This manuscript sets out to test a potential important improvement on SIS models that incorporate co-infection information. This is a significant contribution as it is known that co-infections of multiple pathogens, or multiple genotypes of the same pathogens, can influence transmission, susceptibility, and virulence. This can be particularly important for pathogens such as HPV, where different genotypes can have a wide variety of disease outcomes, and this is a good choice of focus for the manuscript.

The models being developed seem appropriate (see some comments below) and to be an improvement on existing approaches. I think these results are worthy of publication. I particularly found the testing of existing methods followed by the implementation and testing of Approximate Bayesian Computation (ABC) an appropriate and informative approach that illustrates the current shortcomings and the benefit of the ABC modeling well. However, a shortcoming of my review is that, although I am a disease ecologist with experience in many of the concepts of the manuscript, my work relies heavily upon molecular techniques and not this type of modeling. I have experience implementing similar models (ABC with genetic data) but not developing them.

A major area of revision that I recommend prior to submission for publication is the focus of the introduction. The manuscript touches on many relevant points and outlines some clear arguments but these could be better linked and structured. The problem and focus of the manuscript should be introduced earlier. This will allow the reader to follow why the specific arguments and discussion of existing modeling approaches are being outlined.

In addition to the general structure, the introduction needs to be revised for clarity. I recognize this is a pre-print and the main objective is to receive feedback on the study so this is not a major concern for now. The opening sentence "With the advent of next generation sequencing, an increasing number of infections turn out to be coinfections by multiple genotypes" makes it sound like NGS is resulting in an increased number of coinfection, when it means NGS is allowing the discovery of most infections being coinfections.

In other places, it is unclear if multiple infections/coinfections is referring to multiple pathogens or multiple genotypes of the same pathogen. For example, the second paragraph in which they define multiple infections as multiple genotypes but then provide an example of HIV and malaria (i.e., multiple pathogens).

The summary sentence of the introduction (Lines 34-36) needs to be refined for clarity. For example, it is not clear from the rest of the introduction what is meant by "but it is unclear whether these interactions are sufficiently strong to be detected at the population level." Do the authors mean it is unclear if these interactions can be diagnostically detected or if these interactions have biological effects at the population level.

The sections discussing current modeling approaches would benefit from being reworked to have better conclusions of each section- as written, the pitfalls and benefits of each method are not clear. For example, the "parasite combinations" section concludes abruptly (Line 79-80). I recommend explaining why the lack of an explicit epidemiological model is problematic and what important information these current models lack.

The issues I highlight above relate to a general issue with the clarity of the writing throughout (with the exception that the methods section is well written).

If the manuscript is being submitted to a journal in which the methods come after the results (as is the current structure), the results need a bit more information so the reader can follow.

A minor concern with the model that I have is in regards to not considering vaccinations. The fact that natural immunity is low but vaccine immunity is high makes it is unclear why immunization was not included in the model. This would seem to be an important component- specifically as vaccination against specific genotypes could presumably increase frequency of other genotypes.

My major take away is that the analyses are well done and represent a worthwhile contribution to the literature. A careful revision of the text will greatly strengthen the work and is recommended before publication.

A few minor comments:

Line 28: “virus loads also seem to be differ” Should be ‘seem to differ’ or ‘seem to be different’

Line 47-49: For submission to most journals, further explanation is needed as to why a negative binomial distribution is thought to indicate host population structure or a specific functional response. In addition, “a specific functional response” is vague and needs further explanation.

Line 109-110: Revise for clarity “First we use existing methods developed to detect significant associations between parasites from coinfection data.”

Line 228-229: “We do not report it here but the accuracy of the inference varied widely across parameters.” This is interesting and worth reporting in the results.