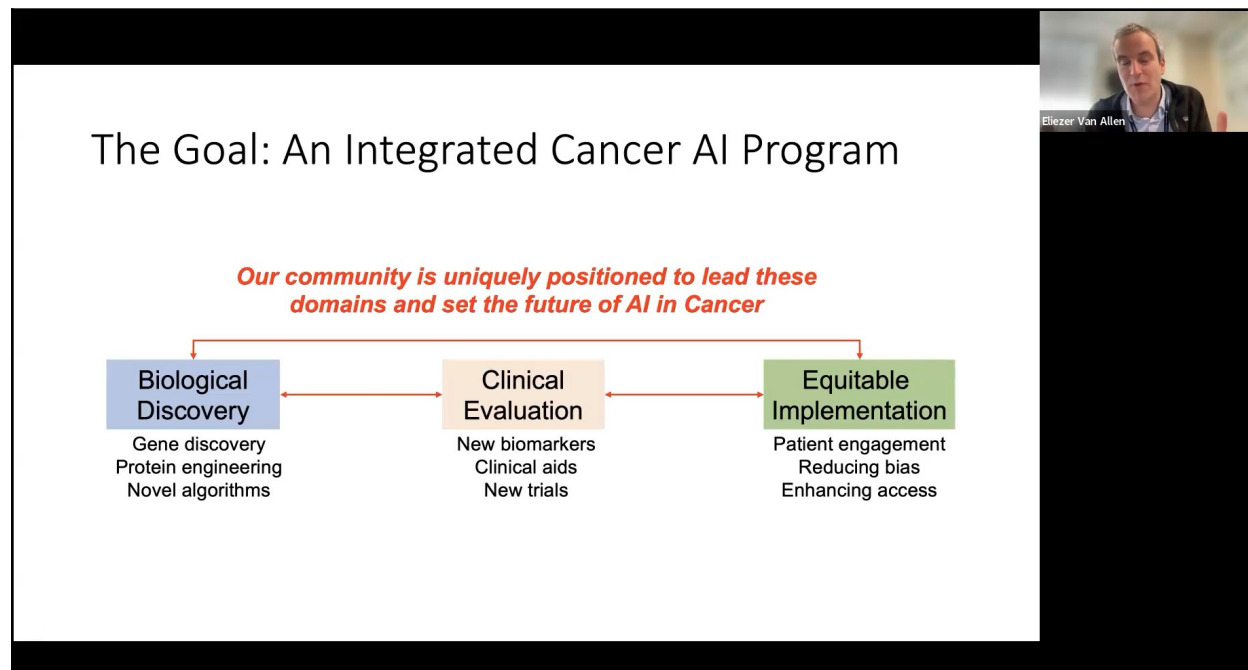
and match their tumor evolution and do some really cool inferences downstream. This is really intended to be a broad overview, a quick snapshot.



The goal of what we're trying to do is to (1) do some biological discovery with this. I hinted at one application, but there are just so many. It's such an exciting time, because there's just so much cool stuff happening in this domain.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

(2) Make better ways of integrating this into clinical care. And then (3) doing it in a way that everyone can access it, which is I think, really where Count Me In and projects like it are at a unique space to really impact the field writ large.



And obviously connect the dots between everything together and are the “royal we”, including all of you, people near and far, if we actually want to make this happen, we as a community are uniquely power to do so. And set the future of AI data sciences and everything else in oncology across the board. Many things to be continued.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Acknowledgements
The Patients

Let's work together!
vanallenlab.dana-farber.org
@vanallenlab

Van Allen Lab

Haitham Elmarakaby	Breanna Titchen
Kevin Bi	Nicita Mehta
Alma Imamovic	Samantha Hoffman
Brendan Reardon	Shreya Johri
Jihye Park	Kevin Meli
Saud AlDubayan	Camille Mathey-Andrews
Sabrina Camp	Theodora Pappa
Alok Tewari	Lauren Miller-Jones
Riaz Gillani	Jingxin Fu
Matias Vergara	Erica Pimenta
Ryan Collins	Maha Shady
Yuxi Liu	Amanda Garza
Gwen Miller	Chris Labaki
Claire Johnson	Eddy Said
Andrew Zhou	

+ Alumni!

Broad Institute

Entire Count Me In team
Gad Getz
Sachet Shukla
John Doench
Genomics Platform
Picard Team
Firehose Team
Carrie Cibulskis
Many others...

DFCI/BWH + Center for Cancer Genomics

Toni Choueiri
Mary-Ellen Taplin
Chris Sweeney
Entire GU Oncology team
Entire CCG team
William Hahn
Rand Arafeh
Felix Dietlein
Many others...

Friends Near & Far

SUZC PCF Dream Team
Justin Hwang
Many others...

Funding

The Mark Foundation for Cancer Research
DAMON RUNYON CANCER RESEARCH FOUNDATION
NATIONAL CANCER INSTITUTE
Prostate Cancer Foundation
Curing Together.
CANCER MOONSHOT
IOTN Immuno-Oncology Translational Network
HTAN HUMAN TUMOR ATLAS NETWORK
PARKER INSTITUTE FOR CANCER IMMUNOTHERAPY

I thought I'd end by again thanking the folks in my lab. I get to work with the entire company in a team, some of whom were on this call. It's really a privilege, again, to work with this team. Day in and day out, folks near and far. Various funding sources. There is our lab website and our social media handle.

Rick Stanton 25:23

For your initial cartoon of AI, you used the genes that were depicted in terms of mutations. I've been a big fan of immunotherapy since they came out and trying to predict response from RNA seq from the tumor microenvironment. I imagine you use that as well.

Eli Van Allen 26:00

It's a great point. The model that I showed you, in some sense, was a toy model. It was a model that snowballed into a Nature paper, which, on the academic side, that's great. It was just like version 0.1. Us and many others are experimenting with different data, modalities, inputs, transcriptional information, not just bulk RNA sequencing data, for which it is hard to actually attribute exactly what cell type the signal is coming from. Is it the tumor cells? Is it immune cells? Is it other cells?

Single cell data, which I was alluding to in one of these pictures, is where you can really do all that. So that's actually just the start. Yes, these models are flexible.

The biggest problem with these models, the biggest philosophical limitation, is that it's built off of prior knowledge. In the end, you've been your “glass half full, glass half empty” person. We only know something. Even that's incomplete about, let's say, depending on if you believe, 3000 to 5000 genes of the 20,000 genes in the genome. These models are incapable of going into the complete unknown, that are not represented in the prior knowledge domain. That's where the

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

foundation models might actually be particularly interesting to explore. If we didn't have these technologies, it would take us like a million years to sort out the old fashioned way. So honestly, I could be wrong. But that's a great point.

Rick Stanton 27:34

My buddy got a liquid biopsy result. I think it was from Guardant. The reports had a 78% variant allele fraction. I thought that seemed really high to me. How could that be? I tried to grapple with it. I really don't know about shed DNA. I would have to think that most of the shed DNA is coming from a tumor, rather than a normal cell to get that kind of variant allele fraction, at a high rate. Do you have any interpretation?

Eli Van Allen 28:24

It's a tricky one. Here's one thought. VAF is the variable allele fraction. If one is considering what our theoretical tumor source reads, it's the abundance of the mutation, let's say in that read. Normally, you'd expect, if it was a typical mutation, where you just needed one mutation to cause a problem, you have two copies of every gene in your genome, you'd expect the VAF to be max 50 percent with some error mode, right? So how do you get above that number? Sometimes it's if there are certain kinds of genes that instead of the mutation turning it on, if the mutation is needed to turn it off, it's called tumor suppressors. There you need to lose two copies. Sometimes what happens is that one copy of the gene is lost by a big deletion, a chromosomal deletion, let's say. And then you have a mutation that keeps the other copy. In that way, with some error, because it's also picking up normal cells and others it's like a big gets a little bit muddy, you end up with something above 50%, somewhere in that range. So that's at least one explanation. It would be hard. I'd have to see the whole report to know there's a few other niche ones as well, but I wouldn't be surprised if for instance, this was a so-called tumor suppressor, that you're actually looking for. Two hits.

Rick Stanton 29:56

Okay, but even approaching 50? Wouldn't you get most of your shed DNA from normal cells from the blood?

Eli Van Allen 30:09

This comes down to each company doing their own thing: how they're filtering, how they're zooming in, what they're looking at, maybe some extreme scenarios. For example, sometimes men who present with de novo metastatic prostate cancer and have a PSA of like 10,000 have a pretty extensive shed tumor burden in their blood. I'm guessing they do some computational work to address some of that, without knowing the details.

Rick Stanton 30:47

It just was bugging me.

[Via the chat:

00:25:36 *Eric Hall: My son went to Camp Kesem this past summer!!!*

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

00:26:52 *Dr. Mike Donohoo: My son's two children (he died Nov 2019 of pancreatic cancer) attended Camp Kesem and loved it!! Thank you for doing this.]*

Eli Van Allen 30:48

It always warms my heart to see folks participating in Camp Kesem. I must say, having been there from the beginning, what is never-endingly wild to me, is that this thing snowballed into what it became or has become, knowing where it began. It's wonderful to hear, and I hope it has been a good experience for those who participated.

David Plunkett 31:26

When you talked about training data, it sounded an awful lot like what I was encountering through the [PROMISE Study](#) (a study to test healthy people who may be at risk for early warning signs of a blood cancer called multiple myeloma.). Do you have access to the data that they've accumulated?

Eli Van Allen 31:41

We're trying to work to get access to those data. Depending on the project, there can be a variety of regulatory reasons that make it difficult to access things in real time. That's also part of the historical norm of the way research has been done. I would give a shout out to the Count Me In folks who had a “day one” ethos to make data available as quickly as we can, to keep up with generating and analyzing it, independent of any one person's research agenda or one organization's research agenda. It still has challenges just to keep up with it, but we've tried to do our best. Some of the fun of those kinds of things has been seeing people use the data, cite the paper, and we have no idea who they are, or what they're doing with it. That's the whole point of what we're doing. We're working with intermediaries, like the PROMISE study and others like it who are doing these kinds of things.

There's a bigger data sharing problem. We all want to share data, and we'll want to access this to make it useful for the “royal we”. It's actually very hard from a regulatory perspective to do so. This is a shameless pat myself on the back. Last fall, I was part of a panel for the United States Senate on educating the senators about AI. This panel was “AI and cancer”. At the end, they asked us all, “If you could fix one thing, what would you fix?” On this panel I was the academic. There was somebody from Google, somebody from Tempus, and somebody from some other independent research lab. Ned Sharpless was moderating. Ned is the former NCI director, and everyone basically said, “If you guys can help simplify all the rules for data access, and patients are demanding this, but we are afraid that we will get in serious trouble with your current rules.” There were some smiles and nods from the senators. They're busy with other things right now, but maybe one day they'll fix that and make it easier to share data in the first place. That's a semi-related tangent and a rant that I always like to have. Certainly, if you ever interact with any of your representatives, let them know that you'd like your data to be shared, if that's something you care about.

Jeff Krolick 34:20

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

I have two somewhat related questions. From a patient perspective, it sounds like you have very granular data, or you can get down into very fine details. At what point might this be able to be individualized? Here's my data set, you run it, what are the possible treatment options for me, rather than looking at the aggregate?

Second one is, many of us also weigh with any treatment options: what are we willing to do to stay alive versus the side effects and quality of life? It's not just overall survival, but it's the quality of that. Is there a way to capture that as you're going through all of this, to see which treatments are somewhat effective, and which are very effective and provide maybe the best quality of life and the fewest side effects?

Eli Van Allen 35:52

Those are two great questions.

On your first question about interpretation. The sequencing vendors try to provide their interpretation. The problem with those things is that they are static in time. Frankly, there's literature on this. It's hard to make them representative. They don't know you, and they don't know where you live, or which trials you have access to. I've seen a bunch of these from patients where a patient lives in Boston, and it recommends a study in Seattle. That doesn't make any sense. It's simple things like that. Or they're not up-to-date. It's a really hard problem to do well and correctly. Shamelessly, we've been thinking about this from a computational perspective. There's an example from one study that has a method that includes a portal where you could conceivably upload your data, which is still not super easy to use, and run this method. It'll try to give you some clues as to what's actionable in your tumor. The trick is, though, this is like your “research use only” tool. It comes back to the prior conversation about what is research and what is clinical care, and who gets to do what. There are a lot of uncertainties there. We're experimenting with some ways to try to overcome that angle. It's a complicated space still, both from a technology perspective, but also from an implementation perspective.

On your second point around other information: that's a really interesting one. In our metastatic prostate cancer project study, we had built in these patient surveys that included a lot of questions around supportive care, other kinds of things you're taking, and other stuff. That's in part driven by patient feedback, as we were building this project, saying, they read the notes that the doctors write, and it doesn't even include a lot of the stuff they're telling in the first place about these things. If it's not included there, it's never going to get measured. If it's never going to get measured, it'll never be able to be included in some of the things that we want to ask questions about down the road, around which ones have a lot of these issues, connected in some other way, with everything else which is a problem. We're certainly encouraging more and more folks to speak up and speak out and contribute that kind of information when available, with the hope being that it might help guide not only decision-making from a purely biology perspective, but also holistically as part of the whole patient's experience. Being self critical as a provider, as a medical doctor, as a physician, we still struggle with balancing a lot of those things. So certainly a lot of work to do there as well.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

Brad Power 39:04

A lot of us are interested in how you can use the data and the information, the AI, the algorithms, and the insights for individuals, which Jeff just touched on a moment ago. Frank and AI, who are teed up behind me, will also touch on this. I understand why you haven't been able to do it today. I think in the past, you told us it would be very hard to return results and very expensive. The vision is we want to find “patients like me” – has anybody who has my profile done similar things, and what were the outcomes? That might steer me to this treatment versus that treatment. Are there any efforts underway given – as you have pointed out – all of the energy going into AI with customized ChatGPTs and so on? What are the efforts underway that we could lend our energy to, that are doing the things that we want as patients, where the beneficiary is the patient, or the doctor deciding for an individual versus for drug discovery or for broader insights into population-level information?

Eli Van Allen 40:19

How do I use it to help me? I'm sympathetic to that being a major need. Within the Count Me In organization, we have one project that's part of a National Cancer Institute consortium, where the return of results is mandated as being used as the sandbox to figure out all the nuts and bolts of all the problems. **There are many challenges; one of which would be, for instance, if patients are driving it, then let's say in prostate cancer, inserting me, Eli, who lives in Boston, into the clinical care of thousands of patients around the country, which is a very complicated thing, and stressful for whoever takes that on. But also on the other end for the providers, the teams, it creates a lot of uncertainty, a lot of things you need to keep track of.** It's a bit challenging to do it, especially if the goal is to come up with real time treatment decision-making, which is an even more complicated element.

But we're trying, and similarly locally, at Dana Farber we're trying. We have some projects that are in real time in my lab, and within collaborations where we're taking our local data, like unstructured electronic health record data from many thousands of patients, and their genomics, and building these patients similarity algorithms, so that you can find how similar you are to other people, see what happened to other people and what was done, and then use that knowledge to guide care. That is a real time, precision medicine goal that we're building now.

The tricky parts become when you want to do that everywhere. There are a lot of challenges to moving these algorithms and testing whether they generalize. A lot of biases related at sites are specific to them. There are even aspects related to which trials are open where, and whether they're relevant for X, Y, and Z. It's still a very hard problem. You as the patient can have access to a lot of the same portals and datasets that we do. There's a great one, for instance, [the cBioPortal for Cancer Genomics](#), which Count Me deposits their data into, that is navigable and does not necessarily require being a super wonk to play. It's got a GUI (graphical user interface) that you can play around with. It's a place to start getting into the weeds of the more sophisticated things.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

In my opinion, I have not seen anything that's so ready for primetime, that everyone should be trying or using it, whether commercial or in an academic setting. But it's what we're all heading towards. But instead, it creates a lot of the same questions and **problems around access, equity, and bias, and whether these models can generalize and be translatable**, such that even if we built that perfect patient similarity thing, and we're able to match you to others and do this for people here, would that translate to somebody in the Midwest, or on the West Coast? Would that even translate to Boston Medical Center, which is three miles away from me, but is a very different medical environment? I don't think we know the answers to those questions. We already have seen in the AI world, in particular, enough examples of where it's already gone bad to give us pause, or at least, I hope, give everyone pause to think about how we're going to do this carefully. That's my two cents.

Frank Nothaft 44:06

When you look at some of the research that you're doing, like a lot of the work in AI for drug discovery, it's very important, but it's very long range, very high risk, high reward. What are some of the things that you see in the research that could come out of a data asset like this that could have short-term patient gains? Because it seems like this is going to be a really valuable asset for forward and reverse translational work. Some of the things that you're showing, like the work around the TIL (tumor-infiltrating lymphocytes, growing immune system cells) prediction from your slides, feels like something that might be short-term diagnostically actionable. I'm curious what you see there.

Eli Van Allen 45:16

You're absolutely right. The biology discovery side of things always has a much longer arc to impact. All that work leads to maybe a target, which is step one of making an actual drug, step one of a 1 billion step process. It's very hard. Obviously people would love to shrink the timelines. But even in the best case scenario, it still takes many years. That's why I wanted to highlight some of the other examples that are much closer to being “shovel ready”. Whether it's the digital pathology space, for which in addition to our own research, there are commercial vendors already setting up shop doing, not the stuff we're doing, but rather stuff for more diagnostic purposes for pathologists. Then, likewise, the precision oncology patient stratification matchmaking kinds of things, for which we're not the only ones thinking about this. There's lots of energy happening there. You could imagine those things becoming quicker to hit the clinic, in hopefully controlled testable environments, but it'll hit the clinic in most cases well before you find the drug.

Now, as a counterpoint to myself, in that example, we used all that retrospective data and **found the drug target that already has drugs that are already in clinical trials**. It's just that they weren't in clinical trials in prostate cancer. That's a huge shortcut. To try to find something that was missed by our own accord and then flip it around becomes just as quickly intervenable as some of this other stuff. But, how many of those examples are there? I don't know, probably not a ton. It's all fair.

Frank Nothaft 47:08

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

You definitely see good examples of line extensions, like Keytruda is now available essentially more or less across the board.

On the regulatory side, you mentioned that you see some pretty substantial regulatory issues blocking data sharing. I was curious to hear about that. I'd love to hear what you see specifically because if you collect the data in clinical practice, HIPAA itself provides pretty involved but straightforward ways to make the data shareable for research reuse. Is this a problem or with the funding model? For instance, I know the NIH and inside of HHS, the NIH and their human subjects policies are different from what plays in HIPAA land. What do you see are the specific barriers?

Eli Van Allen 48:14

HIPAA, just like any other thing, has a layer of legalistic interpretation that is very dependent on who you're talking to, where, when, and why. Likewise, if you're conducting research, even if it is eventually informing clinical care, if it is research, then there's a much bigger layer. Then you add on top of that, where the law starts to get a little bit fuzzier, about what's colliding with what. What's identifiable information? How protected are your genetics, for instance? Where does that get managed and controlled? You see it in the way our consent forms are created for patients, and they keep getting longer and longer. They never get shorter. That's a reflection of the inherent problem.

In addition to that, there are people of a conservative nature. People are always anxious to pull back protections, because people, understandably, get upset about these things. There's a lot of risk there. It's always harder to pull back things that are already happening. Then people are reactive. When one bad thing happens, then people add more things on top, that oftentimes have lots of other problems, that just make everything harder and harder and harder.

The counter where patients have a unique opportunity, and I've been on the soapbox for years, and it's gotten me nowhere so far, you as patients have via HIPAA right of access, the right to follow up with every hospital you've ever been to at every commercial sequencing site you've ever had any data generated and whatever, and demand your data. Not just some stupid little PDF with a little portal. Everything. We did an experiment with some patients, including Bryce (Olson), about five or six years ago, where we had patients call up genomic testing companies to invoke HIPAA right of access and say, “I don't want just some PDF report, I want the raw data.” We learned that we caught some companies off guard. They had no idea what we were talking about. Some companies wanted to help, but they didn't know how. In one case, it may have even been Bryce's, the company we won't name FedExed a thumbdrive of his genomic data to us, with no context, not helpful. Some companies intentionally prevented access, effectively, knowingly violated HIPAA, because – what are you going to do? Then we wrote this thing up, and then a Forbes reporter found us, called us, and put a title of an article and called us “data pirates”. Then my mom called me up and was like, “What the hell is this?” It was a fun experiment to try. But it proved the feasibility. Because let's say tomorrow, a million cancer patients at the same time called up their hospitals and said, “Hey, man, give me my data. I want to give it to these crazy scientists over in Boston, who were going to do some AI on it.” If they all

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

did that, you'd break the system, certainly. But then all of a sudden, all these things go away. Because in the end, it's very clear what the patients can do. Once you have the data, you can do whatever you want with it. You can staple it to your forehead, you can send it to us, you can put it on the moon. That's Eli's provocative call to arms.

I've never been able to successfully get any research funding to support going from that initial pilot phase of like, five of my friends, to let's actually scale this out to see if there's a there there. It scares some people, because there's some profit motives that get destroyed based on that kind of principle. But I haven't given up entirely on it. I just haven't figured out the right way to do it. I'll get off my soapbox, but it's a great question. It's a tough one.

Frank Nothaft 52:26

If you don't mind, I'll follow up with you offline on that last point.

Al Musella 52:49

First, we have a patient program for brain cancer right now that kind of does what you said was tough to do. The way we get around some of those problems is we just give a list of recommendations to the patient and their doctor, and they make the clinical decision. So we're not really setting up a clinical relationship.

Second, we can automate the gathering of all that information. You talked about what we need from the patients, that consent. For our back end, which is run by xCures, they go out and gather all those medical records you talked about and the system uses artificial intelligence to structure the data.

My point I was going to get to is: we're talking about equity. The trouble that we have is we give these patients these ideas of what would be the best clinical trials, for example. And none of them could get into the clinical trials because of location or whatever. So we're working on a way to get these treatments to them outside of clinical trials. The obvious answer is expanded access. But that's not a large-scale solution. That's one-offs. We're working on a way, and I think you might be interested in this, we could talk about later, called **the Promising Pathway Act**. It's a new law that would give us access to all these experimental therapies. But it changes the whole system. It's basically a distributed clinical trial where everybody could get access to these treatments, but they're within a clinical trial using them until we get enough proof to get it approved, then it graduates to have full approval. It's a conditional approval basically.

Eli Van Allen 54:25

That sounds fantastic. I'm reminded of a patient I have with metastatic bladder cancer, who I had to – and I remember this because it took forever – basically beg borrow and steal via a bunch of forms and a bunch of pleading and a bunch of everything with Merck, to get early access to pembrolizumab (Keytruda) for this patient back when we already knew there was signal. It was exactly the type of situation you were talking about. I was able to get it after a ton of work. This was like in 2015, and here it's 2024, and he's alive and well, effectively cured of a metastatic solid tumor, which is remarkable, which is great, except, as you said, it's a huge pain

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

in the industry to be in really challenging and impossible to scale. What you're describing there sounds like a nice story. And if we're not tracking it, we're not learning. By forcing it to be structured somehow, we can all learn collectively

Brian McCloskey 55:39

Our patients at the Cancer Patient Lab have access to a number of different service providers that can sequence their information and interpret it, and then provide treatment options. I, for one, have used several of them to come up with 21 different treatment options. And then, of course, worked with Rana McKay, who I know very well. She spent the time reviewing all of them to figure out what was best for me. But what was interesting is there was no concordance across three different vendors. So you have an interpretation challenge.

I'm just curious if there are sets of data that that we really should be focusing on that can be predictive, or can they really identify targets that will make a meaningful difference in their treatment decisions? For example, whole exome sequencing versus proteomics versus transcriptomics versus spatial phenotyping. There's just a myriad of these different tests.

Eli Van Allen 56:45

Great question. I think at this point, if you wanted to be really cynical about it, even most of the genomics is still not ready for primetime. And even for the things we find in patients that we have known about for decades, we still don't have drugs, like for P53, which is mentioned in the chat. This is not a new cancer target. It's just an old one that everyone knows about. I have no idea how to make drugs. So we move on.

If that's the baseline, a lot of the other tools, which I'm very familiar with in my research lab, fancy schmancy, spatial, and transcriptomics are very cool. We ask interesting questions. There is lots of cool biology. It's very much cart before the horse. Like, how is that going to tell me which drug to pick with confidence? That's tough, and so I would say, at least today, you can speculate and hypothesize. And I certainly do. We all do. But trying to stack or rank the tech by what's going to immediately help me today, that's a tough one. But a lot of these things are cool and really exciting technologies. The newer they are, the harder it is to be clear how useful it is. Unless it's tailored to an existing therapeutic or companion to a therapeutic that's already in development, which most of these are not.

So that's my possibly mildly controversial hot take on that.

Brian McCloskey 58:21

The diagnostic space is moving at lightspeed on the pairing or the matching to drugs that are going to be effective. The data is just not there. Is that fair to say?

Eli Van Allen 58:35

For me, it's mentally jarring flipping between my clinical and research brains. In my research brain, I'm looking at spatial transcriptomics and single cell data of like, millions, billions, or trillions of data points on all these high dimensional things, and we're doing all this crazy stuff.

“Translating Patient Data into Clinical Use” (Eli Van Allen) [#81]

My clinical brain is like, “Alright, I’m doing Lupron and docetaxel, or Lupron or Abi (abiraterone).” It’s like, this or that, and what do I have to guide us on those things? It’s a completely different scope. So the goal, where the data comes in, is you build the bridge, and you also build new bridges. That’s a long term goal.

Brian McCloskey 59:25

This was a phenomenal conversation. We would love to have you back because I know that we just scratched the surface. Maybe we’ll find a way to hone in on some topics and can continue this conversation.