

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

Discrimination of Primary and Chronic CMV Infection based on Humoral Immune Profiles in Pregnancy

Andrew P. Hederman¹, Christopher J. Remmel¹, Shilpee Sharma², Harini Natarajan³, Joshua A. Weiner¹, Daniel Wrapp⁴, Catherine Donner⁵, Marie-Luce Delforge⁶, Piera d'Angelo⁷, Milena Furione⁷, Chiara Fornara^{7,8}, Jason S. McLellan⁴, Daniele Lilleri⁷, Arnaud Marchant², Margaret E. Ackerman^{1,3}#

¹Thayer School of Engineering, Dartmouth College, Hanover, NH, USA

²European Plotkin Institute for Vaccinology, Université libre de Bruxelles, Brussels, Belgium

³Department of Microbiology and Immunology, Geisel School of Medicine, Hanover, NH, USA

⁴Department of Molecular Biosciences, The University of Texas, Austin, TX, USA

⁵Université Libre de Bruxelles (ULB), Hôpital Universitaire de Bruxelles (H.U.B.), CUB Hôpital Erasme, Department of Obstetrics and Gynecology, Brussels, Belgium

⁶Université Libre de Bruxelles (ULB), Hôpital Universitaire de Bruxelles (H.U.B.), CUB Hôpital Erasme, National Reference Center for Congenital Infections, Brussels, Belgium

⁷Microbiology and Virology, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy

⁸Current affiliation: General Clinical Laboratory, with specialist areas of Clinical Pathology, Microbiology and Virology, Istituti Clinici Scientifici Maugeri IRCCS, 27100 Pavia, Italy

Corresponding Author

Margaret E. Ackerman

14 Engineering Drive

Hanover, NH 03755

margaret.e.ackerman@dartmouth.edu

(ph) 603 646 9922

(fax) 603 646 3856

37 **Abstract**

38 Most humans have been infected by Cytomegalovirus (CMV) by the time they reach forty years of
39 age. Whereas most of these CMV infections are well controlled by the immune system, congenital
40 infection can lead to serious health effects and death for the fetus and neonate. With clear evidence that
41 risk to the fetus is lower during chronic maternal infection, and varies in association with gestational age
42 at the time of primary maternal infection, further research on humoral immune responses to primary CMV
43 infection during pregnancy is needed. Here, systems serology tools were applied to characterize antibody
44 responses to CMV infection among pregnant and non-pregnant women experiencing either primary or
45 chronic infection. Whereas strikingly different antibody profiles were observed depending on infection
46 status, more limited differences were associated with pregnancy status. Beyond known differences in IgM
47 responses that are used clinically for identification of primary infection, distinctions observed in IgA and
48 FcγR binding antibody responses and among viral antigen specificities accurately predicted infection
49 status in a cross-sectional cohort. Leveraging machine learning, longitudinal samples were also used to
50 define an immunological clock of CMV infection from primary to chronic states and predict time since
51 primary infection with high accuracy. Humoral responses diverged over time in an antigen-specific
52 manner, with IgG3 responses toward tegument decreasing over time as is typical of viral infections, while
53 those directed to pentamer and glycoprotein B were lower during acute and greatest during chronic
54 infection. In sum, this work provides new insights into the antibody response associated with CMV
55 infection status in the context of pregnancy, revealing aspects of humoral immunity that have the potential
56 to improve CMV diagnostics and to support clinical trials of interventions to reduce mother-to-fetus
57 transmission of CMV.

58 **Introduction:**

59 A member of the herpesvirus family, human cytomegalovirus (CMV) commonly manifests as a
60 mild or asymptomatic infection. However, in very young and immunocompromised individuals, CMV can
61 cause severe disease; it is the leading cause of congenital infection among newborns¹. In the United
62 States alone it is estimated that 40,000 children are born with congenital CMV (cCMV) infection every
63 year¹. This high burden is still likely an underestimate since many cases are asymptomatic at birth. Many
64 of the diagnosed infections are severe, leading to an estimated 400 deaths and an additional 8,000 cases
65 presenting with permanent disabilities, including speech and language impairment, hearing loss, mental
66 disability, cerebral palsy, and vision impairment, annually^{1,2}. Fetal infection results from intrauterine
67 transmission and is most likely to occur when a mother experiences primary CMV infection during
68 pregnancy^{3,4}. The difference in fetal infection risk between primary and reactivated maternal CMV
69 infection is striking, with approximately one third of primary infections leading to CMV infection of the fetus
70 as compared to < 3.5% estimated to result from CMV reactivation or superinfection⁵⁻⁹. New insights into
71 CMV infection during pregnancy that could contribute to identification of pregnancies at greatest risk, and
72 efficient testing of new therapeutic interventions are urgently needed¹⁰.

73 Pregnant women present a population of subjects who are not immunodeficient, but in a unique
74 state of immune regulation in order to ensure fetal tolerance, and can transmit CMV to their fetus during
75 pregnancy. While the differing risk of transmission associated with maternal seropositivity provides strong
76 evidence that pre-conception immunity plays a protective role, studies evaluating the clinical potential of
77 CMV-hyperimmune globulin to improve neonatal outcomes have yielded mixed results¹¹⁻¹³. It has been
78 speculated that these differing outcomes may relate to gestational age and the timing of the intervention
79 following maternal infection and therefore be impacted by both the reliability of the diagnostic approach
80 and strictness in the definition of primary infection cases. Because the suitability of these approaches for
81 application during pregnancy can be supported by comparison of responses in pregnant with non-
82 pregnant individuals during primary infection, developing a deeper understanding of humoral immune
83 profiles in this unique immune state may have implications for the clinical development of diverse small
84 molecule and biologic antiviral interventions.

85 Short of these goals, given the widely diverging congenital infection risks associated with primary
86 and chronic infection, confident discrimination between these two states is crucial for identifying newborns
87 with highest risk of cCMV infection. Complementing virological assessment, discrimination of these two
88 states using serology may allow antiviral treatment in pregnant women and hearing and learning
89 interventions in newborns to be initiated early^{14,15}. Current diagnostic assays for pregnant women are
90 performed using serology measurements for IgM and IgG avidity, while a PCR based urine test, often
91 following a saliva PCR screening test, is the standard for newborns^{14,16-18}. However, although CMV is a
92 common infection, there is no universal standard diagnostic assay across countries and healthcare
93 centers¹⁹, introducing variability in care and posing challenges to clinical trial design for testing of novel
94 interventions. Additionally, many countries, including the United States, do not routinely screen for CMV

95 in pregnant women, whereas some European countries do perform routine screening, adding another
96 layer of complexity into clinical data regarding the impact of maternal CMV status, antibody responses,
97 and timing of infection as assessed by differing diagnostic measures and across populations. Irrespective
98 of these details, following a primary CMV diagnosis during pregnancy, clinicians typically counsel
99 pregnant women about the future risks to their fetus.

100 The dichotomy in fetal risk profiles of primary and non-primary CMV infection has important
101 implications in transmission risk for pregnant women. Beyond this classification, the specific timing of
102 CMV infection is also a crucial factor in potential of severe congenital disease, with a primary infection in
103 the first term of pregnancy leading to higher likelihood of severe cCMV infections as compared to a
104 primary infection in the second or third trimester²⁰. Further insights into relative risk profiles are
105 challenged by imperfect precision in defining and estimating the timing of primary infection, particularly
106 during pregnancy, which has been associated with varying and dynamic immunological state changes²¹.
107 As a result, more accurate and confident diagnosis of primary and non-primary CMV infection would
108 benefit the medical community and affected birthing parents by providing clearer and more definitive
109 insights into their CMV infection status and associated risk to the fetus.

110 However, previous work in a variety of infectious disease settings has shown that antibody
111 responses evolve over time, exhibiting complex patterns in response magnitude and characteristics²²⁻²⁴,
112 and that they can also differ in association with pregnancy²⁴⁻²⁶. Pregnant women with primary CMV
113 infection represent a unique intersection of these two complex antibody response scenarios with
114 established clinical significance. Here, leveraging machine learning and highly multiplexed assays to
115 capture a wide range of antibody response attributes over time, we evaluate how responses to primary or
116 non-primary CMV infection during pregnancy and over time may vary, providing insight into the role
117 pregnancy may play in modifying these responses and new means of predicting and reducing risk of
118 cCMV infection for affected pregnancies.

119

120 **Results**

121 *Antibody Profiles Distinguish Primary from Chronic CMV Infection*

122 Antibody responses were profiled among both pregnant and non-pregnant individuals with either
123 primary or chronic CMV infection (**Table 1**), with primary infection strictly defined by CMV-specific IgG
124 seroconversion, CMV-specific IgM antibody detection, low IgG avidity index, and/or CMV DNAemia.
125 Conversely, chronic infection was defined by seropositivity in the absence of these diagnostic measures.
126 To more comprehensively understand how antibody profiles vary among pregnant primary (n=114), non-
127 pregnant primary (n=29), pregnant chronic (n=30), non-pregnant chronic (n=30) CMV-infected individuals,
128 in addition to naïve subjects (n=9), antibodies specific for CMV tegument protein²⁷ (CG1 and CG2),
129 glycoprotein B (gB), and the pentamer complex were characterized for isotype, subclass, and Fc
130 Receptor (FcR) binding capacity (**Supplemental Figure 1, Supplemental Table 1**). We profiled the
131 antibody response examining a diverse set of antibody features and CMV antigens, including gB and

132 pentamer complexes from several different sources, with the aim of gaining a more holistic view of the
133 humoral immune response to primary or chronic CMV infection. Glycoproteins tested included several
134 that are commercially available (Sino (Sino Biologicals), and NA (Native Antigen)), as well as others that
135 have been characterized in structural studies conducted with an eye toward vaccine development
136 (GSK)^{28,29}, and by academic groups (UT)^{30,31}, including a novel modified form of gB with engineered
137 proline mutations (3p).

138 To define differences in antibody profiles among subject groups that go beyond the measures
139 typically used in clinical diagnosis, we performed UMAP analysis on CMV-specific antibody features
140 excluding IgM (**Figure 1A**). This unsupervised analysis revealed that the main aspects of differentiation
141 among subjects related to infection status; strong distinctions in the antibody response were observed
142 between individuals with primary as opposed to chronic infection. In contrast, limited differences in the
143 global response profiles were observed between pregnant and non-pregnant individuals. While univariate
144 analysis between pregnant and non-pregnant individuals did reveal some nominally statistically significant
145 differences across or within primary and chronic infection groups (**Supplemental Figure 2**), we could not
146 exclude the possibility that these differences resulted from a difference in timing of diagnosis or were
147 impacted by differences in age between cohorts. The former may be likely, as similarly subtle differences
148 were observed between primary subjects in these groups and another cohort (Erasmus Hospital) of
149 pregnant women with primary CMV infection (**Supplemental Figure 3**). Given these observations and the
150 lack of robust differences associated with pregnancy status, this variable was not considered in
151 subsequent analyses.

152 We next investigated which individual features of the immune response might contribute to
153 clustering of subjects in primary and chronic infection groups. Visualizing the degree and confidence of
154 differences in antibody response features between primary and chronic infection status groups revealed
155 aspects of the humoral response that consistently differed (**Figure 1B-C**). Distinct levels of CMV-specific
156 IgA, IgM, and FcR binding antibodies were observed between primary and chronic infection groups
157 across multiple CMV antigens. IgM responses were elevated across all antigens for the primary infection
158 group, which was expected as IgM titers were used to define infection status. Interestingly, total CMV-
159 specific IgA and both IgA1 and IgA2 subclasses were elevated in primary infection subjects. Somewhat
160 surprisingly, total IgG had very modest differences between primary and chronic infection groups.

161 Given the striking differences among groups we next wanted to further examine individual
162 responses for IgM, IgA, and IgG across the panel of CMV antigens. As expected, IgM responses were
163 strongly and significantly elevated among subjects with primary infection, as were IgA responses across
164 diverse antigen specificities (**Figure 1C**). Interestingly, and in contrast with total IgG response levels
165 (**Figure 1C**), which exhibited statistically significant but relatively small differences in response
166 magnitude, FcγR-binding antibodies were both significantly and strongly elevated among subjects with
167 chronic CMV infection (**Figure 1B, Supplemental Figure 4**), suggesting the presence of qualitative
168 differences in the antibodies present during primary and chronic infection.

169

170 *IgG Subclasses and CMV Antigens Display Distinct Profiles in Primary and Chronic Infection*

171 The observation that Fc γ R binding antibodies but not total IgG were reliably elevated in chronic
172 subjects was intriguing, and suggested that differences in induction of IgG subclasses, which differ
173 dramatically in their FcR binding capacity, may exist. However, with the exception of IgG2 responses to
174 tegument, IgG2 and IgG4 responses to CMV antigens were uncommon, and therefore differences in the
175 levels of these more functionally inert subclasses were not observed (**Supplemental Figure 5**). Humoral
176 responses to viral infections are generally dominated by IgG1 and IgG3 antibodies, which both bind well
177 to activating Fc γ R and have the potential to elicit the potent antiviral activities of the complement cascade
178 and innate immune effector cells. While statistically significant differences were observed for a subset of
179 antigens tested, IgG1 responses, which are typically dominant following acute viral infections^{32,33}, were
180 generally similar in primary and chronic infection (**Figure 2A**). Only one of each of the gB, pentamer, and
181 tegument proteins tested showed a statistically significant difference, and each of these demonstrated
182 only a small increase among the chronic infection group. In contrast, IgG3 responses were more
183 uniformly and strikingly distinct between groups. Perhaps most surprisingly, the direction of these
184 differences varied by antigen-specificity. Antibody responses to tegument proteins exhibited elevated
185 levels in primary infection, but elevated levels of gB and pentamer-specific IgG3 were observed among
186 the chronic infection group (**Figure 2A**). While all pentamer complexes tested showed this profile,
187 responses to recombinant gB proteins were more variable, with two gB preparations showing elevated
188 levels among chronic subjects, and one preparation showing elevated levels in primary infection. While
189 some distinctions in IgG subclass composition in the context of primary and latent herpesvirus infections
190 have been reported³⁴⁻³⁸, how well they may support discrimination between primary and chronic CMV
191 infection is not known. In sum, while IgG1 responses tended to either persist or increase between primary
192 and latent infection classes, IgG3 responses differed dramatically by antigen-specificity; responses to
193 tegument were higher among subjects with primary infection, but responses to gB and pentamer were
194 instead elevated in chronic infection.

195 While UT pentamer-specific IgG1 and IgG3 levels were correlated with each other among
196 subjects with primary infection ($R_p=0.33$, $p<0.0001$), they were not well correlated among individuals with
197 chronic infection ($R_p=0.20$, $p=0.08$). This difference in degree of correlation between IgG1 and IgG3
198 responses between infection status states was consistent across pentamer proteins evaluated.
199 Additionally, the strength of correlation observed between these IgG subclasses and the Fc γ RIIIaV-
200 binding capacity of pentamer-specific antibodies changed over time, with stronger correlations between
201 IgG1 and Fc γ RIIIaV observed in primary infection, but IgG3 and Fc γ RIIIaV binding in chronic infection
202 (**Figure 2C**). These patterns were consistent across Fc γ Rs and pentamer proteins tested (**Figure 2D**),
203 suggesting that the pool of antibodies most capable of eliciting Fc γ R-dependent effector functions
204 changes in composition over the course of infection. These temporal differences may have important
205 implications to viral pathogenesis and host defense.

206

207 *Prediction of CMV infection status using machine learning*

208 Next, because unsupervised analysis of antibody profiles revealed clear distinctions between
209 primary and chronic CMV infection, we applied supervised machine learning to explore the ability of a
210 model to accurately discriminate primary and chronic CMV infection (**Figure 3A**). We employed a logistic
211 regression framework with regularization to classify primary or chronic CMV infection based on antibody
212 profiling data while reducing the risk of overfitting associated with high-dimensional data. Again, IgM
213 responses were excluded because they were used in clinical class assignment of the study groups. The
214 model was trained on 80% of primary and chronic subject profiles while the remaining 20% was used for
215 testing. Furthermore, repeated five-fold cross validation was employed so each subject would be part of
216 the test set and representative accuracy across different folds could be defined.

217 Model predictions were highly accurate; across 100 repeated five-fold cross validation runs, the
218 median model accuracy for predicting primary or chronic CMV infection was 94% (**Figure 3B**). In contrast,
219 prediction results were essentially random when training and testing was performed after permutation of
220 class labels, which serves as a means to assess model robustness by measuring the potential for
221 overfitting. Misclassifications were not biased toward one or the other class, as shown for the cross
222 validation run repeat presenting median accuracy as a representative confusion matrix (**Figure 3C**).
223 Classification calls for actual but not permuted class labels were both typically correct and assigned to
224 respective classes with high probability (**Figure 3D**). The relatively few incorrect classifications were
225 typically close to the decision boundary. Given this evidence of model accuracy and robustness, a final
226 model was trained on all subjects in the discovery cohort. When applied to an independent set of 57
227 primary infection and 40 chronic infection samples serving as a validation cohort, this model resulted in
228 similarly high confidence and perfect classification accuracy (**Figure 3D**). Given this validation, the top
229 features employed in the final model were examined (**Figure 3E**). The features with largest positive
230 coefficients, which serve to identify primary subjects, were primarily IgA-related antibody responses
231 directed to tegument antigens (4/5 features). Conversely, features with large negative coefficients, useful
232 to identify chronic infection, were typically related to IgG3 and Fc γ RIII-binding capacity of response to gB
233 and pentamer (4/5 features). Collectively, these modeling results point to novel antibody response
234 attributes, distinct from traditionally considered parameters such as IgM and IgG avidity, as being
235 excellent candidate markers for distinguishing primary and chronic infection status.

236

237 *Modeling Longitudinal Responses to CMV Infection Reveals a Molecular Clock of Antibody Responses*

238 The excellent discrimination of primary and chronic CMV infection in the validation cohort led us
239 to next explore how class predictions related to longitudinal development of humoral immune responses
240 to infection. To this end, subsequent samples were available for the validation cohort over a series of up
241 to four visits that extended out to half a year after initial sampling or infection onset. The model trained on
242 the initial cohort was used to predict class for longitudinal samples from the validation cohort (**Figure 4A**).

243 Whereas chronic subjects were consistently classified as such with similarly high confidence at all
244 subsequent visits, subjects with primary infection at the initial timepoint became less confidently classified
245 as primary at subsequent visits (**Figure 4B**). By visit four, the majority of samples from subjects with
246 primary infection at their initial visit exhibited class probabilities below the lowest observed at the initial
247 sampling. However, because visits were not consistently spaced in time between subjects, these
248 longitudinal profiles can be more meaningfully compared over time post symptom onset (**Figure 4C**).
249 Despite the imperfect reliability of projected timing of infection, the classification model probabilities
250 presented a clear relationship with time. Samples fell below the midpoint on the classification scale as
251 early as 90 days post infection onset, but failed to reach values typically observed among chronic
252 subjects even out at 250 days post infection, suggesting a relatively prolonged transition to reaching the
253 chronic state profile. As expected, the individual antibody response features making the greatest
254 contributions to the classification model also demonstrated clear changes among subjects with primary,
255 but not chronic infection, over time (**Figure 4D**).

256 To this point, machine learning models have only been concerned with making predictions on the
257 probability of a sample belonging to either the primary or chronic infection class. However, given the clear
258 ability for these binary classification models to provide insight into time since infection, we next sought to
259 evaluate models explicitly trained for this specific purpose. For this purpose, longitudinal profiles of the
260 primary infection cases across visits were used to train a model to predict time since symptom onset as a
261 continuous variable (**Figure 4E**). The resulting linear regression model, which showed good robustness in
262 the context of five-fold cross-validation (**Supplemental Figure 6A**), was then used to predict days post
263 infection for the cross-sectional primary samples. This model, which relied primarily on IgG3 features
264 (**Supplemental Figure 6B**) showing strong time-dependence (**Supplemental Figure 6C**), was then
265 applied to the cross-sectional cohort. Predictions of time since symptom onset in the unseen validation
266 cohort exhibited excellent accuracy (**Figure 4F**). Overall, this analysis demonstrated that antibody profiles
267 in CMV-infected individuals exhibit generalizable temporal patterns in their dynamic antibody responses
268 during primary infection.

269

270 Discussion

271 Presently, discrimination of primary and non-primary CMV infection in the context of pregnancy is
272 used in counseling as to the risk of congenital CMV infection based on the lower risk of congenital
273 infection associated with non-primary infection and the prescription of antiviral therapies. Beyond this
274 value, in the absence of an effective vaccine, a deeper understanding of how immune responses differ in
275 association with transmission risk has the potential to contribute to the development of new interventions.
276 Here, high dimensional antibody profiles beyond IgM levels and IgG avidity were developed from a
277 commonly used multiplex assay format and supervised and unsupervised machine learning was used to
278 differentiate primary and chronic infection and pregnancy status.

279 Whereas limited or no differences in humoral responses were associated with pregnancy status,
280 our study showed clear distinctions in antibody profiles between primary and chronic infection cases. Both
281 composite profiles individual antibody features related to antigen-specificity and immunoglobulin isotype,
282 subclass, and binding to Fc receptors demonstrated these distinctions. Importantly, these differences
283 extended beyond those previously known to exist and which are presently applied to support clinical
284 diagnosis. The number and longitudinal profiles of the features that distinguish infection history suggested
285 that a sort of humoral clock could be defined in order to time the onset of primary CMV infection. Indeed,
286 supervised machine learning models that reliably captured how antibody responses to CMV infection
287 varied over time were learned and validated on independent samples.

288 Beyond potential clinical utility, the defining features of how responses to primary CMV infection
289 transition over time is relevant to understanding the evolution of the immune response to this member of
290 the notoriously immune-evasive Herpesviridae family^{39,40}. Current diagnostic methods rely on
291 detection/levels of IgM and of IgG avidity. The presence of IgM alone is insufficient to diagnose a primary
292 CMV infection: poor correlation between commercial tests have been reported⁴¹, and false positive
293 primary status calls can result from both persistence of IgM as well as boosting in response to
294 reactivation^{42,43}. Likewise, a positive test for CMV-specific IgG indicates an individual is positive for CMV
295 but provides little information into the time since infection^{14,44}. However, the combination of IgM positivity
296 and low IgG avidity is generally considered to be a reliable indicator of primary CMV infection, though
297 interpretation of IgG avidity tests are confounded by low levels of CMV-specific IgG, intermediate
298 responses raise classification issues, and these indicators are supported by clinical studies of small size.
299 Coupled to the lack of an international standard serum panel of samples from primary infection, the
300 difficulty inherent to establishing a uniform, accurate, and robust diagnostic method is clear⁴⁵. Our data
301 demonstrates that there are other changes in the secreted antibody repertoire that reliably occur over
302 consistent time periods during primary infection.

303 Prior studies of IgG subclasses of CMV-specific antibodies have consistently reported the IgG1
304 and IgG3 bias typical of anti-viral antibodies but noted that subclass responses and total IgG titers can be
305 discordant^{35,46}. Further, they have attributed relatively greater neutralization potency to this minor portion
306 to IgG3³⁶. To that end, we were surprised to observe increasing levels of IgG3 specific for viral
307 glycoproteins over time. Whereas IgG3 responses are usually associated with acute infections and
308 typically wane over time^{33,47-49}, they were observed to increase across both cross-sectional and
309 longitudinal cohorts, but to do so in an antigen-specific manner. Namely, whereas IgG3 responses to
310 tegument proteins showed the typical pattern and decreased while IgG1 responses were stable or
311 increased, IgG3 responses to pentamer increased while IgG1 responses remained largely stable.
312 Memory B cells (MBC) specific for gB are known to exhibit phenotypic states distinct from those specific
313 for tegument antigens⁵⁰. The reduced frequency of MBCs with effector potential specific for glycoprotein
314 as compared to tegument observed in this prior study may relate to the altered kinetics of IgG subclasses
315 observed here. While the mechanistic underpinning and the biological implications of this unusual pattern

316 have yet to be determined, CMV expresses two Fc binding proteins, gp34 and gp68⁵¹. Unlike the viral
317 FcR of HSV, these proteins have been reported to bind to all human IgG subclasses^{52,53}, though affinities
318 have not been reported. Additionally, how these two surface proteins contribute to viral pathogenesis is
319 incompletely understood⁵⁴.

320 Limitations of this study include the fact that the majority of samples were sourced from a single
321 geographic region, leaving open the possibility that differences associated with host genetics and
322 environmental and infectious disease exposure history impact these observations. While relationships
323 between maternal immune responses and transmission of cCMV are of exceptionally high clinical
324 relevance, this study focused on CMV infection history. However, while risk is considerably greater during
325 primary infection, there are certainly differences in virologic, innate immune, and other factors that
326 contribute to these differing risk profiles⁵⁵⁻⁵⁸. The influence of maternal antibody responses is unclear,
327 particularly given the conflicting results in studies of passive antibody therapy in the context of primary
328 maternal infection^{59,60}. While the dosing and frequency of hyperimmune globulin also differed in these
329 studies, a positive effect was only observed in women with very early infection, pointing to the potential
330 importance of accuracy in the dating of infection recency. Recent studies have started to explore the role
331 that antibody responses play in CMV transmission in a rhesus macaque model^{61,62}. However, data in
332 humans is limited and often confounded by the differing risk of infection associated with primary as
333 compared to non-primary infection⁶³. Additionally, the panel of antigens tested in our investigation was not
334 exhaustive, and the reactivity patterns sometimes differed in association with different sequences,
335 conformational states, expression hosts, and other factors, the contributions and importance of which
336 have yet to be defined. These and other factors may represent worthwhile directions for a future study.

337 Overall, many open questions regarding the role of humoral immunity in the context of CMV
338 infection, transmission, latency, and reactivation remain. Higher resolution and more comprehensive
339 analysis of antibody responses using systems serology approaches has the potential to improve our
340 understanding of the complex virus-host interactions at play. Here, by analyzing highly curated cohorts,
341 we report and validate phenotypic signatures of gB, pentamer, and tegument-specific antibody responses
342 that not only robustly classify primary infection status, but also provide insights into time of infection. It
343 remains to be seen if the atypical dynamic profile of IgG3 responses to envelope glycoproteins elicited by
344 CMV is antigen-intrinsic and might be recapitulated when these antigens are delivered by other means, or
345 if it may represent an evasion strategy dependent on other viral genes or aspects of the innate response
346 to viral infection. In the meantime, this work stands to define novel hallmarks of primary CMV infection
347 and time of infection that may present new opportunities to streamline primary infection diagnosis,
348 potentially impacting current clinical practice and enrolment of pregnant women with primary infection in
349 interventional trials, thereby providing new insights into relative cCMV risk and management strategies.

350

351 **Methods**

352 *Clinical Samples*

353 Serum samples were gathered from subject cohort groups of pregnant primary infection, pregnant
354 latent infection, non-pregnant primary infection, non-pregnant latent infection, as well as CMV negative
355 patient cohort as a negative control group (**Table 1**). Human subjects were recruited from Fondazione
356 IRCCS Policlinico San Matteo, Pavia, Italy, and included healthy, primary, and latent CMV infected
357 subjects, as well as pregnant and non-pregnant subjects. An additional cohort of pregnant women with
358 primary infection was recruited from Erasme Hospital. Diagnosis of primary HCMV infection was based on
359 two or more of the following criteria: CMV-specific IgG seroconversion, CMV-specific IgM antibody
360 detection, low IgG avidity index and CMV DNAemia. Chronic infection was defined by the presence of
361 CMV-specific IgG and absence of CMV-specific IgM antibody, and no detection of CMV DNA in blood,
362 saliva, urine and genital secretion during pregnancy. In 89/114 (78%) pregnant women and 28/29 (95%)
363 non-pregnant subjects, onset of primary infection was defined by the appearance of symptoms, while in
364 19/114 (17%) asymptomatic pregnant women onset of infection was estimated on seroconversion (i.e. in
365 the mid between the last IgG negative and the first IgG positive result), occurring within a ≤ 6 -week
366 interval. Finally, in six asymptomatic pregnant women (5%) and one (3%) non-pregnant subject, onset of
367 infection was estimated on the kinetics of CMV-specific IgM and IgG avidity index. For a subgroup of 40
368 pregnant women and 28 non-pregnant subjects with primary infection, 2-4 sequential serum samples
369 collected until 6 months after onset of infection were available. Samples from pregnant women with
370 chronic infection were collected at 10, 20, and 30 weeks of gestation and at delivery. Sample collection
371 was approved by Fondazione IRCCS Policlinico San Matteo, and each participant gave written informed
372 consent.

373

374 *Antibody Profiling Experiments*

375 Fc receptors were expressed as previously described⁶⁴. A complete list of Fc detection reagents
376 and antigens used is provided in **Supplemental Table 1**. Pentamer and glycoprotein B were sourced
377 both commercially, through industry partners, and expressed in-house. For the latter, HCMV Pentamer
378 (UT, Towne strain) was produced by mixing plasmids encoding for residues 24 to 718 of gH with a C-
379 terminal 6x HisTag, residues 31 to 278 of gL, residues 21 to 171 of UL128, residues 26 to 214 of UL130,
380 and residues 19 to 129 of UL131A, all with artificial signal sequences, at an equimolar ratio³¹. This
381 mixture was then used to transiently transfect FreeStyle293-F cells via polyethylenimine. Plasmids
382 encoding for postfusion HCMV gB (AD169 strain) residues 32–692 with an artificial N-terminal signal
383 sequence and a C-terminal HRV3C protease cleavage site, 8xHisTag, and a TwinStrep tag. One
384 postfusion gB construct (UT, or JSM-1074) contained substitutions Y155G, I156H, Y157R, Y206H,
385 S238N, W240T, L241T, Y242H and C246S, which have been reported previously³⁰ and the other
386 construct (3p, or JSM-956) contained substitutions Y155G, I156H, Y157R, Y206H, W240A, L241T,
387 Y242H, C246S, R456A, R459G, M472P, R491P, G492P and a C-terminal T4 fibrin trimerization motif.
388 Plasmids were transiently transfected into FreeStyle 293-F cells via polyethylenimine. Pentamer was
389 purified using NiNTA resin (Thermo Scientific) and postfusion gB proteins were purified using Strep-Tactin

390 affinity resin (IBA Lifesciences). Affinity-purified proteins were further purified by size-exclusion
391 chromatography using a Superose6 10/300 column (Cytiva) in a buffer composed of 2 mM Tris pH 8.0,
392 200 mM NaCl, 0.02% NaN₃.

393 Characterization of antibody profiles was performed using the Fc array assay^{65,66}. Antigens were
394 covalently coupled to magnetic microspheres (Luminex Corporation) using carbodiimide chemistry.
395 Serum dilutions used in assays ranged from 1:250-1:5000 based on initial pilot experiments and previous
396 experience. Detection of antigen specific antibodies was done using R-phycoerythrin-conjugated
397 secondary reagents specific to human immunoglobulin isotypes and subclasses and by Fc receptor
398 tetramers. Median fluorescent intensity data was acquired on a FlexMap 3D array reader (Luminex
399 Corporation). Samples were run in technical duplicates and results were averaged.

400

401 *Classification of CMV infection status*

402 A binomial logistic regression model with least absolute shrinkage and selection operator (LASSO)
403 regularization was used to prediction infection status. Model training was performed using the scikit-learn
404 (version 1.3) in Python (version 3.9) with default options. The regularization parameter was chosen using
405 the option that gave the lowest classification error. Model accuracy was determined by the test set label
406 predictions compared with true labels. Accuracy was assessed over 100 repetitions of five-fold cross
407 validation. Permutation testing was done to measure model robustness by doing the same procedure as
408 described but on randomly shuffled data. Feature importance was determined from a final model that
409 included all subjects.

410

411 *Prediction of time since infection*

412 The same machine learning model as described for predicting CMV infection status was used, this time
413 minimizing mean squared error in time since infection. The model was trained on cross sectional data and
414 tested on all longitudinal samples.

415

416 **Statistics**

417 Statistical analysis was performed in GraphPad Prism (version 9.7). UMAP plots were generated in
418 Python (version 3.9) using the umap-learn package (version 0.4)⁶⁷ and then plotted using Prism. Volcano
419 plots were generated in R (version 4.3) using ggplot2. Statistical tests are described in the relevant figure
420 legends.

421

422 **Study approval**

423 The study was approved by the IRBs at Fondazione IRCCS Policlinico San Matteo and Erasme Hospital
424 for sample collection, and Dartmouth College for sample testing and analysis.

425

426 **Data availability**

427 Data used in the study is available as **Supplemental Data File 1**.

428

429 **Acknowledgements**

430 This study was supported in part by Infect-ERA, GSK and Sanofi via the CYMAF consortium, as well as
431 by NIH U19AI45825. Pentamer and gB proteins were provided by GSK. A.M. is Research Director at the
432 F.R.S., FNRS, Belgium.

433

434 **Conflict of Interest**

435 A patent application for Systems and methods for identifying and treating primary and latent infections
436 and/or determining time since infection, WO2023154857A2, has been submitted by A.P.H, D. L., A.M,
437 and M.E.A..

438 **References:**

- 439 1 Cannon, M. J. & Davis, K. F. Washing our hands of the congenital cytomegalovirus disease
440 epidemic. *BMC Public Health* **5**, 70 (2005). <https://doi.org/10.1186/1471-2458-5-70>
- 441 2 Schleiss, M. R. Congenital cytomegalovirus infection: molecular mechanisms mediating viral
442 pathogenesis. *Infect Disord Drug Targets* **11**, 449-465 (2011).
443 <https://doi.org/10.2174/187152611797636721>
- 444 3 Dietrich, M. L. & Schieffelin, J. S. Congenital Cytomegalovirus Infection. *Ochsner J* **19**, 123-130
445 (2019). <https://doi.org/10.31486/toj.18.0095>
- 446 4 Pass, R. F. & Anderson, B. Mother-to-Child Transmission of Cytomegalovirus and Prevention of
447 Congenital Infection. *J Pediatric Infect Dis Soc* **3 Suppl 1**, S2-6 (2014).
448 <https://doi.org/10.1093/jpids/piu069>
- 449 5 Stagno, S. *et al.* Congenital cytomegalovirus infection: The relative importance of primary and
450 recurrent maternal infection. *N Engl J Med* **306**, 945-949 (1982).
451 <https://doi.org/10.1056/nejm198204223061601>
- 452 6 Lazzarotto, T. *et al.* Congenital Cytomegalovirus Infection: A Narrative Review of the Issues in
453 Screening and Management From a Panel of European Experts. *Frontiers in Pediatrics* **8** (2020).
454 <https://doi.org/10.3389/fped.2020.00013>
- 455 7 Simonazzi, G. *et al.* Perinatal Outcomes of Non-Primary Maternal Cytomegalovirus Infection: A
456 15-Year Experience. *Fetal Diagn Ther* **43**, 138-142 (2018). <https://doi.org/10.1159/000477168>
- 457 8 Gatta, L. A. *et al.* Clinical factors associated with cytomegalovirus shedding among seropositive
458 pregnant women. *Am J Obstet Gynecol MFM* **4**, 100560 (2022).
459 <https://doi.org/10.1016/j.ajogmf.2021.100560>
- 460 9 Zelini, P. *et al.* Human cytomegalovirus non-primary infection during pregnancy: antibody
461 response, risk factors and newborn outcome. *Clin Microbiol Infect* **28**, 1375-1381 (2022).
462 <https://doi.org/10.1016/j.cmi.2021.09.013>
- 463 10 Kenneson, A. & Cannon, M. J. Review and meta-analysis of the epidemiology of congenital
464 cytomegalovirus (CMV) infection. *Rev Med Virol* **17**, 253-276 (2007).
465 <https://doi.org/10.1002/rmv.535>
- 466 11 Revello, M. G. *et al.* A randomized trial of hyperimmune globulin to prevent congenital
467 cytomegalovirus. *N Engl J Med* **370**, 1316-1326 (2014). <https://doi.org/10.1056/NEJMoa1310214>
- 468 12 Hughes, B. L. *et al.* A Trial of Hyperimmune Globulin to Prevent Congenital Cytomegalovirus
469 Infection. *New England Journal of Medicine* **385**, 436-444 (2021).
470 <https://doi.org/10.1056/NEJMoa1913569>
- 471 13 Kagan, K. O. *et al.* Outcome of pregnancies with recent primary cytomegalovirus infection in first
472 trimester treated with hyperimmunoglobulin: observational study. *Ultrasound Obstet Gynecol*
473 **57**, 560-567 (2021). <https://doi.org/10.1002/uog.23596>
- 474 14 Ross, S. A., Novak, Z., Pati, S. & Boppana, S. B. Overview of the diagnosis of cytomegalovirus
475 infection. *Infect Disord Drug Targets* **11**, 466-474 (2011).
476 <https://doi.org/10.2174/187152611797636703>
- 477 15 Year 2007 position statement: Principles and guidelines for early hearing detection and
478 intervention programs. *Pediatrics* **120**, 898-921 (2007). <https://doi.org/10.1542/peds.2007-2333>
- 479 16 Prince, H. E. & Lapé-Nixon, M. Role of cytomegalovirus (CMV) IgG avidity testing in diagnosing
480 primary CMV infection during pregnancy. *Clin Vaccine Immunol* **21**, 1377-1384 (2014).
481 <https://doi.org/10.1128/cvi.00487-14>
- 482 17 Saldan, A. *et al.* Testing for Cytomegalovirus in Pregnancy. *Journal of Clinical Microbiology* **55**,
483 693-702 (2017). <https://doi.org/doi:10.1128/JCM.01868-16>

- 484 18 Hughes, B. L. & Gyamfi-Bannerman, C. Diagnosis and antenatal management of congenital
485 cytomegalovirus infection. *Am J Obstet Gynecol* **214**, B5-b11 (2016).
486 <https://doi.org/10.1016/j.ajog.2016.02.042>
- 487 19 Abdullahi Nasir, I., Babayo, A. & Shehu, M. S. Clinical Significance of IgG Avidity Testing and
488 Other Considerations in the Diagnosis of Congenital Cytomegalovirus Infection: A Review
489 Update. *Medical Sciences* **4**, 5 (2016).
- 490 20 Chatzakis, C. *et al.* Timing of primary maternal cytomegalovirus infection and rates of vertical
491 transmission and fetal consequences. *Am J Obstet Gynecol* **223**, 870-883.e811 (2020).
492 <https://doi.org/10.1016/j.ajog.2020.05.038>
- 493 21 Aghaeepour, N. *et al.* An immune clock of human pregnancy. *Science Immunology* **2**, eaan2946
494 (2017). <https://doi.org/doi:10.1126/sciimmunol.aan2946>
- 495 22 Zuniga, E. I., Macal, M., Lewis, G. M. & Harker, J. A. Innate and Adaptive Immune Regulation
496 During Chronic Viral Infections. *Annu Rev Virol* **2**, 573-597 (2015).
497 <https://doi.org/10.1146/annurev-virology-100114-055226>
- 498 23 Rouse, B. T. & Sehrawat, S. Immunity and immunopathology to viruses: what decides the
499 outcome? *Nat Rev Immunol* **10**, 514-526 (2010). <https://doi.org/10.1038/nri2802>
- 500 24 Ovies, C., Semmes, E. C. & Coyne, C. B. Pregnancy influences immune responses to SARS-CoV-2.
501 *Science Translational Medicine* **13**, eabm2070 (2021).
502 <https://doi.org/doi:10.1126/scitranslmed.abm2070>
- 503 25 Steinhoff, M. C. *et al.* Influenza Immunization in Pregnancy — Antibody Responses in Mothers
504 and Infants. *New England Journal of Medicine* **362**, 1644-1646 (2010).
505 <https://doi.org/10.1056/NEJMc0912599>
- 506 26 Atyeo, C. G. *et al.* Maternal immune response and placental antibody transfer after COVID-19
507 vaccination across trimester and platforms. *Nature Communications* **13**, 3571 (2022).
508 <https://doi.org/10.1038/s41467-022-31169-8>
- 509 27 Ripalti, A. *et al.* Construction of polyepitope fusion antigens of human cytomegalovirus ppUL32:
510 reactivity with human antibodies. *Journal of clinical microbiology* **32**, 358-363 (1994).
511 <https://doi.org/10.1128/jcm.32.2.358-363.1994>
- 512 28 Chandramouli, S. *et al.* Structure of HCMV glycoprotein B in the postfusion conformation bound
513 to a neutralizing human antibody. *Nat Commun* **6**, 8176 (2015).
514 <https://doi.org/10.1038/ncomms9176>
- 515 29 Chandramouli, S. *et al.* Structural basis for potent antibody-mediated neutralization of human
516 cytomegalovirus. *Sci Immunol* **2** (2017). <https://doi.org/10.1126/sciimmunol.aan1457>
- 517 30 Ye, X. *et al.* Recognition of a highly conserved glycoprotein B epitope by a bivalent antibody
518 neutralizing HCMV at a post-attachment step. *PLoS Pathog* **16**, e1008736 (2020).
519 <https://doi.org/10.1371/journal.ppat.1008736>
- 520 31 Wrapp, D. *et al.* Structural basis for HCMV Pentamer recognition by neuropilin 2 and neutralizing
521 antibodies. *Sci Adv* **8**, eabm2546 (2022). <https://doi.org/10.1126/sciadv.abm2546>
- 522 32 Nachbagauer, R. *et al.* Age Dependence and Isotype Specificity of Influenza Virus Hemagglutinin
523 Stalk-Reactive Antibodies in Humans. *mBio* **7**, 10.1128/mbio.01996-01915 (2016).
524 <https://doi.org/doi:10.1128/mbio.01996-15>
- 525 33 Yates, N. L. *et al.* Multiple HIV-1-specific IgG3 responses decline during acute HIV-1: implications
526 for detection of incident HIV infection. *Aids* **25**, 2089-2097 (2011).
527 <https://doi.org/10.1097/QAD.0b013e32834b348e>
- 528 34 Doerr, H. W., Rentschler, M. & Scheifler, G. Serologic detection of active infections with human
529 herpes viruses (CMV, EBV, HSV, VZV): Diagnostic potential of IgA class and IgG subclass-specific
530 antibodies. *Infection* **15**, 93-98 (1987). <https://doi.org/10.1007/BF01650204>

- 531 35 Gilljam, G. & Wahren, B. Properties of IgG subclasses to human cytomegalovirus. *J Virol Methods*
532 **25**, 139-151 (1989). [https://doi.org/10.1016/0166-0934\(89\)90028-1](https://doi.org/10.1016/0166-0934(89)90028-1)
- 533 36 Gupta, C. K., Leszczynski, J., Gupta, R. K. & Siber, G. R. IgG subclass antibodies to human
534 cytomegalovirus (CMV) in normal human plasma samples and immune globulins and their
535 neutralizing activities. *Biologicals* **24**, 117-124 (1996). <https://doi.org/10.1006/biol.1996.0015>
- 536 37 Joassin, L., Somze, A. & Reginster, M. Detection by enzyme-linked immunosorbent assay of
537 specific immunoglobulin G isotypes in primary and established cytomegalovirus infections.
538 *Journal of Clinical Microbiology* **27**, 139-144 (1989). [https://doi.org/doi:10.1128/jcm.27.1.139-](https://doi.org/doi:10.1128/jcm.27.1.139-144.1989)
539 [144.1989](https://doi.org/doi:10.1128/jcm.27.1.139-144.1989)
- 540 38 Rodriguez, G. E. & Adler, S. P. Immunoglobulin G subclass responses to cytomegalovirus in
541 seropositive patients after transfusion. *Transfusion* **30**, 528-531 (1990).
542 <https://doi.org/https://doi.org/10.1046/j.1537-2995.1990.30690333484.x>
- 543 39 Powers, C., DeFilippis, V., Malouli, D. & Früh, K. Cytomegalovirus immune evasion. *Curr Top*
544 *Microbiol Immunol* **325**, 333-359 (2008). https://doi.org/10.1007/978-3-540-77349-8_19
- 545 40 Jackson, S. E. *et al.* CMV immune evasion and manipulation of the immune system with aging.
546 *Geroscience* **39**, 273-291 (2017). <https://doi.org/10.1007/s11357-017-9986-6>
- 547 41 Genser, B., Truschnig-Wilders, M., Stünzner, D., Landini, M. P. & Halwachs-Baumann, G.
548 Evaluation of five commercial enzyme immunoassays for the detection of human
549 cytomegalovirus-specific IgM antibodies in the absence of a commercially available gold
550 standard. *Clin Chem Lab Med* **39**, 62-70 (2001). <https://doi.org/10.1515/cclm.2001.014>
- 551 42 Rasmussen, L. *et al.* Virus-specific IgG and IgM antibodies in normal and immunocompromised
552 subjects infected with cytomegalovirus. *J Infect Dis* **145**, 191-199 (1982).
553 <https://doi.org/10.1093/infdis/145.2.191>
- 554 43 Naumnik, B. *et al.* Comparison of serology assays and polymerase chain reaction for the
555 monitoring of active cytomegalovirus infection in renal transplant recipients. *Transplant Proc* **39**,
556 2748-2750 (2007). <https://doi.org/10.1016/j.transproceed.2007.08.040>
- 557 44 Razonable, R. R. *et al.* Clinical Diagnostic Testing for Human Cytomegalovirus Infections. *J Infect*
558 *Dis* **221**, S74-S85 (2020). <https://doi.org/10.1093/infdis/jiz601>
- 559 45 Revello, M. G. *et al.* Comparative evaluation of eight commercial human cytomegalovirus IgG
560 avidity assays. *J Clin Virol* **48**, 255-259 (2010). <https://doi.org/10.1016/j.jcv.2010.05.004>
- 561 46 Joassin, L., Somze, A. & Reginster, M. Detection by enzyme-linked immunosorbent assay of
562 specific immunoglobulin G isotypes in primary and established cytomegalovirus infections. *J Clin*
563 *Microbiol* **27**, 139-144 (1989). <https://doi.org/10.1128/jcm.27.1.139-144.1989>
- 564 47 Dugast, A.-S. *et al.* Independent evolution of Fc- and Fab-mediated HIV-1-specific antiviral
565 antibody activity following acute infection. *European Journal of Immunology* **44**, 2925-2937
566 (2014). <https://doi.org/https://doi.org/10.1002/eji.201344305>
- 567 48 Sadanand, S. *et al.* Temporal variation in HIV-specific IgG subclass antibodies during acute
568 infection differentiates spontaneous controllers from chronic progressors. *AIDS* **32**, 443-450
569 (2018). <https://doi.org/10.1097/qad.0000000000001716>
- 570 49 Damelang, T., Rogerson, S. J., Kent, S. J. & Chung, A. W. Role of IgG3 in Infectious Diseases.
571 *Trends Immunol* **40**, 197-211 (2019). <https://doi.org/10.1016/j.it.2019.01.005>
- 572 50 Dauby, N. *et al.* Limited Effector Memory B-Cell Response to Envelope Glycoprotein B During
573 Primary Human Cytomegalovirus Infection. *J Infect Dis* **213**, 1642-1650 (2016).
574 <https://doi.org/10.1093/infdis/jiv769>
- 575 51 Corrales-Aguilar, E. *et al.* Human cytomegalovirus Fcγ binding proteins gp34 and gp68
576 antagonize Fcγ receptors I, II and III. *PLoS Pathog* **10**, e1004131 (2014).
577 <https://doi.org/10.1371/journal.ppat.1004131>

- 578 52 Atalay, R. *et al.* Identification and expression of human cytomegalovirus transcription units
579 coding for two distinct Fcγ receptor homologs. *J Virol* **76**, 8596-8608 (2002).
580 <https://doi.org/10.1128/jvi.76.17.8596-8608.2002>
- 581 53 Antonsson, A. & Johansson, P. J. H. Binding of human and animal immunoglobulins to the IgG Fc
582 receptor induced by human cytomegalovirus. *J Gen Virol* **82**, 1137-1145 (2001).
583 <https://doi.org/10.1099/0022-1317-82-5-1137>
- 584 54 Vezzani, G. *et al.* Human immunoglobulins are transported to HCMV viral envelope by viral Fc
585 gamma receptors-dependent and independent mechanisms. *Front Microbiol* **13**, 1106401
586 (2022). <https://doi.org/10.3389/fmicb.2022.1106401>
- 587 55 Ssentongo, P. *et al.* Congenital Cytomegalovirus Infection Burden and Epidemiologic Risk Factors
588 in Countries With Universal Screening: A Systematic Review and Meta-analysis. *JAMA Network*
589 *Open* **4**, e2120736-e2120736 (2021). <https://doi.org/10.1001/jamanetworkopen.2021.20736>
- 590 56 Wang, C., Zhang, X., Bialek, S. & Cannon, M. J. Attribution of congenital cytomegalovirus
591 infection to primary versus non-primary maternal infection. *Clin Infect Dis* **52**, e11-13 (2011).
592 <https://doi.org/10.1093/cid/ciq085>
- 593 57 Picone, O. *et al.* A series of 238 cytomegalovirus primary infections during pregnancy:
594 description and outcome. *Prenat Diagn* **33**, 751-758 (2013). <https://doi.org/10.1002/pd.4118>
- 595 58 Lazzarotto, T., Guerra, B., Gabrielli, L., Lanari, M. & Landini, M. P. Update on the prevention,
596 diagnosis and management of cytomegalovirus infection during pregnancy. *Clin Microbiol Infect*
597 **17**, 1285-1293 (2011). <https://doi.org/10.1111/j.1469-0691.2011.03564.x>
- 598 59 Nigro, G., Adler, S. P., La Torre, R. & Best, A. M. Passive Immunization during Pregnancy for
599 Congenital Cytomegalovirus Infection. *New England Journal of Medicine* **353**, 1350-1362 (2005).
600 <https://doi.org/10.1056/NEJMoa043337>
- 601 60 Jückstock, J., Rothenburger, M., Friese, K. & Traunmüller, F. Passive Immunization against
602 Congenital Cytomegalovirus Infection: Current State of Knowledge. *Pharmacology* **95**, 209-217
603 (2015). <https://doi.org/10.1159/000381626>
- 604 61 Gong, Y. *et al.* Mathematical Modeling of Rhesus Cytomegalovirus Transplacental Transmission
605 in Seronegative Rhesus Macaques. *Viruses* **15**, 2040 (2023).
- 606 62 Otero, C. E. *et al.* Relationship of maternal cytomegalovirus-specific antibody responses and viral
607 load to vertical transmission risk following primary maternal infection in a rhesus macaque
608 model. *PLoS Pathog* **19**, e1011378 (2023). <https://doi.org/10.1371/journal.ppat.1011378>
- 609 63 Semmes, E. C. *et al.* ADCC-activating antibodies correlate with decreased risk of congenital
610 human cytomegalovirus transmission. *JCI Insight* **8** (2023).
611 <https://doi.org/10.1172/jci.insight.167768>
- 612 64 Boesch, A. W. *et al.* Highly parallel characterization of IgG Fc binding interactions. *mAbs* **6**, 915-
613 927 (2014). <https://doi.org/10.4161/mabs.28808>
- 614 65 Brown, E. P. *et al.* Multiplexed Fc array for evaluation of antigen-specific antibody effector
615 profiles. *J Immunol Methods* **443**, 33-44 (2017). <https://doi.org/10.1016/j.jim.2017.01.010>
- 616 66 Brown, E. P. *et al.* Optimization and qualification of an Fc Array assay for assessments of
617 antibodies against HIV-1/SIV. *J Immunol Methods* **455**, 24-33 (2018).
618 <https://doi.org/10.1016/j.jim.2018.01.013>
- 619 67 McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for
620 dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

621

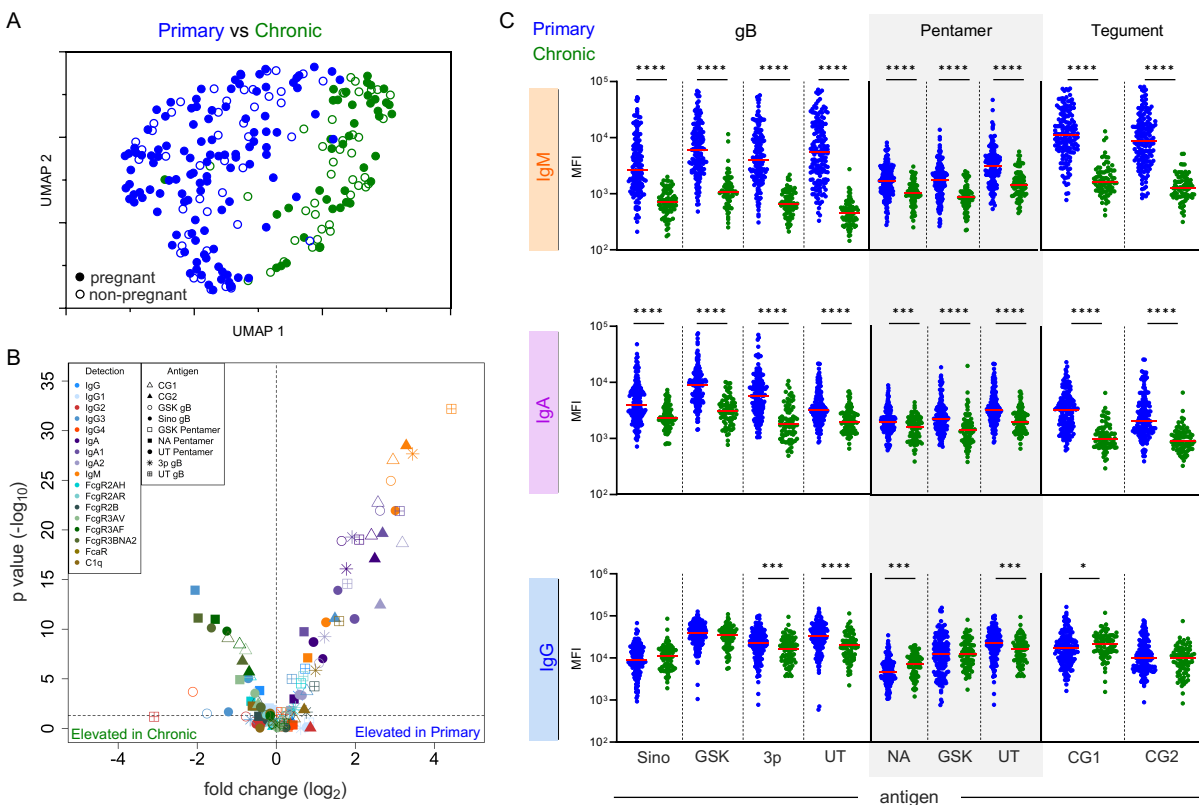


Figure 1: Antibody features distinguish primary from Chronic CMV infection but not pregnancy status. **A.** Uniform manifold approximation (UMAP) biplot of antibody features excluding IgM. Distinct clusters of subjects with primary (blue) and chronic (green) infection but not pregnancy status (hollow and filled symbols) are observed. **B.** Volcano plot of each CMV-specific antibody feature assessed. Volcano plot represents the \log_2 fold change (x-axis) against the $-\log_{10}$ p value (Mann Whitney test: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$). Antibody specificities (Antigen) are indicated by shape and Fc characteristics (Detection) indicated by color. **C.** IgM, IgA, and IgG binding to CMV antigens (further described in Supplemental Table 1). Data are the mean median fluorescent intensity (MFI) values of technical replicates. Solid red line indicates median.

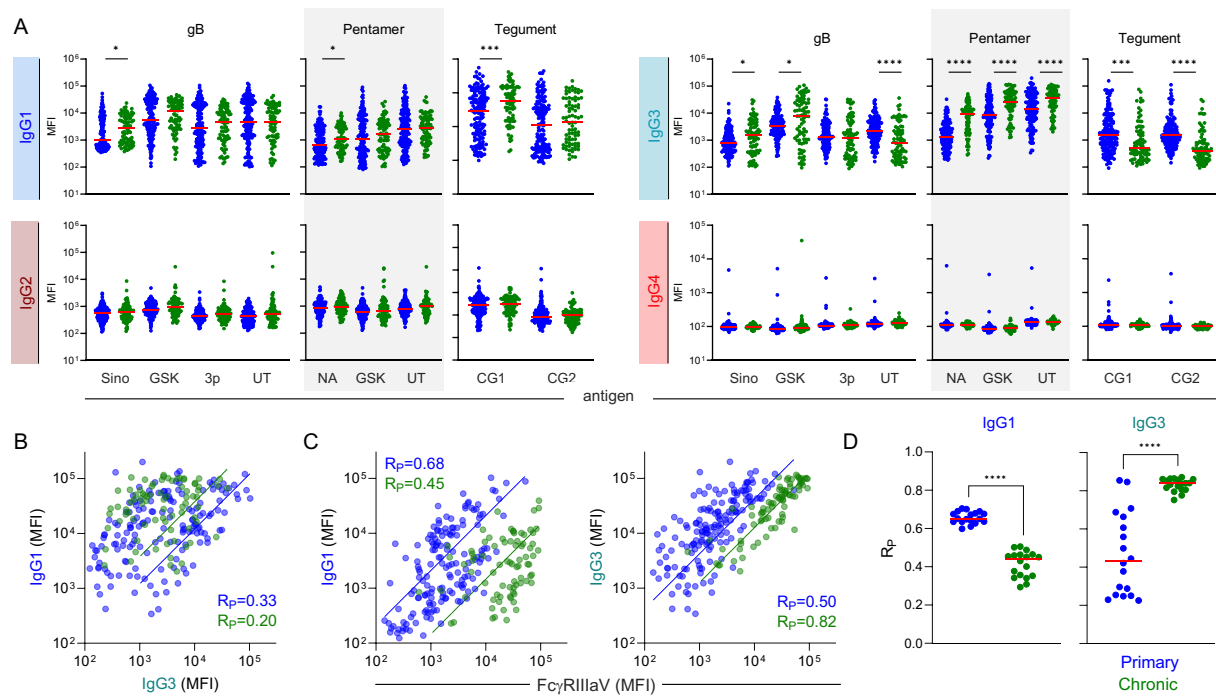


Figure 2: IgG Subclasses display distinct profiles in primary and chronic infection depending on antigen specificity. **A.** Levels of gB-, pentamer-, and tegument-specific IgG1, IgG2, IgG3, and IgG4 antibodies in individual with primary (blue) or chronic (green) CMV infection. **B.** Scatterplot and fit lines of levels of (UT) pentamer-specific IgG1 and IgG3 responses by infection status. Pearson correlation coefficient (R_p) indicated in inset. **C.** Scatterplot and fit line of levels of (UT) pentamer-specific IgG1 (left) and IgG3 (right) versus (UT) pentamer-specific Fc γ RIIIaV-binding antibodies. Pearson correlation coefficient (R_p) indicated in inset. **D.** Pearson correlation coefficients (R_p) for each pentamer antigen tested across Fc γ R by infection status. Solid lines denote group medians; differences between groups were assessed by Mann-Whitney test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$); values presented are median fluorescent intensities (MFI).

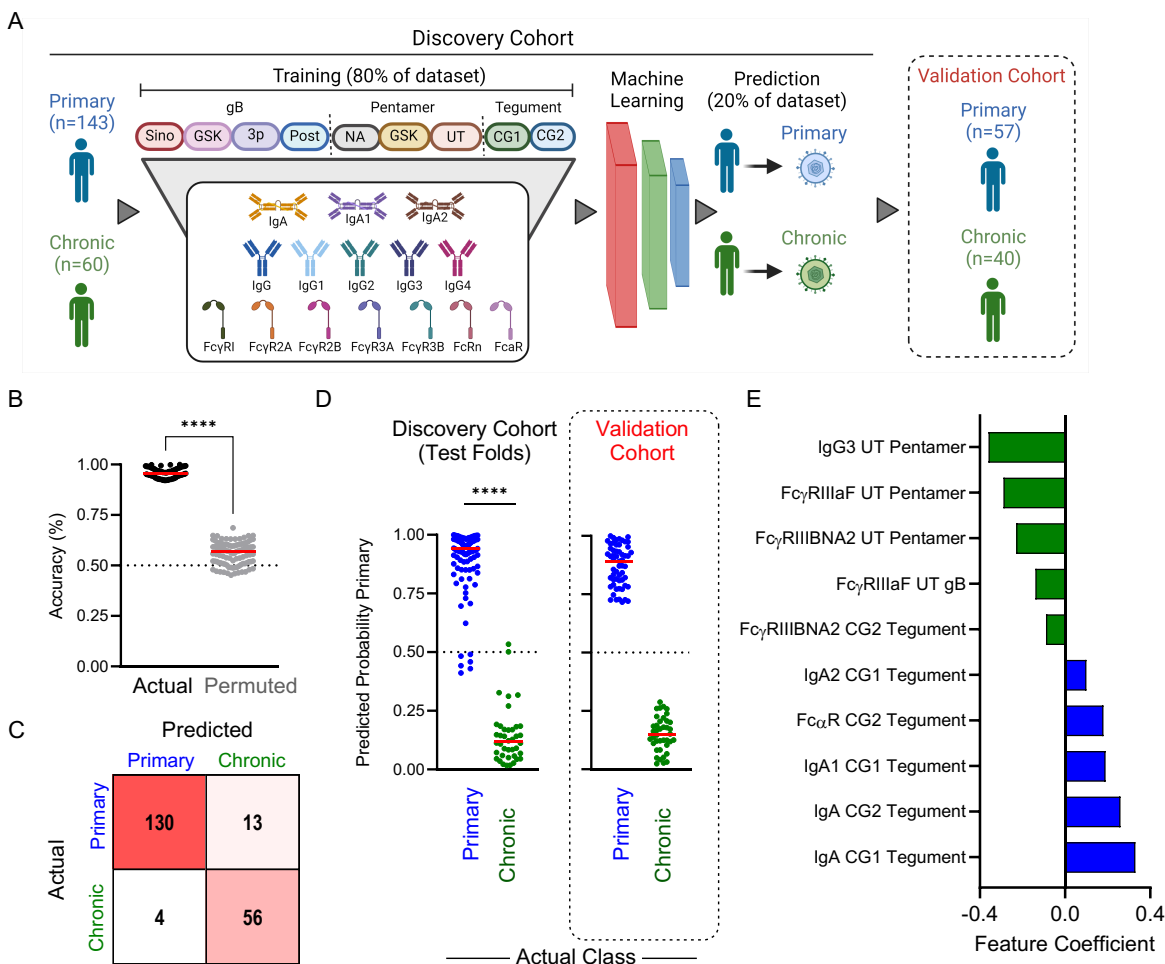


Figure 3: Machine learning accurately predicts primary or chronic CMV infection status. A. Schematic overview of cross-validated machine learning workflow employing antibody profiling data to discriminate between primary and chronic infection status in discovery and validation cohorts. **B.** Prediction accuracy for 100 repeated 5-fold cross-validation runs on actual (black) and permuted (gray) class labels. **C.** Confusion matrix of predicted versus actual class labels in the median model for 5-fold cross validation. **D.** Class probabilities of each sample in the discovery set when evaluated as a test sample in the cross-validation run exhibiting median performance (left) and for the validation cohort using the final model (right). **E.** The identities and coefficients of the features making the largest positive (n=5) or negative (n=5) contributions to the final model. Solid lines denote group medians; differences between groups or conditions were assessed by Mann-Whitney test (*p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001).

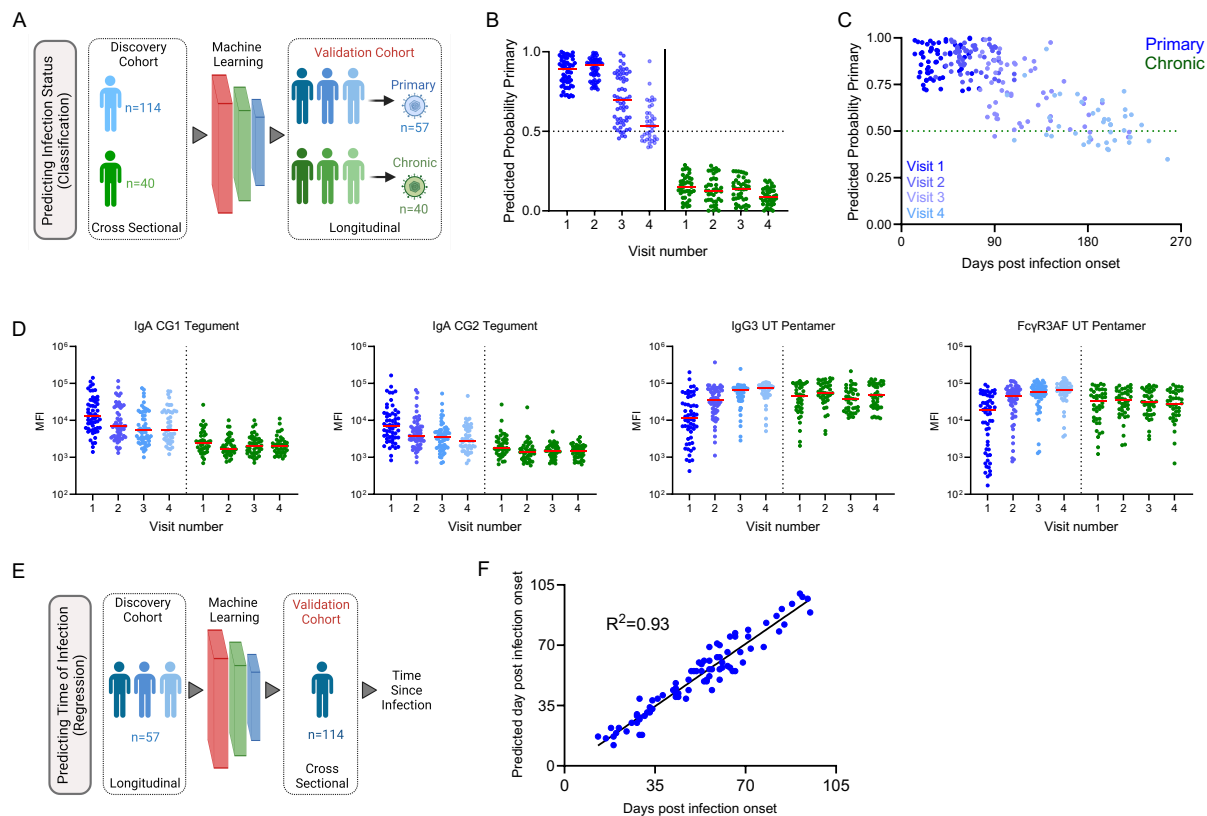


Figure 4: Longitudinal models define a molecular clock of CMV primary infection. **A.** Analysis overview. The infection status classification model trained on the cross-sectional cohort was applied to longitudinal samples available from the validation cohort. **B.** Class probabilities of each sample in longitudinal cohort over sample collection visits for individuals with primary (blue) and chronic (green) infection. **C.** Scatterplot of class probabilities for subjects defined as having primary infection at visit 1 over time. **D.** Scatterplots of features employed by classification model to predict infection status over time in the longitudinal cohort. **E.** Analysis overview. Primary infection samples from the longitudinal cohort samples were used to train a regression model to predict time since infection (days post symptom onset) that was applied to the primary samples from the cross-sectional cohort. **F.** Scatterplot of model predictions of time since infection when primary samples used for predicting days post symptom onset. Data shows the measure of the predicted label and its closeness to the true label.

