

1 Overview of serological modelling

1.1 Introduction

Serological samples can be analysed to detect the presence of biomarkers made in response to an infection long after the infection has cleared.[?] Therefore, analysing serological samples allows researchers and healthcare professionals to deduce crucial information about the epidemiology of a pathogen at the individual and population level, which active virological surveillance systems may otherwise miss.

On the individual level, after measuring antibodies to a specific pathogen, infection is usually inferred using either i) an antibody threshold level (seropositive) or ii) a threshold fold-rise between a pair of samples (seroconverted).[?] Often, researchers are interested in understanding how seropositivity and seroconversion rates change according to controlled host factors, such as age, XXXX etc.[?] On the population level, serological samples which are representative of a population (e.g. cross-sectional samples) can be used to estimate the prevalence of infectious diseases (seroprevalence) and determine how seroprevalence changes over time according to host factors.[?] Estimates of seropositivity, seroconversion, and seroprevalence can help the understanding of the immune system's ability to combat various pathogens, aid in developing new targeted intervention programmes and provide insights into the transmission dynamics of infectious diseases. The methods used to analyse serological samples to inform epidemiology and public health have been termed 'serodynamics' and have recently been reviewed. [?]

Serological samples play an increasingly important role in public health efforts to combat and control infectious diseases.[?] However, inferring infection through seropositivity or seroconversion requires deriving an absolute or relative threshold hold value, and these are often determined by rule-of-thumb heuristics (e.g. flu, 4-fold-rise for conversion, titre of 1:40 HAI for seropositivity).[?] However, antibody responses vary greatly between individuals for many pathogens; therefore, relying on these heuristics to determine infections in serology studies can lead to incorrect infection status being inferred, potentially leading to biased estimates.[?] Consequently, a better understanding of the kinetics of antibody trajectories post-vaccination and infection can help establish and ascertain the accuracy of existing heuristics and be used to more accurately infer infection status.

1.2 Antibody kinetics

Modeling antibody kinetics involves using mathematical and statistical techniques to simulate the trajectories of antibodies in response to an infection or vaccination. [?] Typically, this involves the use of mathematical equations and statistical methods to describe the time-dependent changes in antibody levels within an individual or a population. This process is essential for understanding how antibody levels evolve and, therefore, potentially provide protection against infectious diseases. Various functional forms have been used to model the individual-level kinetics of antibody trajectories,[?] typically it follows a three-stage process:

- *Initial Response:* The trajectories often start by capturing the initial antibody response to a pathogen or vaccine. This phase is characterized by a rapid increase in antibody levels as the immune system recognizes and mounts a defence against the antigen.
- *Peak Antibody Level:* The trajectories then rise to a peak antibody level, which is the highest concentration of antibodies reached during the immune response. This peak can vary depending on factors like the strength of the immune response and the nature of the antigen.
- *Decay Phase:* After the peak, there is a decline in antibody levels. Antibodies have a finite lifespan in the bloodstream, and their concentration gradually decreases as the pathogen is cleared or the vaccine antigen wanes. Often antibodies secreted from now-establish long-lived plasma cells, the decay phase trajectories fall to a set point titre.[?]

Modeling antibody kinetics provides several important benefits. First, by understanding the rate of antibody decline, models can estimate how long an individual's immunity is likely to last after infection or vaccination. This information is critical for designing vaccination schedules and determining the need for booster shots and permits the creation of better heuristics for determining infections using serology. Kinetics are also useful for optimizing vaccination strategies and can help identify the optimal timing and frequency of booster vaccinations to maintain protective antibody levels within a population. This is especially important for vaccine-preventable diseases with varying immunity levels such as influenza and COVID-19.

1.3 Correlates of protection

A correlate of protection is an immune function or biomarker that correlates with and or may be biologically responsible for protection against infection or disease. Correlates of protection have been established for influenza, Hepatitis A and B, Measles, Polio, Rabies, Yellow fever and more.[?] Correlates of protection help researchers understand the specific immune responses needed to prevent or control an infectious disease. This knowledge is pivotal in designing and optimizing vaccines to induce the required immune response effectively.[?]

Correlates of protection can be established by i) comparing the immune response of those protected by the vaccine with so-called 'breakthrough cases', where clinical disease manifests despite prior vaccination and ii) human challenge studies.[?]

A natural biomarker for a correlate of protection from infection is the amount of neutralising antibodies in the serum as this measures a serum's ability to prevent viral particles from infecting vulnerable cells. Therefore, those with high levels of neutralising antibodies could abort an infection by neutralising viral particles in-host even if exposed to a virus. Determining the immunological profiles of individuals are exposed to an infection but manage to abort it due is crucial for determining a universal correlate of protection. Usually, these individuals are identified through challenge studies, allowing for a direct evaluation of correlates of protection, however these studies are expensive, difficult to run, can only be performed on healthy individuals, and are only possible for pathogens with general pathogenesis.

Using serological samples to establish a correlate of protection would solve the aforementioned problems with challenge studies as they are cheaper to conduct and less invasive. However, determining correlates of protection through serological studies is challenging to establish because those who have been exposed but then experience an abortive infection generally leave no measurable antibody imprint. Therefore serological studies can be augmented with either immunological profiling to determine other immunological biomarkers that indicate abortive infection[?] or include intensive contact tracing with the serological studies to determine exposure rates between individuals[?]. Augmented serological studies with these additional increase the complexity and cost of a study and, therefore, are not feasible in many settings.

1.4 Overview of modelling framework

If antibody trajectories are known for every individual throughout a study and the time of infection is known, then we can determine the probability of infection for a given titre value. Though not strictly a correlate of protection, as we do not know whether non-infected individuals have been exposed, this can be useful for determining associations between infected and non-infected individuals for a given population. However, relying on this sample to establish correlates of protection can lead to bias estimates XXX.

In this document, we present a single modelling framework which takes individual-level serological sample data and uses changes in antibody titres over time to determine i) which individuals have seroconverted throughout the study, ii) subsequent antibody kinetics of these infection individuals, iii) which individuals are exposed throughout the study, and iv) the correlate of protection preventing exposed individuals from becoming infected. Though the infection and exposure status is often unknown for most individuals in a serological study, we find that by using broad, biologically-informed mechanistic forms for the antibody

kinetics and correlation of protection, the infection and exposure status of individuals and the population are recoverable through the interdependencies of mechanisms i)–iv) within a Bayesian probabilistic framework.