Differential Calculus Project

Sigmoid Functions and Vaccine Dose-Response Curve

Navya Jain, Katiana Nadolsky, Kaira Nair

Description:

This question needs the logistic sigmoid function and general differential calculus skills. We may need some additional biology to explain the parameters for the gradient, midpoint, and maximum given a specific vaccine.

A sigmoid function is any mathematical function whose graph has a characteristic S-shaped or sigmoid curve. If the drug dose is plotted on a base 10 logarithmic scale, this produces a sigmoidal dose-response curve.

The specific sigmoid function we are interested in is:

$$\text{Saturating (dose)} = \frac{\text{maximum}}{1 + e^{\text{gradient}(\text{midpoint} - \text{Dose})}}$$

where the maximum is the maximum effectiveness of the vaccine, the gradient is the slope at the midpoint, and the midpoint represents the concentration needed to produce 50% of the maximum possible effect. Alternatively, the midpoint could be defined as the dose of a vaccine required to produce a specific desired effect in 50% of the population.

Note that, in reality, it is ligand concentration (and resulting receptor occupation) that affects response - the term ‘dose-response curve’ assumes that the drug dose and ligand concentration are closely linked.

Question 1 - Group

We are interested in the math behind sigmoid functions and their application in biology, specifically regarding vaccine dosage-response relationships. The questions that everyone in class will do are the following.

Suppose you have a vaccine whose dose-response curve is modeled by a logistic sigmoid function below. Find the point of inflection. How does this relate to the original equation? What happens as the dosage approaches infinity?

$$f(x) = \frac{M}{1 + e^{g(m-x)}}$$

where M, m, and g are all constants.

Question 2 - Navya

The Hill equation (https://en.wikipedia.org/wiki/Hill_equation_(biochemistry)) is a way to model the relationship between ligand concentration and protein saturation in a biochemically accurate manner (given some constraints). This equation looks like

$$\theta = \frac{[L]^n}{K_d + [L]^n}$$

where theta is the protein saturation, [L] is the ligand concentration, K_d is the dissociation constant, and n is the Hill coefficient. My follow-up question is what is the inflection point for this equation, and how can we potentially relate this to the solution for the above equation (given that these are both used in dose-response)?

Question 3 - Katiana

How do these equations relate to the actual dose-response curves of vaccines administered during the Covid19 outbreak? At what point is the dosage considered optimal?

Question 4 - Kaira

How do we determine the parameters (M, g, m) of the sigmoid function from experimental data? What is the effect of over- or under-estimating these values when fitting real-world vaccine response curves?

Visuals

We modeled sigmoid functions using Desmos, with parameters for different initial potency, gradient, and midpoint.

Here is a Desmos model with adjustable variables of a Sigmoid function: https://www.desmos.com/calculator/yuc6glfnjc

Here is a Desmos graph with different Sigmoid functions (from Dr. Loveless):
https://www.desmos.com/calculator/mnqg38eved

Solutions

Question 1 - Group:

1. Find the first derivative of the logistic sigmoid function
2. Find the second derivative
3. Set the second derivative equal to 0
4. Solve for x. Relate to original equations

Here are the worked out steps.
First, we have the equation

$$\text{Saturating (dose)} = \frac{\text{maximum}}{1 + e^{\text{gradient}(\text{midpoint} - \text{Dose})}}$$

We can say that the dose is x, the maximum is M, the midpoint is m, the gradient is g, and the saturation is f. This gives us the below function.

$$f(x) = \frac{M}{1 + e^{(g(m-x))}}$$

We can rewrite this as shown below.

$$f(x) = M\left(1 + e^{(g(m-x))}\right)^{-1} = M\left(1 + \frac{e^{gm}}{e^{gx}}\right)^{-1}$$

Next, we can find the first derivative and simplify. These steps are shown below.

$$f'(x) = -M\left(1 + \frac{e^{gm}}{e^{gx}}\right)^{-2}(e^{gm})\left(-ge^{-gx}\right)$$

$$f'(x) = \frac{Mge^{gm}e^{-gx}}{\left(1 + \frac{e^{gm}}{e^{gx}}\right)^2}$$

Next, we can find the second derivative and simplify.

$$f''(x) = (Mge^{gm})\frac{-ge^{-gx}\left(1+\frac{e^{gm}}{e^{gx}}\right)^2 - 2e^{-gx}\left(1+\frac{e^{gm}}{e^{gx}}\right)(e^{gm})(-ge^{-gx})}{\left(1+\frac{e^{gm}}{e^{gx}}\right)^4}$$

$$f''(x) = (Mge^{gm})\frac{-ge^{-gx}\left(1+\frac{e^{gm}}{e^{gx}}\right) - 2e^{-gx}(e^{gm})(-ge^{-gx})}{\left(1+\frac{e^{gm}}{e^{gx}}\right)^3}$$

$$f''(x) = \left(-Mg^2e^{gm}\right)\frac{e^{-gx}\left(1+\frac{e^{gm}}{e^{gx}}\right) - 2e^{-gx}\left(e^{gm}\right)\left(e^{-gx}\right)}{\left(1+\frac{e^{gm}}{e^{gx}}\right)^3}$$

$$f''(x) = \left(-Mg^2e^{gm}\right)\frac{e^{-gx}\left(1+\frac{e^{gm}}{e^{gx}}\right) - 2e^{gm-2gx}}{\left(1+e^{gm-gx}\right)^3}$$

$$f''(x) = \left(-Mg^2e^{gm}\right)\left(\frac{e^{-gx}+e^{gm-2gx}-2e^{gm-2gx}}{\left(1+e^{gm-gx}\right)^3}\right)$$

$$f''(x) = \left(-Mg^2e^{gm}\right)\left(\frac{e^{-gx}-e^{gm-2gx}}{\left(1+e^{gm-gx}\right)^3}\right)$$

Setting the second derivative equal to 0, we get

$$0 = \left(-Mg^2e^{gm}\right)\left(\frac{e^{-gx}-e^{gm-2gx}}{\left(1+e^{gm-gx}\right)^3}\right)$$

Solving for x, we get

$$0 = \left(\frac{e^{-gx}-e^{gm-2gx}}{\left(1+e^{gm-gx}\right)^3}\right)$$

$$e^{-gx} = e^{gm-2gx}$$

$$-gx = gm - 2gx$$

$$gx = gm$$

$$x = m$$

This means that the point of inflection where the rate of saturation starts decreasing is the midpoint when 50% of the population has been reached!

In the original equation, as the dosage approaches ∞, the saturation approaches M, or the maximum. This justifies the presence of the M in the original logistic equation.

Question 2 - Navya:
https://drive.google.com/file/d/1bQO1SGABTIIbogMlrwixoBRXcC-Qei9c/view?usp=sharing

The question and solution can be found at the above link. I couldn't figure out how to embed LaTeX in Google Docs, so I just included a link instead.

Question 3 - Katiana:
Source: https://pmc.ncbi.nlm.nih.gov/articles/PMC9574467/

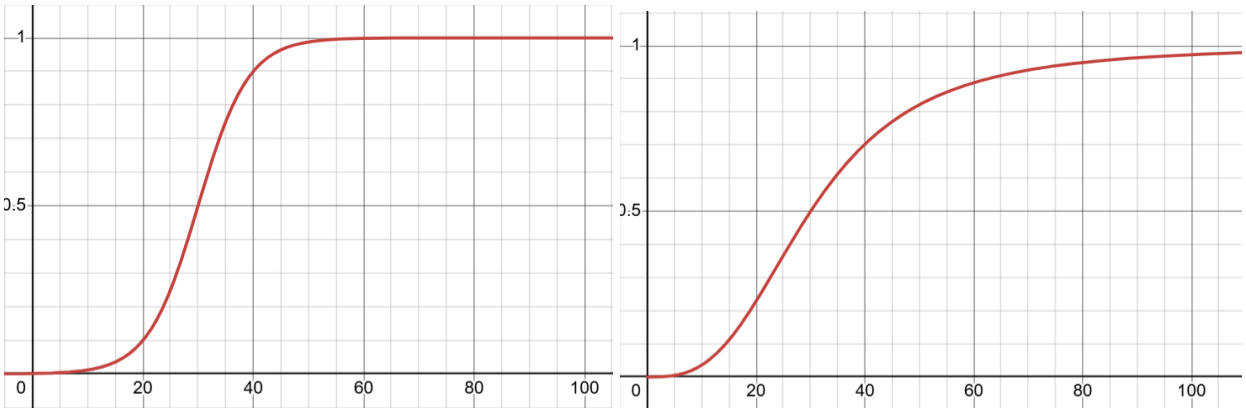
Using this study from the National Library of Medicine on finding the optimal doses summarized across vaccine types from published clinical data on Covid-19. Focusing on the saturating curve, the study chose to use a sigmoidal curve like our example, modeled by:

$$Response = \frac{R_{max}}{1 + \left(\frac{R_{50}}{Dose}\right)^{p}}$$

where Rmax is the saturation maximum, p is the gradient as defined earlier, and R50 is the value where the response is 50 % of the saturation maximum. While they may look similar, in application, the equation greatly differs from the logistic curve we used for our example.

Here is a Desmos graph modeling the dose-response curve equation used in this study:
https://www.desmos.com/calculator/vfdovojrs8

I will compare this model to our previous Desmos model by setting both equation's parameters to M=1, m=30, and p=3 (g=0.22 so the graphs will have a similar response at 5 units of dose). Our logistic curve is on the left and the study's saturating curve is on the right.

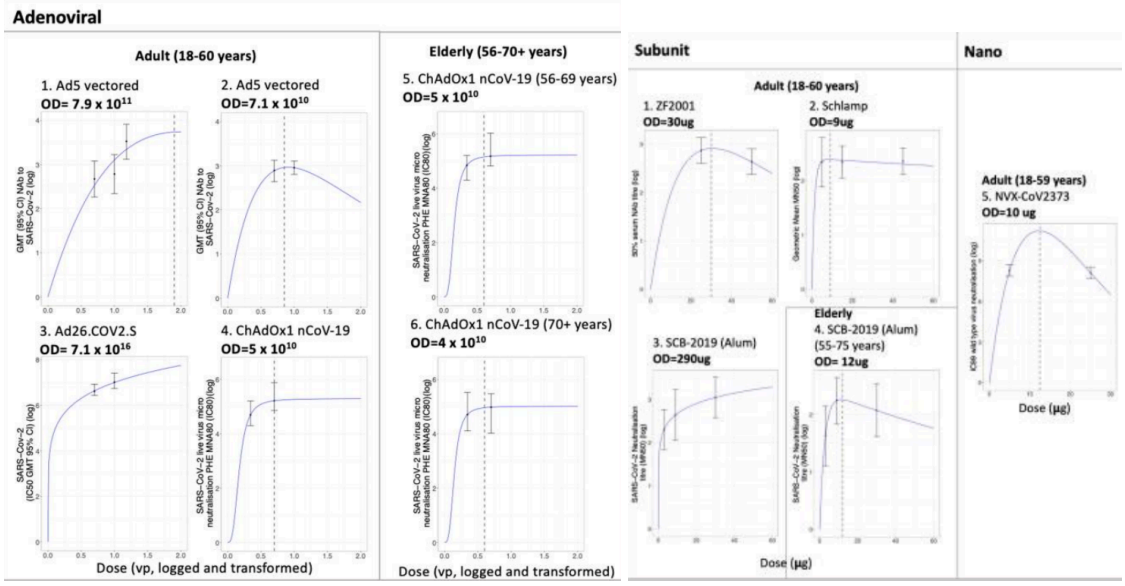


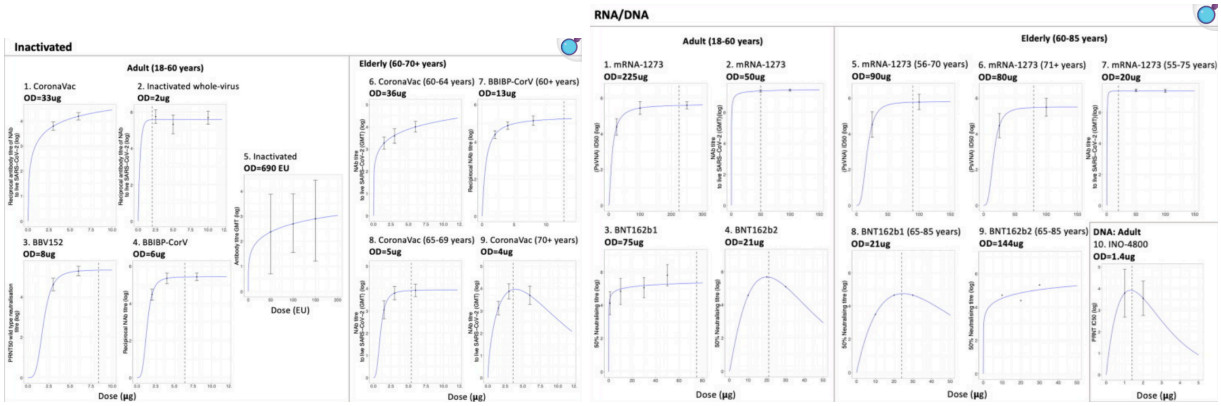
.

Looking at these models side by side, we see that despite having the same maximum, midpoint, and dose-response at 6 units of dose, the graphs differ greatly. The logistic curve peaks much sooner and is steeper with much sharper bends. While the curves appear more similar prior to the midpoint, after the midpoint the model from the study approaches 1 much more gradually. This is why I did not choose to make g equal to p, because doing so made the logistic curve's slope practically vertical and the graphs were incomparable. This amount of difference is interesting to me as I wonder where each model would be more applicable in real world contexts. These differences could have to do with how the 'response' on the y-axis is measured, or it may just be that logistic curves aren't as accurate in practice as our initial research. This would be interesting to look into further.

The optimal dose for a saturating dose-response curve is defined as the smallest dose for after which there is a negligible increase in response for an increase in dose. Since the curve will continue to increase towards the saturation maximum, a negligible increase should be defined. To define a negligible response, it is when a unit of dose leads to a less than 0.1% increase in response.

The following are the predicted dose-response curves modeled from the study, sorted by vaccine type and labeled by age group:





As you can see, many of the curves are modeled with peaking curves (the response decreases after the optimal dose), rather than a saturating curve like we chose to study.

In practice, the doses administered compared to the predicted models varied from below, within, and above the predicted optimal doses. Looking at the two most common vaccines used in the USA, Pfizer and Moderna (both RNA/DNA vaccine types), the doses administered were within and below the predicted amounts overall respectively. Personally, these results make sense to me as the vaccines were developed and administered quickly, and since many dose-response curves were peaking, it makes sense that the doses administered would be within or below the predicted amount as to not administer more than necessary for cost and health risk related reasons.

Question 4 - Kaira:

<u>How do we determine the parameters (M, g, m) of the sigmoid function from experimental data?</u>

To determine the parameters (M, g, m) in our sigmoid function, from experimental vaccine dose-response data, we can use a statistical technique called non-linear regression, specifically the four-parameter logistic regression model.

We can use this by :

1. Gathering Experimental Data: Conduct experiments where different doses of a vaccine (our 'x' values) are administered, and the corresponding immune response or protective effect (our 'f(x)' values) is measured. This gives us a set of data points, for example, (Dose 1, Response 1), (Dose 2, Response 2), and so on.

2. Plotting the Data: When these data points are plotted, they typically form the characteristic S-shaped curve that the sigmoid function describes.

3. Using a curve fitting software: Use GraphPad Prism, or programming languages with statistical libraries such as Python's SciPy or R to perform the curve fitting. We input our experimental dose and response data into the software.

4. Mathematical Optimization: The software then uses algorithms to find the values for M, g, and m that make the sigmoid function's curve fit our experimental data points as closely as possible. It does this by minimizing the difference between the observed experimental responses and the responses predicted by the sigmoid function. In other words, it is finding the "best-fit" S-shaped line through all the scattered data points.

The software will essentially provide us with :

1. M (Maximum): The software estimates the highest point the curve reaches, representing the maximum effectiveness or response the vaccine can achieve.

2. m (Midpoint): It estimates the dose at which the vaccine achieves 50% of its maximum effect. As our calculations show, this also corresponds to the inflection point of the curve, where the rate of response change is greatest.

3. g (Gradient): The software determines the steepness of the curve at its midpoint, indicating how sharply the vaccine's effect increases with dose around that point.

This process provides us with statistically robust estimates for M, g, and m, along with measures of their uncertainty (like confidence intervals), reflecting the quality and variability of the experimental data.

<u>What is the effect of over- or under-estimating these values when fitting real-world vaccine response curves?</u>

Accurate estimation of M, g, and m is critical because these parameters define the entire shape and implications of the vaccine's dose-response curve. If these values are over- or under-estimated due to poor data, measurement errors, or flawed fitting, our mathematical model will misrepresent the vaccine's true behavior, leading to potentially significant real-world consequences:

1. Misestimation of M (Maximum Effectiveness):
    a. Overestimation: If we incorrectly estimate 'M' to be higher than the vaccine's true maximum effectiveness, our model would suggest the vaccine can offer a higher level of protection than it actually can. This could lead to an overconfidence in the vaccine's capabilities.
    b. Underestimation: Underestimating 'M' would imply the vaccine has a lower maximum effectiveness than it truly possesses. This might lead to under-utilization of a highly effective vaccine or a perception that it is not very good, potentially discouraging its use.

2. Misestimation of m (Midpoint):
    a. Overestimation: If 'm' is estimated to be higher than the true value, the model would suggest that a greater dose is required to achieve half of the maximum response. This could result in recommendations for unnecessarily high vaccine

doses, which might increase the risk of side effects, raise costs, and strain vaccine supply.

b. Underestimation: If 'm' is underestimated, the model would imply that a lower dose is sufficient to achieve half the maximum response. This could lead to recommendations for doses that are too low to provide adequate protection, leaving individuals vulnerable despite vaccination.

3. Misestimation of g (Gradient/Steepness):
    a. Overestimation: An overestimated 'g' would make the curve appear much steeper, suggesting that small changes in dose around the midpoint lead to very dramatic shifts in effectiveness. This might create an illusion of extreme sensitivity to dosing, making optimal dose determination seem overly critical.
    b. Underestimation: An underestimated 'g' would flatten the curve, implying a very gradual increase in effectiveness across a wide range of doses. This could mask the crucial "sweet spot" where the vaccine's effectiveness significantly improves, potentially leading to inefficient dosing strategies.

In summary, any misestimation of these parameters distorts the vaccine's true dose-response. This can lead to inaccurate predictions of vaccine efficacy, and suboptimal dosing recommendations.