Machine Learning Meets the PICU-Predicting ECMO Outcomes in Pediatric Chronic Lung Disease - Pediatric Grand Rounds-5-16-2025-Meeting Recording

May 16, 2025, 12:29PM

52m 24s

Kamat, Deepak M started transcription

Kamat, Deepak M 0:26

Good morning.
It's 730 in San Antonio and it's time to start our grand rounds before the grand rounds. I have a quick announcement.
To remind you all that next week, next Friday, May 23rd is the Pediatric Research Forum, which starts at 12:30 PM in room 409 L.
Lunch will be available and our senior speaker this this year is our own doctor Morerera and he will be giving his screen presentation at 1:00 PM.
So please, please, please remind your medical students, residents and fellows who are working with you on research project to submit abstracts as soon as possible.
Please contact Doctor Mustafa if you have Shamim Mustafa. If you have any questions about the pediatric Research forum.
So again, we are continuing with our graduating fellows, presenting grand rounds and today it will be doctor Shyan Mangold from pediatric critical Care Medicine and Doctor Veronica Armijo will be will be introducing Dr. Mangold.

Armijo-Garcia, Veronica 1:34

Good morning, everyone.
So I am very excited to be able to introduce Doctor Mein Gold, who most of you I'm sure have had the opportunity to learn or to to get to know over the tenure that she's been here with us.
Doctor Mangold grew up in San Antonio.
And in her young adult years.
Went off to Texas Tech Health Science Center.
And her to get her both her.
Sorry, she she completed her mph and her MD at Texas Tech.

Then she came back to San Antonio, thankfully for residency.
So she was here for residency and then progressed onto pediatric critical care, which she's done a fantastic job as a fellow progressing through our program and we're very excited that she'll be joining.
Us on faculty.
Starting in June, so her interests include ECMO, which is nearly dear to my heart. You guys know.
And iatrogenic withdrawal syndrome.

Wards team 9th floor 2:47
I can tell my dad something and then.

Armijo-Garcia, Veronica 2:51
As well as machine learning, which is far from my expertise, but she's done a fantastic job.
With her research project, which is machine learning needs to pick U predicting ECMO outcomes in pediatric chronic lung disease.
So I'm going to give the floor to Doctor Mangold so she can.
Teach us about her project.

Mangold, Cheyenne M 3:19
Awesome. Thank you so much, Doctor Armijo. Good morning, everyone.
Thank you so much for being here today. So today I'm going to be presenting on my fellowship research project, which is machine learning where it intersects the PICU and specifically combining some NIU and PICU problems.
So we're going to be trying to predict ECMO outcomes in kids who have some chronic lung disease including BPD pulmonary hypertension.
To start off, I have no financial disclosures, conflicts of interests.
To disclose and I'd like to start off by acknowledging my mentors.
So in addition to my pick U mentors, which include Doctor Armijo, Dr. Tyler Hughes, Doctor Meyer, additionally, I also would like to acknowledge my NICU faculty, who was a major contributor to this project and has also been the teacher who's been teaching me how to do machine Lear.
Who's doctor? Amorea who?
Many of you know and will be the keynote speaker at our next next week's research

week.
So let's set the stage.
So I want to highlight the specific patient population that motivated this work.
So the majority of the patients are infants with bronchopulmonary dysplasia that we'll see referred to as bpd throughout the rest of the presentation. As many of you know, BPD is the most common chronic lung disease. We see an infants, especially those who are born preterm. So you'll see.
On the left hand side of the screen there's about 130 to 140 million.
Pregnancies worldwide?
Per year of those pregnancies, the majority, about 120, are full term, but about 13.4 million are preterm.
Of those, there's those that are under 1000g at birth, about 70% of those are going to go on to develop some form of bronchopulmonary dysplasia and then those that are born above 1000 grams, about 30%, will go on to develop BPD.

Pat Herndon 5:03
Yes.

Mangold, Cheyenne M 5:19
Why this is important is this is where the PICU comes into play. As you know, infants who have who are born extremely premature are going to have a prolonged mechanical ventilation and prolonged oxygen therapy, and some of them even are discharged home with either tracheostomies or home 0.
But we know that the respiratory morbidity doesn't end at NICU discharge.
We know that infants with BPD do have higher rates of admission.
After, especially in the first year of life.
With recurrent respiratory infections, poor growth or ongoing oxygen or ventilatory adjustments that need to be made, these complications also will increase their likelihood of requiring ICU level care. And sometimes this can be multiple admissions throughout the first years of life.
Now we know ECMO is well known to the PICCU, but can be uncommon to those who work outside of the ICU. According to our neonatal or national ICU database.
And else.
And what we know from our LSO data.
What we understand currently is that about four percent of pediatric patients, those

under the age of 21, ultimately will.
Require ECMO during their ICU stay and we know that patients who have severe bpd, and especially those that have also developed secondary pulmonary hypertension, there's already an increased risk due to right heart strain, impaired gas exchange and their limited pulmonary reserve.
So we're talking about a subset of already very complex patients.
Who then?
Go on to need one of the most resource intensive forms of support life support that we offer.
Let talk a little bit about ECMO and shift to ECMO itself and ECMO.
Kind of came to the forefront during the COVID pandemic, but those of us in the neonatal and pediatric world knew it far long before the COVID pandemic, because it really had revolutionized our ability to support children with severe respiratory or cardiac failure.
It's life saving, but it does come in a substantial cost and it does carry a high risk and so you'll see on the left hand side of the screen I have two photos, essentially one from a world view and then one closer up of a patient of just.
How intensive the resources are to run ECMO so you'll see in this first picture that very small.
You almost can't.
You can't even see the patient in this first picture.
But you can see the large number of machines, so to the right hand of the neonatal bed, you'll see.
Ecmo circuit itself, which contains multiple components. And then you'll also see oftentimes the large number of different drips and medications these patients require.
And then in the picture below, you'll see a more zoomed up picture of essentially the positioning and the type of support that these children require. And also you think of how small neonates are, and look at the size of that tubing, the amount of.
Flow and blood that these kids require.
So we know in the last five years alone, there's been over 110,000 ECMO runs that have been reported worldwide in neonate survival to discharge via the the LSO data is about 59% and it's about 60% in Pediatrics. And those are really.
Promising numbers, but that does mean that about one in three to two and five patients still don't survive even with this maximal support.

But we also know that ECMO is one of the most expensive interventions in critical care.
Where the cost?
Range from 5000 to $11,000 per hospital.
And most runs a typical ECMO course run you'll see at the bottom is about five to seven days.
But we also know that some kids can take much longer if it's a bridge to transplant. In some cases, kids can be on ECMO for months.
So this can translate to 10s of thousands of dollars even before you factor in. Once they do come off ECMO, the long term rehabilitation treating the complications or even subsequent hospitalizations these kids face.
And we also know beyond the cost of the support itself that these kids are at increased risk of neurological injury from strokes, brain bleeds, hemorrhage due to the anticoagulation and risk factors that come with ECMO itself, infection and other multi organ dysfunction.
And so.
So we when we have to initiate ECMO, especially in kids who are chronic lung disease, we have to, we're making this decision that carries quite enormous clinical and ethical weight. But we often don't have great tools to estimate how these kids are going to do in the long.
Term and to give families good ideas about likelihood of survival.
So that's where our project comes in.
We want to explore whether we can use machine learning to assist in our early risk stratification.
To preface, not trying to replace clinical judgment, but give a secondary source of information.
To both clinicians and families.
Though our objective was twofold, we wanted to identify risk factors for in hospital death in children with BPD and those who had pulmonary hypertension requiring ECMO and then to develop a machine learning model to try and help predict which kids may be at the highest risk for death.
During their hospital stay.
So let's talk about where we got our data from.
So our data came from the extra corporeal life support organization or also for short, which many of you may or may not be familiar with, but for those that aren't as

familiar with alzo, it's an international consortium of ECMO centers that has been collecting detailed case level data.
Since the 1980s on the right hand side of the screen, you'll see just how many EC mod centers exist throughout the world and contribute to this database.
As indicated by all of the dropped pins on that map.
It's the largest ECMO registry in the world.
And it serves not only as a research resource, but also as a quality improvement and benchmarking tool for all of the participating centers at the time of our analysis.
The registry included over 112,000 ECMO runs in the preceding five years alone from both neonatal pediatric and additionally, that didn't include the 100,000 adult cases as well.
What makes LSO particularly valuable from modeling is its breadth and depth of data, and includes patient demo.
Graphics primary diagnosis, comorbidities, ventilator settings, ECMO, cannulation, details, flows, gas exchange parameters, complications and outcomes, including survival to discharge.
Importantly, it also captures ECMO mode transitions. Complications such as hemorrhaging neurologic injuries and standardized timestamps.
So it takes all of your info at the beginning of the ECMO, right before a patient goes on to ECMO, and then repeats a subset of data at the 24 hour mark after the initiation.
Of ecmo. So that does allow gives us a good clinically relevant window for prediction.
And while registry based studies do have limitations, including variability on your data entry and missing data points.
Also scope and international representation does provide a real world data set that reflects current practice.
And it does make for a strong foundation for building generalizable machine learning models.
So I'm going to walk everyone through how we started to filter down this data set.
To determine which patients would be included in our machine learning model.
So we included all patients from birth to 18 years old. As you can see in our inclusion criteria side, who had a documented diagnosis of bronchopulmonary dysplasia, chronic lung disease or pulmonary hypertension.
These were identified using ICD 9 and 10 codes.
I will preface that the ICD 10 codes were much more specific when it came to the

neonatal bronchopulmonary dysplasia and pulmonary hypertension.
The ICD 9 codes did have a little bit more laxity in these kind of like neonatal chronic lung disease criteria.
Additionally, we also obviously they had to receive ECMO support between the years of 2012 to 2020 and narrowed to this time window just to allow for more current practice but still ECMO itself has changed so much in the last couple of years, especially with the new.
Data we obtained during the COVID pandemic.
And additionally, we also used cardiac and respiratory support only. We did apply several exclusion criteria that you'll note on the right hand side of the screen.
So those where we floated any patients over 18 or any patients with adult Physiology at that point and obviously protocols in the adult ECMO population might differ from the pediatric and the neonatal ECMO cohorts.
We also excluded patients who underwent central cannulation.
Meaning that any patient that was essentially had their chest opened and was cannulated directly into the vessels around the heart or the heart itself were excluded.
Mostly because these are often in the surgical context, either in the OR post cardiac repair, and that also is going to be different in peripheral ECMO.
And then finally, we did exclude any patients who underwent extracorporeal cardio pulmonary resuscitation or ecpr.
These cases represent a fundamentally different clinical scenario.
Where the timing of cannulation and the associated arrest Physiology would confound our model. So by refining the cohort this way, we were able to ensure that the model learns from a more clinically coherent group. Those were patients with chronic respiratory disease who had peripheral ecmocannulation for either resp.
Or cardio pulmonary failure, but not in the setting of arrest or open heart surgery.
And again, this also allows for the model to focus on patients where ECMO was a considered a clinical decision.
With you can actually intervene and have actionable items.
So let's talk about why machine learning.
And critical care were constantly flooded with data.
I mean, we love labs.
We love vitals.
Our Q1 vitals are Q 15 minute vitals and obviously there's often a lot more in play in

the ICU, especially if children are intubated. You have a whole now set of ventilator settings.
Now you have blood gasses.
Additionally, if kids are intubated, they often require sedation, so there can be a lot more clinical points in the ICU that we have to.
Account for.
And pick you Pediatrics. Unfortunately, does lag behind in terms of machine learning compared to the adult world where there's lots of adult machine learning models.
And now in the last couple of years, you'll notice that in the AAP, all the different journals that are coming out, you're starting to see more and more machine learning studies coming out.
So you'll notice on the right hand side of the screen these are just three examples of both in the PICU and in the NICU of some of the most recent literature.
That's come out where they're trying to build machine models to also potentially be integrated into whether it's the EMR or clinical decision making.
The reason why we love machine learning is it allows us to analyze large multidimensional data sets. It can uncover non linear patterns and interactions that it can be difficult to detect using conventional regression techniques.
Think of it as a tool that learns from the data without us having to necessarily predefined how variables may relate to the outcomes and the PICU machine learning has been explored for predicting sepsis.
Ards mortality.
But there's a couple studies out there with different ECMO.
Predictions subgroups.
But our study is aiming to fill the gap by applying these methods to a large ECMO registry with a specific subset of the population that hasn't had a machine learning model developed before.
Look about machine learning development process.
So you're going to start by obtaining your data set so that data came from the also registry.
So obviously you notice that there was a large amount of ECMO cases. Obviously not all of those pertain to our age category.
Not all of those pertain to our diagnosis codes, so we submitted our ICD 9 and 10 codes to help them narrow and refine, and then they sent us essentially about eight different Excel spreadsheets with all of the data that all had to be processed.

So what does that mean?
That goes into your data cleaning.
So we had to essentially take these seven XL8XL spreadsheets, combine them into one cohesive spreadsheet.
By their essentially like their ECMO code marker.
But obviously we know that there's going to be duplicates in there, so you have to remove any duplicates. If there's inconsistencies and then this is where you also need to consider handling missing values.
And so in our specific subset, what we did is if a category had greater than missing 25% of the data, that category was automatically excluded from the evaluation.
So for example.
In the last couple of years, the also.
Registry has started including lactate levels pre and post ECMO initiation, which is fantastic, but because it's only been in the last couple of years there's not enough data for the eight years that our study is covering for us to include that variable.
But when we have less than 25% missing data, we can actually use machine learning techniques to actually input data for both continuous and categorical variables using some different coding process.
And so we are able to then for those that have only less than 25%, generally about 5 to 10%, we can use the data we have to predict what those values should be in that in the missing data.
And so then we go into feature engineering.
So this step is critical.
We can transform critical inputs such as your ECMO mode, your pre ECMO vitals, your gasses into structured features. The model can learn from.
So for example, we can also.
Develop features as well or variables.
So we did some research and one of the things that research has recently found is that if you so obviously a lot of reasons why patients have to go on to ECMO is generally their inability to have good gas exchange.
They generally have very high PCO 2 levels at the time of ECMO initiation.
What research has recently found is that if you reduce that number too quickly.
Can actually have detrimental outcomes for patients.
And So what we did is using all of the data we had is we created variables.
So we actually developed and calculated what the relative change in the PCO two

was in that first 24 hours as a new variable.
Additionally, I know some of my NICU colleagues are on here.
We love a good oi and oxygenation index, so we calculated that from the ECMO data as well.
And in the pic you also love O is and PF ratios.
So we also calculated the PF ratios.
Additionally, during our feature engineering, we can also identify variables that are important that could potentially or should be included in the model itself.
But now we need to split up our data and so this can be done whether so your training data is going to be about 70 to 80% of the data from this data set and then your test data is going to be that 20 or 30%.
Remainder specifically for our study, we did a 7030 split.
So your model is trained on the training data and then to see if your model works well. You then use your testing data to see how well that model plays out.
So essentially your training set is going to teach a learning algorithm.
So the learning algorithm is where your different machine learning types come in, and in this process we did supervised machine learning. So we told what we wanted the outcome to be and then you can we'll go into some different machine learning algorithms as well to kind of talk.
About how we picked the best model.
Cool. And so we use multiple algorithms trying to find the best one, which you'll see next, and then finally, once we had models, we then used the testing data to score it.
And so this model evaluation essentially is where we look at our AUC, our sensitivity, our specificity and the interpretability of the model itself. And so these steps really mirror clinical approach, right, you gather data.
You clean it, you define what your question is.
And you can test a few different diagnosis or algorithms and then assess how well it works.
Let's talk about some different types of machine learning that you might often hear.
So we'll start from the basics so often.
A lot of us know linear regression pre machine learning, but in the early days of machine learning and a lot of the early studies you'll see for machine learning, they still were using a linear regression.
They were just using machine learning to do the linear regression.
So an extension of linear for the generalized linear regression. It's an extension.

Of linear regression model where it models the relationship between.
One or more predictors or your independent variables and has a binary outcome using a Lync function for binary classification like survived or died.
Logistic regression is most common because it models the log odds of the outcome for linear function of predictors.
So the pros of this is that it's highly interpretable. Your coefficients directly show the strength and direction of associations.
It's familiar to clinicians and researchers.
It's equivalent to multivariable regression models.
It's fast to compute, even on large.
Data sets, but the cons of using this is that it assumes that there is linearity between predictors and the odds of the outcome, which may not hold up in complex clinical settings. It's limited in modeling interactions and non linear relationships, and it struggles with high dimensional data so.
Lots of predictors relative to the sample size, so a nice analogy for this is calculating an odds ratio and logistic regression table.
It's clean and informative, but it may miss some of the complex.
Physiologic interactions.
One of them after the linear regression, a lot of what was coming out was neural networks.
So here you'll see.
A kind of different inputs with your hidden layer. Your output layer for the kind of concept of how neural networks work.
So it's composed of layers of these interconnected nodes, or neurons that stimulate brain like processing.
So each node computes a weighted sum of its inputs, passes it through a non linear function.
And sends the result to the next layer. So you'll see that blue is your input layer. Then the purple is that hidden layer where it sends it to the next layer, and then eventually makes it to the output layer.
So it learns complex non linear relationships through back propagation, updating internal weights that minimize the error.
The pros of using a neural network.
It's excellent in modeling complex high dimensional data and it can automatically learn non linear interactions between variables.

It's scalable to large data sets and adaptable to various data types, but the issues with it is that it has low interpretability, often called black boxes, because we can't easily trace why these decisions were made. It requires more data and a longer training time.
And it's prone to overfitting if it's not properly regulized.
So it's like watching an experienced clinician make a decision, often correct.
But difficult to fully explain their internal reasoning.
Kind of like that gut feeling you get.
And you pretty much know what the diagnosis is.
But you can't quite necessarily give like a like a very good scientific clinical reasoning for why. You just know that's what it is.
Then the next ones you'll look into are kind of the ensemble models.
So random Forest is the first one.
You'll see them some decision trees there.
This is an ensemble model built from hundreds of decision trees. Each tree is trained on a random subset of the data and features.
So it picks a random group of the data a couple of the features included, and then each the final prediction is based on essentially majority vote or the classifications or the average prediction from all of the trees combined.
The pros of this is that it again handles non linear relationships and interactions very well, but it's and it's robust to overfitting, especially compared to single decision trees and it can manage missing data.
And imbalanced datasets fairly well, but it's slower than simpler models.
It can become less interpretable as the number of trees increases and its performance gains plateau with very large forests.
So analogy here would be. It's like a multidisciplinary team would be each tree voting on a diagnosis more stable than one expert alone.
But it may not give you a clear explanation.
So think of a very complex kid and you have five or six different teams all consulting trying to figure out what exactly is wrong with that kid.
And each team is kind of voting on what they think the diagnosis is.
And then at the end, everyone kind of comes together based on all of those votes to decide what they think the diagnosis is.
And then last up is extreme gradient boosting which is the model that ended up being the most successful in our study that was used.

So this is also a boosting algorithm that builds trees sequentially, but the difference here is that each new tree corrects the errors of the previous one, and so you'll notice as you go from left to right on the screen that you start with the original data set. And then it modifies, modifies, and continues this until you kind of, as it builds and corrects its own errors, and it uses a gradient descent approach, so it minimizes the loss function with each new tree, and it also includes regularization to prevent overfitting.
And it can also have built in handling of missing data.
The pros of this is that it's state-of-the-art performance for tabular clinical data.
It's highly accurate, especially.
In complex data sets that have many variables.
And there's also built in tools for feature importance which can help what the model has learned. Issues with gradient boosting is that it is more computationally intensive than random forester regression.
It does require tuning of a lot of hyper parameters like your learning rate, the tree depth, the number of trees, and it's still a lot more complex to interpret than linear models.
So again clinical analogy.
You're a physicians.
Who is improving their diagnosis?
Over several days or learning, you know, seeing acute liver failure multiple times and every time they learn how to be more efficient, how to minimize lab draws to essentially interpret the data, use the most effective labs and exams in response to therapy and overall to figure out the.
Best way to work up, assess and treat that problem so it's a learning process that gets better over time.
So that was a little overview of machine learning. Hopefully those of you that weren't as familiar with it feel a little bit more comfortable kind of understanding how it differs from our traditional logistic regression.
So let's talk about how our model came out.
So we used all four of those model types and used the data in hundreds of combinations to identify the best model. At the end, our model had over 30 or sorry at the end, the best model had 30 different.
Variables and those are going to be broken down over the next three slides.
So this is our univariate analysis of those 30 variables that were included.

So you'll notice that on the screen starting from the top the most these are our demographics and kind of some of the past medical history.
So congenital diaphragmatic hernia. If there was a history of congenital diaphrhrmatic hernia had a substantial impact on patient mortality.
What's really important is also how important race and ethnicity.
Was which.
We'll talk about more in the discussion aspect and then obviously whether that was a history of chronic lung disease or bpd versus pulmonary hypertension also was an incredible important variable.
And then the other ones that were included in the demographic data was sex, age and weight of note.
That was the age and weight at the time that they went on to ECMO.
The next subset is what their pre ECMO support was, and so this included whether they what type of drips they were on.
Sometimes it was specific drips versus more general.
Also what their pre laboratory vitals were or values were and their vital signs so of the pre ECMO data you'll see that there was quite a bit of significant data that was included.
So specifically a pre ECMO dopamine drip was associated with mortality, but also the use of any.
Face oppressor or dionotropic drug. But as we know, obviously those are going to be sicker kids that are going on ECMO at that point.
Mil renown use were also associated with mortality from a lab standpoint and a vital standpoint.
Your SA 02, your pH, your bicarb and your pco.
Two pre ECMO were statistically significant.
Additionally, your diastolic blood pressure and your PF ratio, which we all understand is incredibly important in that ARDS lung, respiratory distress picture.
And additionally, the pre ventilator, respiratory rate and systolic blood pressure were included in the model, but in the univariate analysis.
We're not statistically significant, nor was the use of a neuromuscular blocker pre ECMO.
The following is the ECMO support your laboratory values and the vital signs after 24 hours.
So the mode of ECMO was statistically significant.

Other statistically significant was your mean arterial pressure at 24 hours?
Again, the ventilator rate at 24 hours and then as we had mentioned earlier, the relative change in the PCO two and you'll notice.
That it was higher in the groups that died compared to the group that survived, as is supported in the current literature.
Additional variables that were included, but we're not significant in the universe analysis where the Fio two at 24 hours.
The PIP on the ventilator at 24 hours.
The PAO 2.
The bicarb on the gas.
The P 24 hours and then the ECMO pump flow at 24 hours. Of note in the also registry. The pump flow data is collected 4 hours post ECMO initi.
Initiation and at the 24 hour mark during the model development, the ECMO pump flow at the 24 hour mark was always subsequently identified as an important variable in model development where the four hour pump flow was less so.
Let's talk about what this means in terms for our model.
So here's our receiver operating characteristic curve or the Roc curve for short of that final xg boost model with those 30 variables that you saw on the previous slide, the area under the curve or the AUC was 0.725 with a 95% conf.
Interval of .671 to .779.
For those less familiar, an AUC of .5 is no better than chance, while one is perfect.
Discrimination, so, .725 suggests moderate accuracy, which is consistent with and in some cases slightly better than existing models in similar populations. Most importantly, the model maintains stable performance across bootstrap samples, which does support internal validity.
And we'll explore more interpretability and feature importance in the next few slides.
So here's our confusion matrix.
So if you remember going back to how the machine learning algorithm works, that confusion matrix specifically includes the testing data set that was used to test the model on.
So these are the actual numbers for the confusion matrix from that testing data set.
So we know that for those a quick reminder from our you know, board testing days, we know that for a confusion matrix are true positives.
Are patients who survived and were correctly predicted to?
So that's this one right here. We know that.

Our two negatives are patients who died but were and were correctly predicted to die.
And then we know that we have our false positives. Who are patients who are predicted to survive.
But did not.
And then the false negatives are those that the model predicted would die but actually survived.
So one key observation here is our model did demonstrate a relatively high sensitivity surf for survival more so than mortality. In other words, our model was very accurate at predicting which patients would live, but had less precision in identifying those at risk of death, although the pattern's.
Not uncommon in clinical.
Predictive modelling, especially those with high mortality settings and may reflect both the limitations of the input data and the fact that death is often the result of complex.
Dynamic processes.
Not fully captured by early variables, we'll revisit the limitations later on in the presentation.
So the next is and I apologize for how small this is, but this does include all 30 variables, and then you'll notice from the top it is the most important variable all the way to the bottom that had the least impact on model output.
So this is the Schaeff plot, or a shapely additive ex planations.
It's a visualization tool that we use to explain how machine learning model made its predictions, and specifically which variables were most important and how they influenced the model's output.
So you'll notice at the bottom the colours, so you'll notice the feature value being low versus to high. So purple means there is higher more clinical relevance whereas yellow had lower or less impact on the model output.
So they shop. Values are based on our shapely values, which is the concept from game theory that determines how much each player, or in this case, each fear feature or variable contributes to the final outcome.
So imagine a prediction as a team score and each variable.
As a player on that team, the shop diagram tells us how much each player helped out or hurt the prediction.
And so you'll see that congenital diaphragmatic hernia was the most important. The

use of dopamine pre ECMO was the second most important on the outcome of death in the model race was the third most important variable.
And so I'd like to take a moment to pause here and discuss.
The impact that that has and as we know in medicine.
And we tried to be as equitable as possible, and we know that we've made strides, especially in the world of Pediatrics, to improve that.
But the fact that in every single model we ran at times, both race and sex were incredibly important variables. It does limit using this model in the real world because we don't want this model to further.
Target people based on their race for determining whether or not to offer this service.
Which is why this overall model can be used to help identify or help give families and clinicians information on the likelihood of survival.
But it shouldn't be used as a clinical decision making tool so that way it doesn't further.
Potentially contribute to racial practices in medicine.
So let's talk about some of the current models and model evaluation out there.
So initially an AUC of .72, we can put our findings comparatively to some.
The recent literature that's been posted, but since there have been a limited number of studies out there.
So specifically, the first one is that study from 2018, which used looked specifically at ECMO use in congenital diaphragmatic hernias and then that model they used both pre post ECMO data and also had an AUC of 0.73 which is very close to where our.
AUC was as well.
And then following two years.
In 2020 was the extracorporeal membrane oxidation, specifically in just pediatric pulmonary hypertension.
Now they made three different models using the also registry, so they used built a model only using the pre ECMO data which had an AUC of 0.62.
They did a ECMO related model, which was just everything at the initiation of ECMO to post ECMO, which had an AUC of 0.72.
And then they combined all of those variables and had an AOC of .75.
So again, our model falls right in the middle there.
And then the last one, which is the most recent one, was the development of the model for pediatric survival after vino arterial extracorporeal membrane oxygenation

score also referred to as the PD save score which came out in 2022.
And they also did a precannulation model and a post cannulation model, and the Precannulation model had an AUC of 0.64 and the post cannulation model had an AUC of 0.71.
So this also.
Helps highlight that while our model was.
Within the range of prior models and in some cases slightly better, despite focusing on a population that may be more heterogeneous and difficult to model it.
It does also highlight that this data set that we get from the Elsa registry does have its limitations, and it is very limited. And so despite us narrowing the years, narrowing our subset of patients, that data is still still has a lot of variability in it and does.
Limit.
How accurate our models can be?
So if I'm going into that, no model, no study as we know is without limitations and ours is no exception.
So as I mentioned, the model was more accurate.
It accurate in predicting survival than mortality, but this asymmetry likely reflects several factors.
So limited mortality, specific features, survivor bias and available data, and the fact that early variables may not fully capture deteriorating trajectories.
Second, our aim was to build a pre ECMO prediction tool.
But we did find, as you'll notice in that previous slide, the pre ECMO models often performed far less, far worse than including the post ECMO data.
Or including all of the data we had that similar issue once we included that post 24 hour ECMA data data. Our model improved substantially, which does present a challenge for developing real time clinical use.
It's the best model.
Does require post initiation data.
Then obviously.
Including race and sex is variables in your model.
It does limit using it as a clinical decision augmentation for when to deciding when to offer ECMO.
And obviously missing data is always going to be a constraint in most of studies.
And so again, while the LCO repository of data is amazing, it's very limited in the data that it collects.

You get one set of vitals, one blood gas, and one set of ventilator settings pre ECMO initiation, and then you get a repeat set of those labs vitals invent settings 24 hours post ECMO initiation.
But we know that there's a lot more than just those variables that go into deciding when to offer ECMO, how the process of placing on ECMO and what they look like 24 hours post ECMO. We know that a lot can happen in that 24 hours.
But we also understand that this massive data repository.
Can't collect hundreds and hundreds of these data points cause all of this data that Elso collects is manually inputted by a person at each of those LSO data centers. Which aids to then go into that. There can also be errors within that data set since it is reliant on human imputation.
And so obviously as the last point points out, because of the limitation of those two time points, obviously we're not really capturing the full dynamic clinical course of patients.
Prior to ECMO and in that first 24 hours prior to getting that next set of labs.
So where do we go from here?
So to wrap it up again, this model is not meant to be a replacement for clinical judgment and not really meant to aid in deciding at this point when to place a child on ECMO, because we are relying on those ECMO parameters at the 24 hour mark. But it is giving us more information so families and clinicians can be informed on at that 24 hour mark what this patient's likelihood of survival to discharge is.
So from a future standpoint.
From a future focus after we understand that.
We can continue to refine the model.
We know that our knowledge of ECMO during the COVID pandemic has provided a lot more clinical features and we know that even from when we started using ECMO in the NICU in the PCU, you know everyone went on VA ECMO and now we use a combination of V.
And VA ECMO, depending on what the patient's needs are. And so over the last 2010 five years, the practice of ECMO itself has changed.
Changed. I mean you could send there are hundreds.
You know you could send hours just talking about how anticoagulation over the last couple of years has changed substantially in ECMO and so limiting yourself to the last like the last five years, the last three years and trying to narrow to eliminate as much heterogeneous data in the.

LSO data set is a potential way. We can also proceed forward to try and get a more accurate model and obviously work on eliminating.
The like the race and the sex characteristics from our model.
So that way it can be equitable and be more interpretable in terms of using it. If you do want to use it as a clinical decision making device.
And again, our whole goal for this project is just to have a transparent, equitable resource allocation, especially in a high cost, high stakes therapy like ECMO.
Ecmo and this is kind of the first step as the other studies that use that I presented are just tools like this to help us get there.
So that way we can make sure that.
We're putting patients who are hopefully gonna have the best outcomes and the best position in trying to also use it as a kind of notifier to hopefully get ECMO initiation started sooner when we think patients will have a good outcome with it as well.
So that is my whole presentation.
I will happily take any questions at this point, and I also recognize that obviously tried to do a quick overview of as much machine learning and our project as possible.
So I'm happy to take any questions at this point.

Kamat, Deepak M 44:13
Doctor Mangold, thank you very much for that outstanding presentation on machine learning in patients with ECMO on in with chronic lung disease.
There's already one question in the chat box by Doctor Lukoffer.
Did you control for cofactors?
That may explain the race findings. For instance, maternal age, education, etc.

Mangold, Cheyenne M 44:37
No, unfortunately we did not.
Just because again, the there's very limited on the data 'cause. Obviously the other challenge with that is we didn't get the from a HIPAA standpoint since it's in DE identified registry.
Obviously, we also have international data with U.S. data based on the number of ECMO centers in the world.
And so that was one of the limitations as well as that because we didn't have the

regional data, we also couldn't account for.
How also different?
Countries different areas of origin were playing into the demographic data as one of our limiting factors as well.

Kamat, Deepak M 45:15
Thank you.
There is a question by Doctor Perlman.
What software did you use to do the machine learning?

Mangold, Cheyenne M 45:21
Sure. So all of my coding was done in R, the coding software.
And then I the most recent version, I don't have the version off the top of my head, but using the most recent version of R is where we did all of our coding.

Kamat, Deepak M 45:36
Any other questions?
Comments for Rathman good doctor Brooks, have we considered including physician in implementation of likely or impression of likely success of ECMO I think.

Mangold, Cheyenne M 45:51
We did not again.
Very limited on what our L sale registry was, and so and obviously the total number of patients was also about 1600 patients. And so since it was from that data registry, we didn't have any data like that.
Obviously it's something you know, we could have some of our LSO docs take a look at.
They'd have to look at each, you know, data individually and set the likelihood of survival based on the pre ECMO markers.
But.
But you know, like that's the again the challenge with these repositories is that there's not any subjective data included in them.
It's mostly it's all objective data.

Kamat, Deepak M 46:32

Thank you.
Are there questions, comments for Doctor Manko?

Armijo-Garcia, Veronica 46:43
Doctor Mangold's Dr. Armijo.

Mangold, Cheyenne M 46:45
Yeah.

Armijo-Garcia, Veronica 46:47
Knowing what you know now, having gone through this machine learning model.
Are there any other ECMO populations that you think might?
Favor.
In a positive way, giving us survival slash mortality outcomes.
Maybe for another project.

Mangold, Cheyenne M 47:16
Yeah. You know, I think obviously this was a combination kind of Nick, you pick you overlap.
But when I was looking at the current literature, obviously you know there was pulmonary hypertension, congenital diaphragmatic hernias.
This one was kind of a combination of chronic lung disease, pulmonary hypertension, study, a lot of them. I felt like focused on more NICU related causes for ECMO use.
And so and then obviously there was quite a few that were cardiac surgery specific.
But when I was looking at the literature, there wasn't a lot that focused on more of our PICU population.
So acute respiratory failure in like kids who have normal Physiology or asthma.
Some of the other like more adolescent type lung diseases. I didn't really find any literature or any studies that looked at those populations in their ECMO outcomes.
And so I do feel like there is quite.
A gap in terms of more PICU related patients.
Where there's less of our NICU problem overlap, that is kind of an untapped area that could easily build prediction machine learning models.
And obviously that's all going to be dependent on the ICD 9 and 10 codes that are put in.

But you know, I think even looking at what the number of asthma patients who require ECMO, knowing how risky, you know, intubating and ventilating an asthma patient or, you know looking at pediatric influenza who require ECMO.
So I do think there is quite a few.
Untapped areas that could benefit from looking at subset populations in the PICU population.

Kamat, Deepak M 48:54
This question from Doctor Parmar when do you think raw continuous writer science data will be able to be used in machine learning?

Mangold, Cheyenne M 49:04
I mean, we're seeing it used in.
So, you know, Epic is a fantastic resource and epic. Does you know on our kids who have Q&R vital signs. And so I think that data does exist in individual EM Rs being able to pull that large number of data and get by the minute V.
Signs or even you know, by the hour, vital signs and use much more data. I unfortunately think from an ECMO and also repository standpoint.
And obviously also does look.
And I'm sure Armijo can speak to that, too.
They do look at the data they collect, but obviously it would be impossible for someone to have to manually input, you know, every hour. Vital signs 24 hours leading up to ECMO.
24 hours post ECMO directly into that registry.
And so I think until we can figure out how to integrate, you know our EMR and and also from a hipa standpoint and all the ethics.
And legal ramifications that go without, you know, I don't think that would be possible.
For these large datasets, until there's some way to integrate the EMR directly to feed that data.
Or some sort of agreement between hospitals to be able to kind of transmit that data. And I know Vps which is one of the ones that's used utilize and pick you quite a bit does have a little bit more of that integration ability.
But I know right now the LSO data set doesn't quite have a system like that set up to get those kind of hour by hour vital signs.

And obviously there's a lot of lab work missing. You know, a blood gas.
Doesn't tell you a whole story about the patient.
You know you don't have electrolytes.
We don't have a CBC.
We don't have their coagulation pre ECMO initiation or post ECMO initiation, but we know that all of those factors play a large role in patient stability pre and post ECMO.
And so, you know, there's there's 1000 things I wish I had and would be able to add to this model. But, you know, I think the next step to kind of get more accurate data points would be trying to use.
You know EMR data from maybe one hospital institution and looking at these more as you mentioned like hourly vital signs, more lab data to give yourself and give us more information to kind of build our prediction models.

Kamat, Deepak M 51:19
Thank you.
Any other questions?
Comments for Doctor Mango.
I don't see any, just a quick reminder next week, next Friday, May 23rd is a pediatric research forum which starts at 12:30 PM in room 409 L.
Lunch will be available and Doctor Meyer is will be our senior speaker and he'll be giving his pH presentation at 1:00 PM.
Also, we will not have grand rounds next week because of Memorial Day holiday.
Weekend, we'll see you in two weeks for the next.
So I'm graduating fellow presentation.
Thank you all for attending this morning's round.
Have a wonderful weekend.
Thank you.
Thank you, Doctor Mangold.
Really appreciate for a wonderful, outstanding presentation.

Mangold, Cheyenne M 52:13
Thank you so much.
Have a great weekend everyone.

Kamat, Deepak M stopped transcription