## Supplementary Methods

### Bayesian models of IFN-I resistance of HIV-1 plasma isolates from longitudinally sampled participants

To create a simple model of the temporal dynamics of type I interferon (IFN-I) resistance, IFNα2 and IFN IC50 values were each modeled using a Bayesian change point hierarchical model. The model is based on a segmented regression of the log IC50 making the following simplifying assumptions:

• Each participant has a level of resistance at the acute infection stage drawn from separate population-level distributions for typical, non- or fast progressors.

• Each participant has a drop (or rise) in IFN-I resistance from acute levels drawn from separate population-level distributions for typical, non- or fast progressors.

• Each participant has a time to nadir drawn from separate population-level distributions for typical, non- or fast progressors.

• Resistance changes linearly from onset of symptoms to time of nadir.

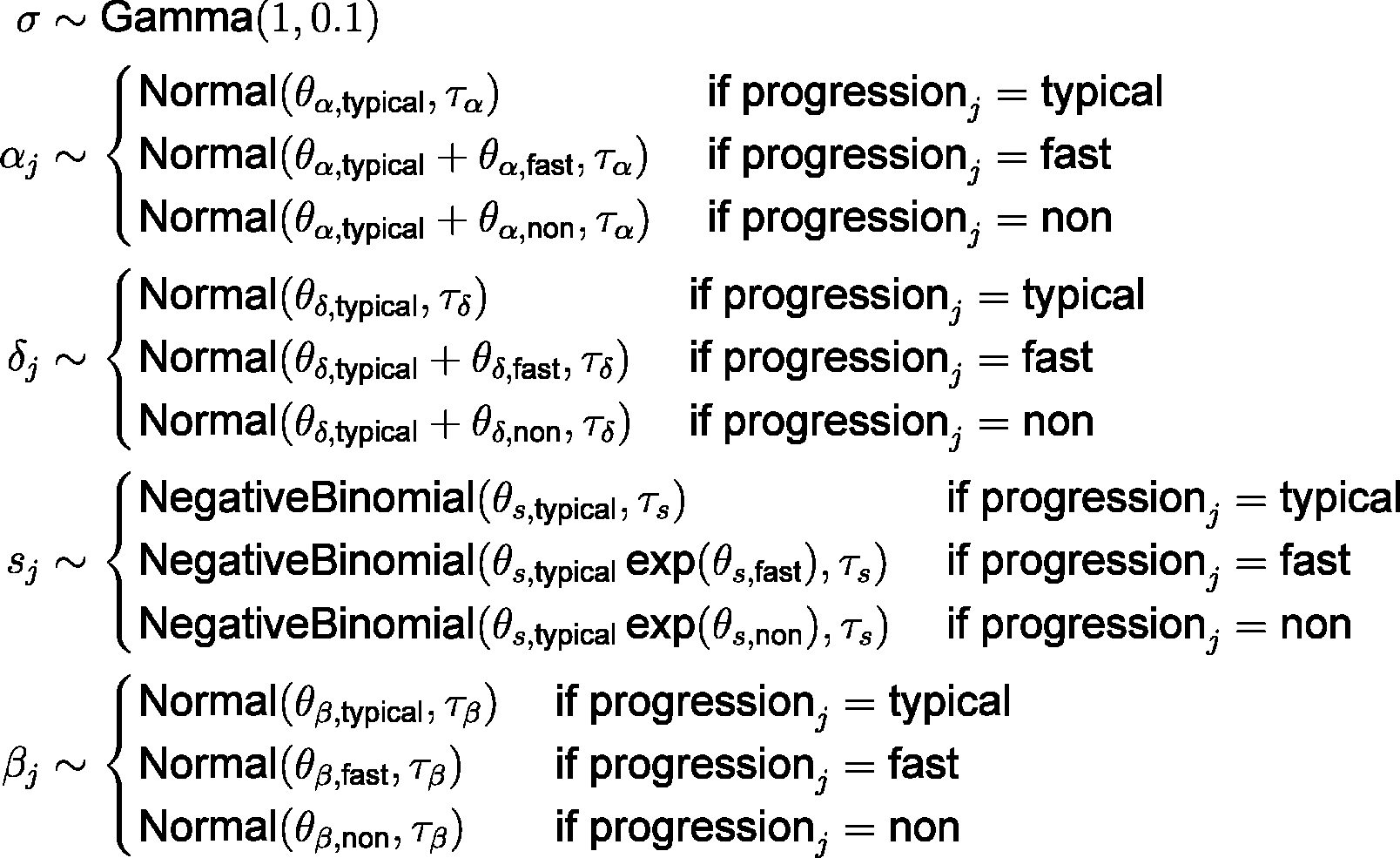
• The nadir of IFN-I resistance represents a changepoint in the data. Following this point, changes in IFN-I resistance are modeled as a linear function of CD4+ T cell count changes away from the count found at nadir.

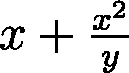
• Where data is not present, CD4+ T cell counts are assumed to be linearly interpolated between adjacent observations.

The log IC50 observation from each viral isolate  was modeled as a normal distribution  with mean  where:

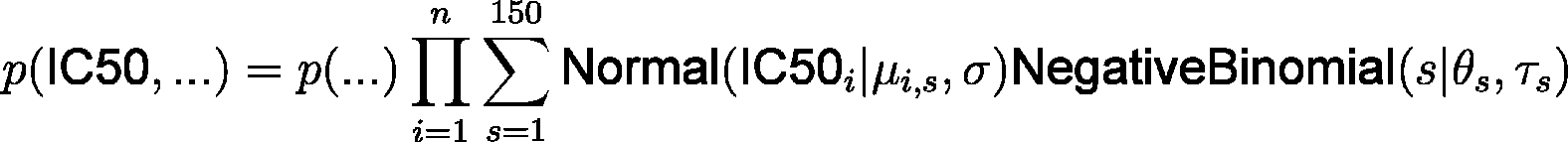


where the parameters  represent the level of IFN-I resistance at symptom onset,  represents the change from symptom onset to nadir and  represents the time of nadir in person . Study participant data is represented by  corresponding to the time since onset of symptoms and  recording the participant from which isolate  was collected,  containing the estimated CD4+ T cell count for person  at time  and  is the disease progression type (fast/non/typical) for participant . The hierarchical probabilities for these parameters were:



where  indicates each participant and  represents a negative binomial distribution parameterized such that the expected value is  and the variance is . All hyperparameters were given prior probabilities of  for parameters representing the means of a distribution and  for parameters representing standard deviations other than  and  which were given a flat prior and .

For computational efficiency, the nadir time parameter  was discretized to weekly intervals, assumed to fall within 1 to 150 weeks after symptom onset and marginalized out of the joint probability:



where  represents all parameters other than  and  is defined the same as :



Posterior probabilities were estimated with 50 Markov chain Monte Carlo chains of 5000 iterations each using Stan (*90*).

### Bayesian models of IFN-I resistance of outgrowth and rebound HIV-1 isolates

To compare the IFN-I resistance of viral isolates derived from plasma samples collected during acute, chronic and rebound infections, as well as from viably frozen PBMCs collected during ART suppression (QVOA), IFNα2 and IFN values were modeled using a Bayesian hierarchical model. The model is based on the assumptions that:

• Isolates found at acute infection form a base level of IFN-I resistance for a given person. Resistances in virus isolated from chronic, ART suppressed and rebound infection for this person are modelled as changes from this initial level.

• The mean IC50 level within each person for acute isolates and the change from acute levels for chronic, QVOA and rebound isolates are drawn from a population-level distribution for that type.

• QVOA isolates are separated into to two populations; a “pre” group composed of QVOA viruses isolated from study participants prior to or in the absence of treatment interruption (ATI) and a “post” group of QVOA viruses isolated from participants following ATI and reinitiation of ART.

• In both QVOA populations, the viruses can include some proportion of rebound-like isolates. This mixture is modeled in both pre- and post-treatment so that differences in mixture proportion between the two populations can be assessed.

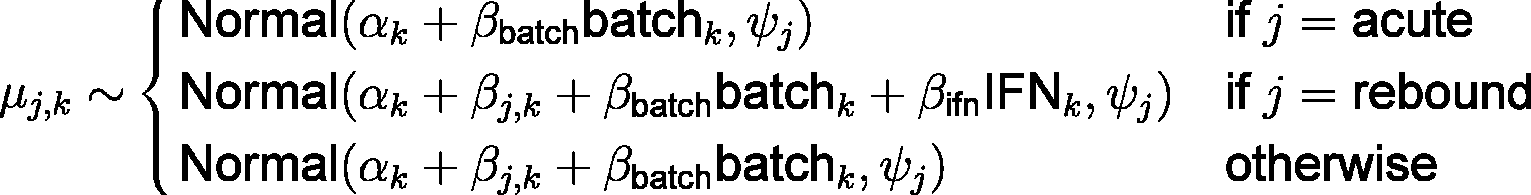
• Variation in the potency of INF-I used to experimentally determine IC50 values may shift the inferred resistance for isolates tested in other studies. This effect is modeled as a multiplicative shift in IC50 for all isolates measured outside this study (acute recipient and chronic donor isolates from ref. 50).

• Isolates from participants who received exogenous IFNα2 during treatment interruption may display altered interferon resistance. This effect is modeled as a multiplicative shift in IC50 for all rebound isolates from such participants (participants 004, 030, and 044 from ref. 38).

The log IC50 observation from each viral isolate  from acute, chronic and rebound isolates was modeled as a normal distribution:

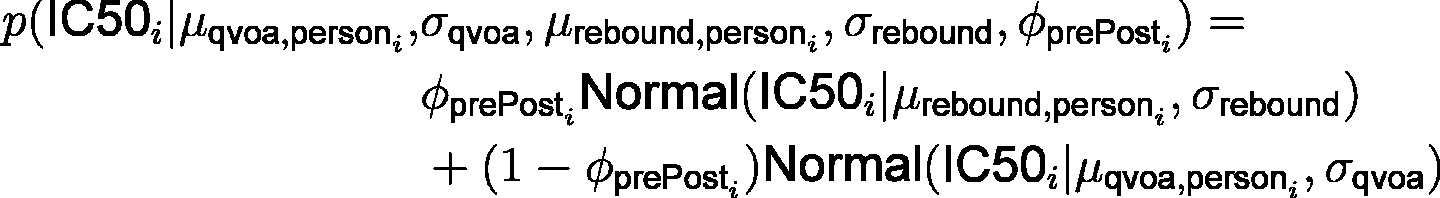


with the mean resistance for isolate type  from person :



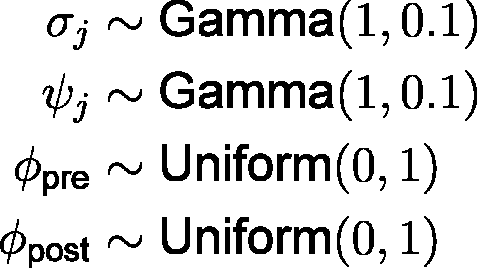
where  indicates whether isolate  was isolated during acute, chronic, QVOA or rebound infection from participant ,  indicates when isolates from person  were tested in another study and  indicates when person  was treated with exogenous IFN prior to and during treatment interruption. Parameters are included for the mean resistance level during acute infection for each person , standard deviation of isolates of type  within a person , standard deviation of mean resistance for type  isolates among people , change from acute levels in isolates of type  in a given participant , the effects of exogenous IFN treatment  and batch to batch variation in IFN in isolates assayed in previous studies .

For QVOA isolates, the IC50 was modeled as a mixture of two populations such that:



where  indicates whether isolate  was isolated pre- or post-ATI and and  and  represent the proportion of rebound-like virus present in pre- and post-ATI QVOA isolates.

The hierarchical parameter priors were modeled as:



where  indicates the isolate type (acute, chronic, QVOA, rebound). All ,  were given flat priors and  and .

Posterior probabilities were estimated with 50 Markov chain Monte Carlo chains of 5,000 iterations each using Stan (90).