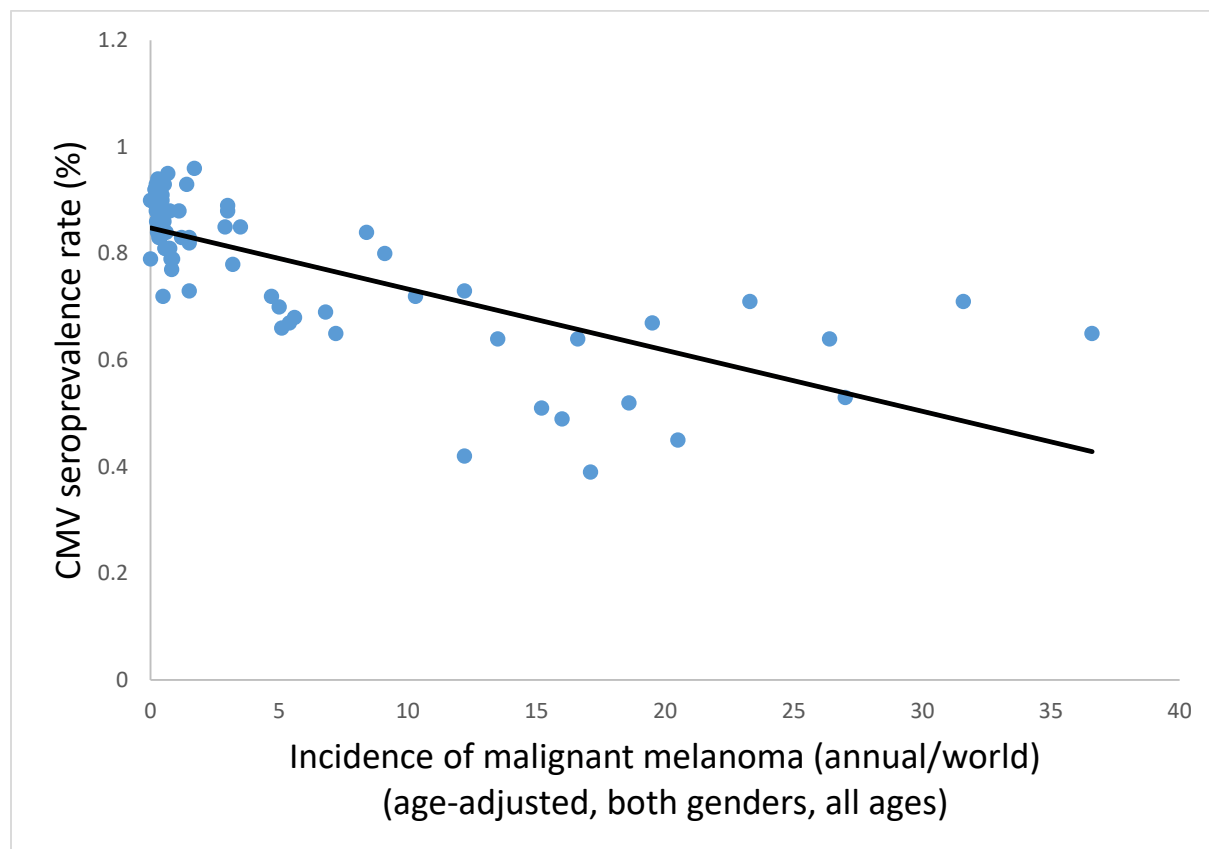
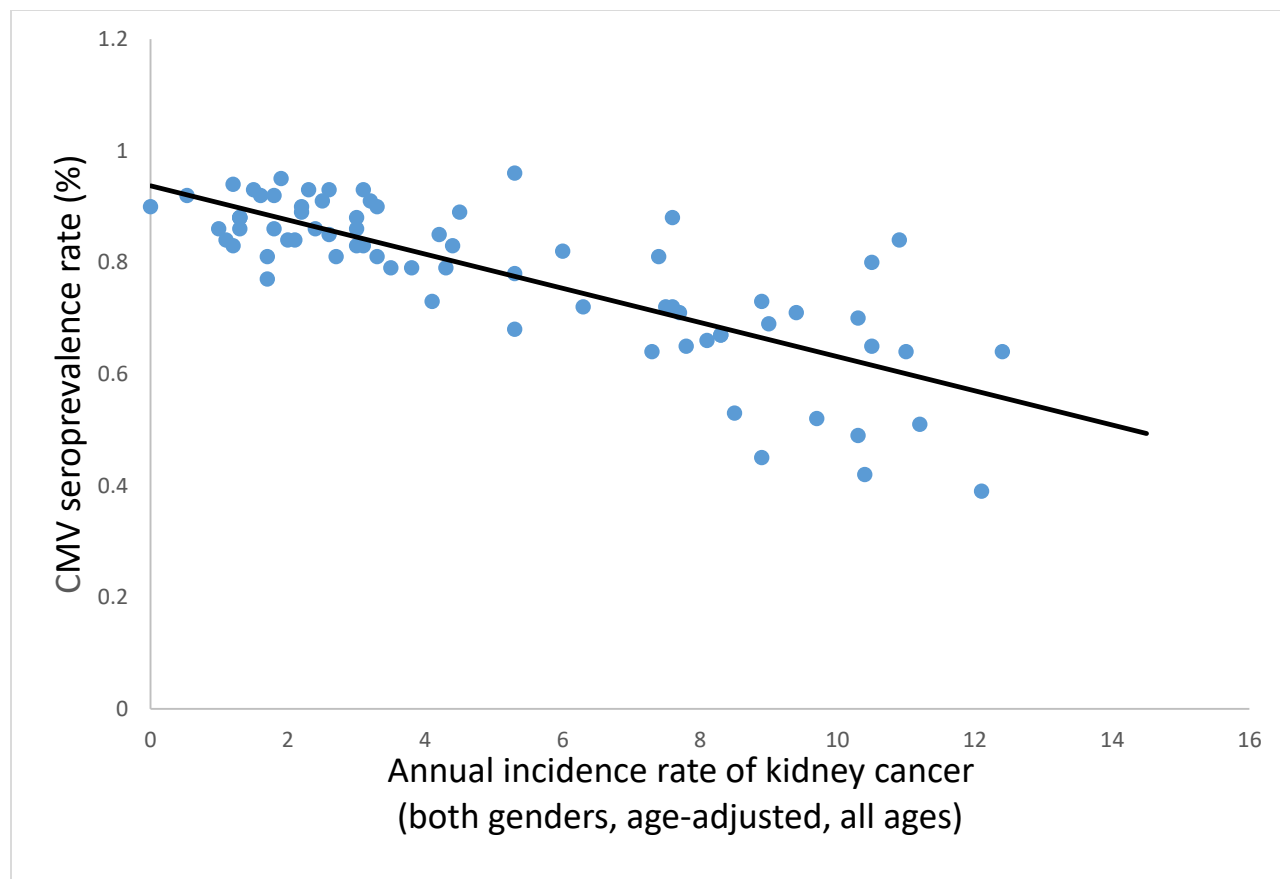


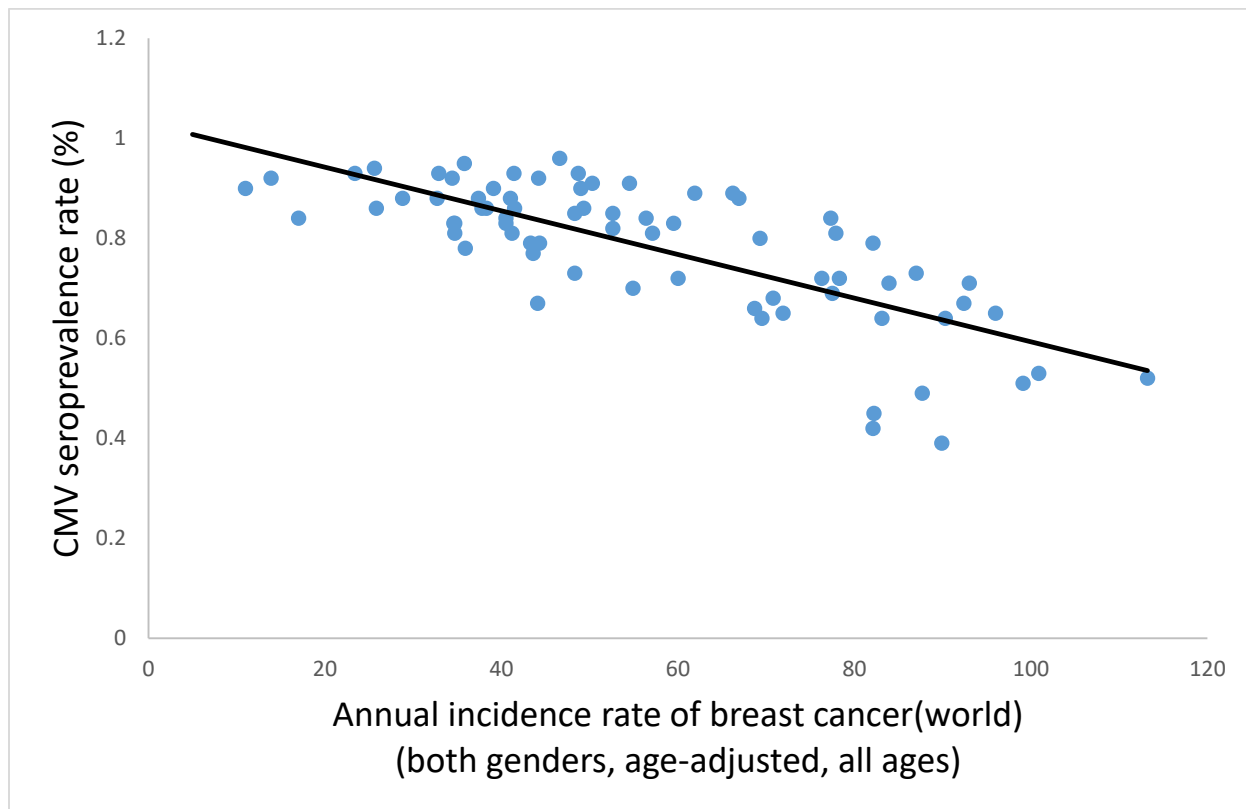
**Figure 1.** New cancer (all invasive types combined, all sites)/ $10^5$  population is plotted against these country CMV seropositivity rates [28]. Inverse relationship over 73 countries [51] (Spearman's  $\rho = -0.732$ ;  $p < 0.001$ ), signals a potential, globally pervasive, protection against tumorigenesis afforded by CMV.



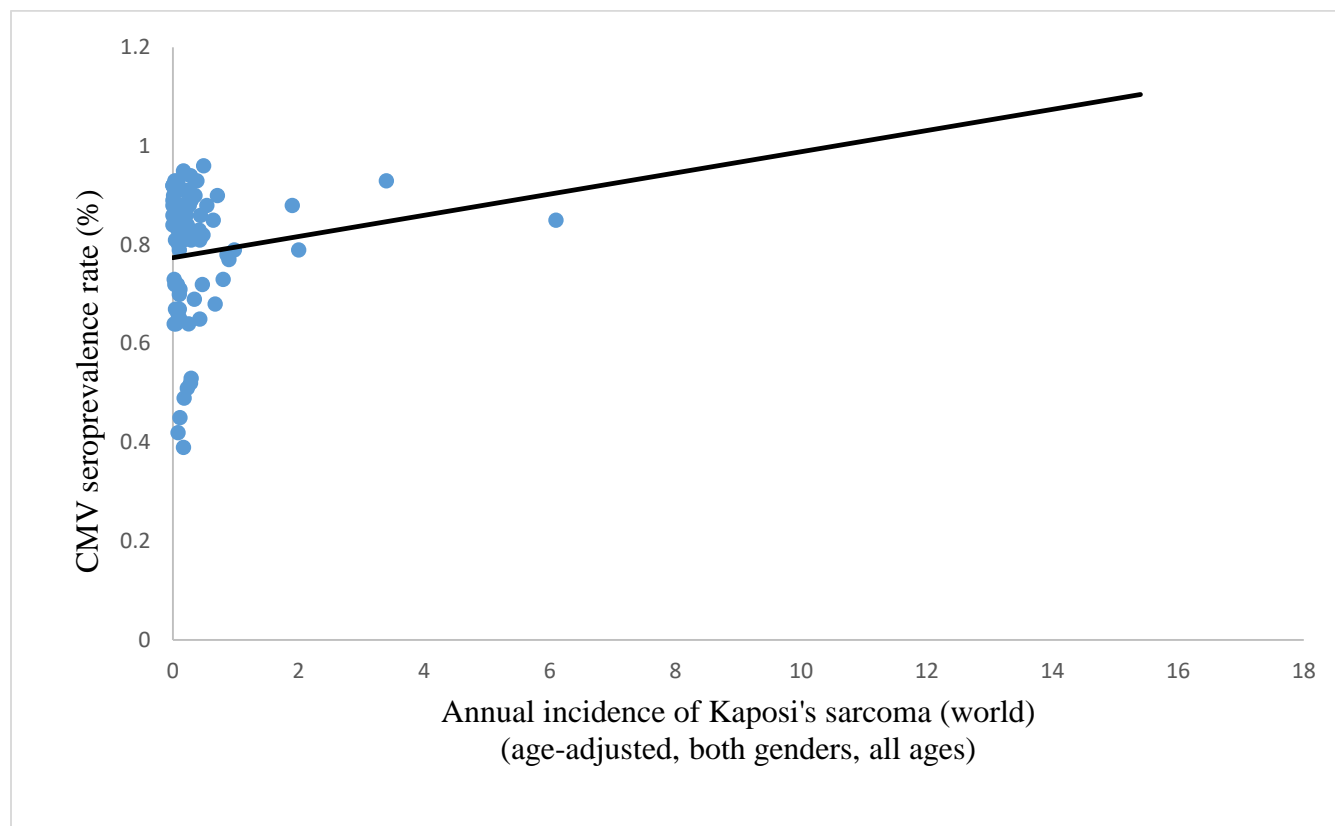
**Figure 2.** Malignant melanoma/ $10^5$  persons vs. country specific CMV seropositivity [28] correlate strongly inversely (Spearman's  $\rho = -0.763$ ;  $p < 0.001$ ) across 73 countries [51].



**Figure 3.** Kidney cancer incidence ( $10^5/\text{year}$ ) correlated to the country specific CMV seroprevalence [28]. A strong and inverse correlation across 73 countries [51] (Spearman's  $\rho = -0.754$ ;  $p < 0.001$ ).



**Figure 4.** Breast cancer incidence/ $10^5$  population/year and country specific CMV seropositivity rate [28] spanning 73 countries [51]. A significant inverse (possibly protective) association is suggested by Spearman's  $\rho = -0.719$ ;  $p < 0.001$ .



**Figure 5.** Incidence of Kaposi's sarcoma (mostly, HIV+ individuals)/10<sup>5</sup> population/year and the specific country-level rate of CMV seropositivity [28]. Seventy-three countries [51] lack an inverse correlation between variables (Spearman's  $\rho = -0.007$ ;  $p=0.953$ ). This suggests weak or absent protection afforded by a concomitant CMV infection.

1 **Title:** *Cytomegalovirus seropositivity relates inversely to cancer incidences across races and*  
2 *ethnicities: implications for oncoprevention*

3  
4 Marko Janković, MD, PhD<sup>\*a</sup>, Ognjen Milićević, MD<sup>b</sup>, Milena Todorović-Balint, MD, PhD<sup>c</sup>,  
5 Irena Đunić, MD, PhD<sup>c</sup>, Biljana Mihaljević, MD, PhD<sup>c</sup>, Tanja Jovanović, MD, PhD<sup>a</sup>, Aleksandra  
6 Knežević, MD, PhD<sup>a</sup>

7  
8 <sup>a</sup>Institute of Microbiology and Immunology, Department of Virology, Faculty of Medicine,  
9 University of Belgrade, 1 Dr Subotića street, 11000 Belgrade, Republic of Serbia.

10  
11 <sup>b</sup>Institute of Medical Statistics and Informatics, Faculty of Medicine, University of Belgrade, 15  
12 dr Subotića street, 11000 Belgrade, Republic of Serbia.

13  
14 <sup>c</sup>Clinic of Hematology, Faculty of Medicine, University Clinical Centre of Serbia, University of  
15 Belgrade, 2 dr Koste Todorovića street, 11000 Belgrade, Republic of Serbia

16  
17 \*Corresponding author:

18  
19 Marko Janković, MD, PhD [e-mail: [marko.jankovic@med.bg.ac.rs](mailto:marko.jankovic@med.bg.ac.rs)]

20 Institute of Microbiology and Immunology | Department of Virology

21 Faculty of Medicine | University of Belgrade,

22 1 dr Subotića Street, 11000 Belgrade,

23 Republic of Serbia

24 **Abstract**

25 **Background.** Race and ethnic disparities in cancer incidence rates and the prevalence of  
26 *cytomegalovirus* (CMV) are known to exist in the United States (U.S.) but also across broad  
27 geographic expanses. CMV prevalence seems to inversely contrast tumor incidence rates both in  
28 ethnic groups and globally. Is there a biological link between cancer and CMV infection? Most  
29 recent clinical results seem to certify it.

30 **Methods.** Global cancer data were retrieved from the World Health Organization (WHO)  
31 database. Incidence of cancer and CMV seroprevalence (73 countries) were subjected to  
32 Spearman's correlation test. The Bayesian framework was adopted for CMV seropositivity  
33 variables. Relevant data for the U.S. were extracted from publications based on the Surveillance,  
34 Epidemiology, and End Results (SEER) registries and the National Health and Nutrition  
35 Examination Surveys (NHANES), 1988-2004.

36 **Results.** An inversely directed coupling between cancer and CMV seropositivity across diverse  
37 ecologies and cultural domains suggest a global oncoprotective effect of the CMV (Spearman's  $\rho$   
38 = -0.732;  $p < 0.001$ ). Rates of all cancers combined and CMV seropositivity show an opposite  
39 association ( $p < 0.001$ ) among the races and foremost U.S. ethnic groups.

40 **Conclusion.** The racial/ethnic incidence of cancers and CMV seropositivity are inversely  
41 proportional both in the U.S. and globally. This would support a view that CMV is a potential  
42 driver against tumorigenesis. An absence of CMV infection abrogates protection against  
43 malignant clones afforded to an infected host. Abating CMV seroprevalence may relate causally  
44 to the buildup of malignancies in U.S. and the West world countries with thriving hygiene and  
45 healthcare systems.

46 **Keywords:** Cancer, cytomegalovirus, oncoprotection, global, United States

47 **Importance.**

48 Increasing evidence substantiates the potential of cytomegalovirus (CMV) to counteract tumors,  
49 particularly in the field of anti-cancer vaccinology, leading to extended periods of survival. This  
50 research unveils a robust and inverse correlation between the prevalence of CMV and the  
51 occurrence of cancer both within the United States and on a global scale (73 countries), hinting  
52 at the ability of CMV to inhibit tumor development. Furthermore, this phenomenon remains  
53 consistent across various racial and ethnic groups within the United States and applies to a  
54 diverse range of cancer types. A notable pattern emerges: the higher the prevalence of the viral  
55 infection, the lower the incidence of tumors within a given country. These findings support  
56 existing insights from clinical and experimental investigations, underscoring the notion that this  
57 effect becomes apparent at the level of entire nations and populations worldwide.

58

59

60

61

62

63

64

65

66

67

68

69



70

71 **Introduction**

72

73 Human cytomegalovirus (CMV) is a ubiquitous  $\beta$ -herpesvirus typically causing an asymptomatic  
74 to mild infection in healthy children and adults. In the immunocompromised patient and the  
75 unprotected fetus CMV can present as a significant cause of morbidity and mortality. After  
76 primary contact, the virus persists within the host as a lifelong infection awakening at times from  
77 dormancy and leading to reactivation and virion shedding. Over millennia of evolutionary time,  
78 CMV had interlocked in a complex interplay with the immune system of its human host, with  
79 growing evidence pointing to T lymphocytes as being galvanized by CMV and directed against  
80 specific cancer cells. Recently, we drew attention to an inversely correlated incidence of human  
81 B lymphocyte malignancies and CMV seroprevalence in a single-center clinical setting and in  
82 humans across races/ethnicities the world over (Spearman's coefficient  $\rho = -0.625$ ,  $p < 0.001$ )  
83 [13]. This hinted at a possible oncopreventive faculty of the virus (**Fig. 1**). Similar findings were  
84 described in animal *in vitro* tissue models [6-8] covering several oncological diseases [9-11,12].  
85 This indicates that CMV infection may bestow a capacity on its host to resist carcinogenesis  
86 somewhat more effectively than does a normal immunity in an uninfected subject. In effect,  
87 CMV and cancer may be profoundly connected. Virus/host interaction has not been investigated  
88 in a racial/ethnic background or across the globe. Inquiring into this association we used the  
89 wealth of data derived from authorized literature based on the U.S. registries and global  
90 publications (World Health Association [WHO]).

91 An intriguing disparity in overall health status between Whites, Blacks, and Hispanics  
92 and the CMV seroprevalence patterns by race and ethnicity in the U.S. evades a full explanation

93 by factors of family size, household income level, education, marital status, census region, area  
94 of residence, country of birth or type of medical insurance. Differential exposure to CMV might  
95 partially explain these incongruities in health status, though [5,14,15].

96 The U.S. Hispano-Americans (Latinos) have significantly better health and mortality  
97 outcomes than the average population [16], contradicting their low socioeconomic status (SES).  
98 Medical experts have known for some time that Latinos living in the U.S. have on average a  
99 better life expectancy than non-Hispanic Whites. The so-called “Hispanic paradox” was recently  
100 supported by new data from the U.S. Centers for Disease Control and Prevention (CDC) [17-20].  
101 Of note, the “paradox” seems not to hold true for follicular lymphoma (FL) and chronic  
102 lymphocytic leukemia (CLL) [21].

103 In this work we correlated age-adjusted incidence rates of all primary invasive cancers  
104 combined with the age-adjusted CMV seropositivity estimates among races and minorities in the  
105 U.S. The presentation of the reports used (**Table 1**) shows that, comparatively, CMV  
106 seropositivity is higher in the U.S. minorities than in non-Hispanic Whites. The significance of  
107 this observation is confirmed by statistical analysis. Moreover, we show that this effect is global,  
108 predominating across nations, societies, and histological types of malignant tumors (**Figs. 1-4**  
109 and **Table 3**).

110

## 111 **Materials and Methods**

112

113 **Study hypothesis.** Our primary outcome of interest is an inverse association between rates of  
114 new cancer and CMV seropositive status the world over and within race/ethnic groups in the  
115 U.S.

116  
117 **Study populations – worldwide data (73 countries).** In order to evaluate whether the proposed  
118 cancer incidence/CMV seroprevalence association holds true at a global scale, we accessed the  
119 World Health Organization Global Cancer Observatory (International Agency for Research on  
120 Cancer [IARC]), for data on worldwide cancer statistics. Worldwide annual incidence of new  
121 cancers (age-adjusted, both genders, all ages) was used. Country specific seroprevalence of  
122 CMV comes from work of Zuhair *et al.* [28].

123  
124 **Study populations – the U.S.** We examined a connection between an overall rate of new  
125 cancers (combined at all anatomical sites, all ages) among races/ethnic groups in the U.S. and the  
126 demographic pervasiveness of CMV seropositivity in these populations.

127 Patients with primary diagnosis of cancer (invasive) evidenced in the Surveillance,  
128 Epidemiology, and End Results (SEER) Program database between January 2007 and December  
129 2015 were included in the study. The population was categorized into non-Hispanic Whites, non-  
130 Hispanic Black, Non-Hispanic Asian/Pacific Islanders (API), Non-Hispanic American  
131 Indian/Alaskan Natives (AI/AN), and Hispanics.

132 We collated national information from the SEER registry, an authoritative high-quality  
133 resource and error-proofed data source for the burden of cancer among the U.S. populaces. The  
134 SEER database is updated yearly. It contains patient demography, the primary site of the tumors,  
135 histology, and cancer stage at medical detection time [30,31]. The current study is based upon  
136 the reports of meticulous work of researchers on sure and accessible data on cancer rates for the  
137 U.S. race/ethnic groups (the Cancer Planet website and publications from the North American  
138 Association of Central Cancer Registries or NAACCR and the ACS). Oncologic parameters of

139 race/ethnicity groups are defined by the American Joint Committee on Cancer (AJCC). Also,  
140 SEER program is the only source for historic population-based incidence and survival data  
141 (1975-2018). SEER 22 Incidence provided coarse rates (2000-2019) with the total registries data  
142 for all cancers combined, including sex, race, and ethnicity [34-50]. Numerical derivatives based  
143 on this data were made available by published reports herein referred to.

144 For data on CMV, we consulted a series of cross-sectional surveys drawn from NHANES  
145 collected by the National Center for Health Statistics (NCHS). The reports utilized here are  
146 based on the 2003-2004 wave which included CMV latency in the population and recognized  
147 consequences of persistent CMV infection on human health [29-31]. We made use of CMV  
148 prevalence data from the Third National Health and Nutrition Examination Survey (NHANES),  
149 1988–1994. NHANES III was cross-sectional and stratified to allow for heterogeneity,  
150 multistage probability sample of civilian non-institutionalized U.S. population aged 2 months to  
151 90 years.

152 To obtain current nationally representative estimates of the prevalence of CMV in the  
153 U.S., we used NHANES III study data (1988-1994) from Staras *et al.* [5]. NHANES is a series of  
154 cross-sectional surveys supervised and managed by the National Center for Health Statistics  
155 (NCHS) of the Centers for Disease Control and prevention CDC [38,39]. Also, we leaned on  
156 information and arguments from Bate *et al.* (**Tables 1, 2** and *Results* section [33]). The overall  
157 age-adjusted prevalence of CMV seropositivity seems not to have changed significantly in the  
158 U.S. for the intervals 1988-1994 and 1999-2004 (**Table 1** in [33]).

159 Also, we used query tools to collect literature related to CMV seroprevalence and cancer  
160 burden in the U.S. and elsewhere through MEDLINE, PubMed database search engine [terms  
161 'cytomegalovirus', 'prevalence', 'IgG', 'race/ethnic', 'global burden of cancer', and their

162 synonymous expressions]. Initially, we focused on CMV prevalence data for information  
163 regarding sex, race/ethnicity and SES, inversely concordant with a rate of cancer across  
164 geographic domains [13].

165 We have drawn on data from the primary literature, recognized reports and authoritative  
166 reviews on cancer incidence and CMV seroprevalence rates published heretofore (**Table 1**, [10]).  
167 The collection periods vary, spanning multiple years (see *Limitations*). We entirely relied on  
168 systematic reviews and meta-analyses of the epidemiological burden of CMV in the U.S.  
169 extracted from Medline and LILACS (Latin American and Caribbean Health Sciences Literature  
170 (10 October, 2020) [2,32,33].

171

172 **Statistical analysis.** For the analysis of aggregated data points at the ethnicity level for CMV  
173 seropositivity and cancer incidence, we utilized available descriptive statistics to empirically  
174 determine the significance of correlation. For the CMV seropositivity variables (proportion  $p$  and  
175 the number of subjects  $N$ ), we adopted the Bayesian framework with the assumed uniform  $U(0,1)$   
176 prior and binomial  $Bi(p,N)$  data distribution yielding a beta  $Be(p*N+1, (1-p)N+1)$  posterior  
177 distribution. For the cancer incidence variables (estimated incidence and the 95% confidence  
178 interval (Henley *et al.* 2020 [44]) we assumed the maximum likelihood normal distribution; the  
179 Bayesian framework was omitted due to the variable methodology of incidence calculation and  
180 the assumption of a large number of observed data points. All the variables for all the ethnicity  
181 groups were independently simulated from their distributions 10 000 times under the null  
182 hypothesis that the data points are uncorrelated, and the Pearson coefficient of correlation was  
183 calculated each time. The empirical distribution of the correlation coefficient was estimated from  
184 the data giving the mean and 95% confidence interval. The 2-sided significance is calculated

185 from the value of 0 and the significance level of 0.05. The analysis was done with custom scripts  
186 and the SciPy package of the Python programming language. Descriptive statistics, including  
187 frequencies and percentages were used for defining baseline characteristics of populations.  
188 Results are presented as counts, percentages (in parentheses) or median (interquartile 3 12 range)  
189 and frequency distributions depending on data type. Data was organized using Microsoft Excel  
190 software 2010 (Microsoft Corporation, Redmond, WA, U.S.). Spearman's correlation coefficient  
191 ( $\rho$ ) served to capture significance of correlation between variables assessed across the globe.  
192 Based on the method of covariance, it is a preferable method of measuring the agreement  
193 between variables of interest. Also, it provides information on the direction of the relationship.  
194  $P < 0.05$  tested against an artifact of chance.

195

## 196 **Results**

197

198 The prevalence of CMV seropositivity correlates inversely to cancer incidence rates across a  
199 broad range of histology and worldwide (**Table 3**; the top 3 malignancies with strongest  
200 correlations are represented in **Figs. 2-4**). Correlations are obtained in 30/34 (88.2%) tumors and  
201 are highly significant over 25/34 (73.5%) histology types. Also, an inverse association attains a  
202 significance if annual incidence of new cancers of all histologic types was combined  
203 (Spearman's  $\rho = -0.732$ ;  $p < 0.001$ , **Fig. 1**). These results favor a possible oncoprotection that  
204 CMV infection provides to its host.

205         Conversely, CMV seroprevalence has a significant and co-incremental relation with an  
206 incidence of malignancies of nasopharynx and the gallbladder, suggesting a potential pro-  
207 oncogenic effect regarding these malignancies.

208 Notably, country-specific prevalence of CMV did not correlate inversely with an overall  
209 incidence of Kaposi's sarcoma (**Table 3** and **Fig. 5**, Spearman's  $\rho = -0.007$ ;  $p=0.953$ ). This is in  
210 line with a requirement of an operative T cell-mediated oncolytic response as being critical for  
211 the expression a protective capacity of CMV, supporting a hypothesis of specific T cell viral  
212 oncoprevention. We propose that a severely disrupted T cell immunity in HIV/AIDS patients,  
213 who comprise a vast majority of population with Kaposi's sarcoma, effectively precludes a  
214 CMV-associated anti-tumor response.

215 **Table 1** exhibits comparative data on race/ethnic estimates of the incidence of all cancers  
216 combined (both genders, all anatomical sites, all ages) in the U.S, and the territorially matching  
217 prevalence of CMV seropositivity. The estimate of cancer incidence for Mexico is included as a  
218 contrasting population but was discarded from statistical calculations. The aggregated results of  
219 studies mentioned therein provide a clear representation of a higher CMV pervasiveness in the  
220 U.S. minorities than in non-Hispanic Whites.

221 **Table 2** represents a statistical analysis of the SEER data on racial/ethnic rates of cancer  
222 (Henley *et al.* [44]) and the corresponding prevalence of CMV. Also pointed in **Table 2** is a  
223 highly significant and inverse correlation ( $p<0.001$ ; mean:  $-0.674$ ;  $SD=0.018$ ;  $CI=-0.711$ )  
224 between the standardized cancer incidence and the CMV seroprevalence in the U.S. An outlying  
225 group are the Eskimos (AI/AN), perhaps because of an impaired T cell function which  
226 distinguishes this ethnicity from the others [66,67]. Consequently, CMV-specific cytotoxic T cell  
227 oncolysis in Eskimos is not an efficient suppressor of clonogenic processes (see *Discussion*). The  
228 results also support broad evidence (Figs. 1-4) suggesting a tendency of CMV seropositivity to  
229 monotonically decrease as SES improves across a range of developed and underdeveloped

230 countries. This may explain a higher incidence of malignancies in economically progressive  
231 countries.

232

## 233 **Discussion**

234

235 Cytomegalovirus is wreathed in mystery, a Gordian knot of biological knowledge and  
236 understanding [82]. Cytomegalovirus had evolved a peculiar virus-host symbiosis by a taciturn  
237 tactic that co-opts or escapes immune pathways in order to arrange its persistence in the host. A  
238 strong inverse statistical correlation between the incidence of cancer and CMV seroprevalence  
239 covers quite a swath of cancer spectra and territorial ranges. Although correlation does not  
240 categorically infer causation, this one may underscore a causal meaning. Besides the benefits of  
241 vaccine against CMV [53], it may suggest an impeded carcinogenesis by generating the CMV-  
242 induced tumoricidal T cells. CMV-based vaccine antagonistic to tumorigenesis would possibly  
243 result in a global regression of cancer. Comparing non-Whites to Whites, Cannon *et al.* [2]  
244 evaluated that CMV seroprevalence is consistently 20-30 percent points higher in the former  
245 than in the latter (summary PR=1.59, 95% CI=1.57-1.61). Some ethnic groups had CMV  
246 seroprevalence ~ 100% (Fig. 4A in [2]). In agreement with the current work, Fowler *et al.* [32]  
247 discounted the importance of ethnicity *per se* as a risk factor for the CMV infection. Similar  
248 conclusions were reached by Lantos *et al.* [54], Rook [55], and Dowd *et al.* [56,57].

249       Elimination of health disparities as a consequence of racial and ethnic differentials has  
250 been earlier recognized as important [47,48,83]. Socioeconomic disparities across race/ethnicity  
251 categories impact the level of CMV seroprevalence [83]. These may have biased lower estimates  
252 of cancer incidence at the time. Rather than the race/ethnic divides of themselves, we propose



253 CMV infection as an oncopreventer both of some hematologic malignancies at our clinic and of  
254 cancers worldwide [13]. Although CMV prevalence and education level, SES, and household  
255 income are associated with race and ethnicity, we speculate here that latent CMV infection is a  
256 fundamental cause underlying disparities in cancer incidence among race/ethnic groups in the  
257 U.S. and worldwide. The NHANES III data (the U.S. 2011-2012) report CMV prevalence  
258 (race/minority) in children 1-5 years of age as 37% among non-Hispanic other/multiracial, 31%  
259 among Hispanic, 15.9% among non-Hispanic Black, and 10.6% among non-Hispanic White  
260 ethnicities [33]. This data is inversely proportional to the incidence of all cancers (combined) in  
261 populations of these youths, indicative of a higher risk of cancer in CMV seronegative  
262 populations in the U.S. Also, unlike Hispanos, cancer rates in Cubans were comparable to non-  
263 Hispanic Whites, and Puerto Ricans and Cubans in Florida had rates of some solid cancers  
264 similar to non-Hispanic Whites despite the rates of these cancers being significantly lower in  
265 their countries of origin [22]. High incidence of cancer despite a high prevalence of CMV in the  
266 Inuit (Eskimos) may be a consequence of deficient T cell immunity in this ethnic population  
267 [23,24]. This is to be expected and, indeed, we found no global correlation between the incidence  
268 of Kaposi's sarcoma (mostly diagnosed in HIV-positives with compromised T cell immunity)  
269 and the prevalence of CMV (**Fig. 5**). CMV does not exert protection unless T cell immunity is  
270 functional.

271 Immigrants to U.S. experience decreasing incidence rates of cancer of infectious origin  
272 (hepatitis B virus, *Helicobacter pylori*, human papillomavirus) which are prevalent in their  
273 countries of origin. On the contrary, incidence rates of lung, breast, colorectal and prostate  
274 cancer have been on the rise despite remaining relatively low in the host nations [57-59].

275           The statistical analysis indicates an inverse correlation between CMV pervasiveness and  
276 the race-specific cancer incidence, a valuable hint at a possible oncoprotective effect of the  
277 pathogen. Previously, we speculated that CMV may confer a protection against B cell dyscrasias  
278 [13]. For example, there is a highly significant inverse link (**Fig. 1**. Spearman's  $\rho = -0.754$ ;  
279  $p < 0.001$ ) between all invasive cancers combined (both genders, all ages) and the country specific  
280 CMV prevalence profile across the mainland and sea-coastal regions of 73 countries (**Table 3**).  
281 We draw attention to a possible protective effect of the CMV infection as an unappreciated  
282 factor which may subvert oncogenesis.

283           We sought to grasp a better understanding of fluctuating incidence of new cancers among  
284 immigrant and established indigenous populations by consulting relevant reports. We envision  
285 that dilution of the prevalence of CMV, due to a progress in economic opportunity and an open  
286 access to competent medical patronage, may have resulted in increased cancer rates, indeed the  
287 epidemics of neoplasms [60]. Societal development aside, poverty remains a considerable  
288 medical concern [30,33,62,72]. Higher prostate and lung cancer rates are reported in established  
289 immigrant enclaves from Japan and China in the U.S. than are found in Japan and China,  
290 probably because improved SES in the U.S. (i.e. a better hygiene and improved medical care)  
291 with a consequent decline of CMV-mediated oncoprotection in the U.S. Also, individuals with  
292 low SES have higher antibody titers to CMV [31] and presumably more protection against  
293 cancer.

294           Crucially, CMV induces a specific, CMV-determined, T cell mediated antitumor effect in  
295 immunocompetent persons but fails in patients with inoperative T cell immune surveillance, like  
296 Kaposi sarcoma (**Fig. 5**). The role of CMV acquisition and the consequent T lymphocyte-  
297 specified inhibition of tumorigenesis may prove of importance in pre-empting various

298 malignancies, particularly those that can be detected at an early stage, such as breast and colon  
299 cancers.

300 Disaggregation (decomposition) of ethnic data in racial/multiethnic studies of cancer vs.  
301 rates of cancer cases combined and the CMV prevalence may, while quite information-rich, also  
302 be a source of bias due to unmeasured confounders or mediators [79]. Disaggregation may mask  
303 true biological linkages which would be more accented if global bulk of data were analyzed.  
304 Hoshiba *et al.* [61] examined the long-term (1980-1998) dynamics of CMV seroprevalence in  
305 pregnant women in Japan. Complement-fixing antibody and specific IgG antibody, as measured  
306 in sera, decreased gradually from 93.2% to 66.7%. CMV-IgG seropositive rates were 87.4% in  
307 1985, and 75.2% in 1996 to 1997. This provides a separate hint at a possible protective effect of  
308 CMV in this population. Of note, the incidence of cancer increased in parallel to decrease of  
309 CMV seropositivity in this Japanese population.

310 Most recent clinical research on colorectal [86] and bronchogenic cancer [87] confirm  
311 our evidence for the protection against cancer (**Table 3**) conferred by the CMV infection.

312

### 313 **Limitations**

314

315 Notable limitations of this work are difficult to lift. A major limitation are the data used from  
316 literature and reviews with varying collection periods which span numerous years (**Table 1**).

317 Disparate time scales occasionally underlie epidemiological studies of CMV in the low-income  
318 greater neighborhoods. There, the prevailing cultural settings and areas of higher deprivation  
319 (with highest prevalence of CMV infection) frequently overlap confounding the measured  
320 estimates (**Table 2**). Small numbers yield unstable prevalence rates. The roundabouts of more

321 disadvantageous neighborhoods, culturally and genetically distinct yet sharing remotely  
322 entangled backgrounds, often overlay and may affect accuracy and consistency of the estimates.  
323 Race and ethnicity were differently categorized by the investigators. Also, the options were  
324 defined by study participants. Marked disparity between African Americans and White patients  
325 might have resulted from different indication settings for CMV testing and in unknown  
326 proportions. Also, there still remains a possibility that, besides the CMV, there could have  
327 existed yet another causative tumor-suppressive agent. Consequently, formal adjudication of  
328 causality of statistical significance is not allowed by the cross-sectional nature of the data in  
329 many studies we used. Still, prospective studies suggest that causality could be at work behind  
330 significance of statistical tests; high degree of disparity between Blacks and Whites is unlikely to  
331 result solely as an artifact of testing biases. Pronounced variance in cancer incidence in AI/AN  
332 reported by Henley *et al.* [44] and Cronin *et al.* [45] are possibly due to improved financial  
333 attainment and hygiene, cleaning, and the privilege of preventive medical care although a drop of  
334 CMV prevalence in this populace has not yet been accurately registered. The New Zealand  
335 Islanders' CMV seropositivity estimates presented here (**Table 1**) may crudely parallel those in  
336 APIs who live on the U.S. Pacific Island Territories rather than that of API migrants living on the  
337 U.S. mainland. The estimates should be considered with circumspection and checked with regard  
338 to future vaccination strategies to fight CMV infection [64]. We have to take the results  
339 presented here within the context of these limitations.

340

## 341 **Conclusion**

342

343 Our work is the first to analyze demography of cancer incidence rates relative to the prevalence  
344 of CMV seropositivity both in U.S. race/ethnic groups and worldwide. A biological connection  
345 between the ubiquity of CMV and the incidence of neoplasms is predicated by a worldwide  
346 correlation between the two. This result may further inspire an oncologic initiative for  
347 development of the CMV-based antitumor vaccine constructs that could result in a significant  
348 reduction of human cancers the world over. Here, we aimed at just this.

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367 **Acknowledgements**

368 We credit the staff of the SEER and NHANES III who disseminated publicly visible data,  
369 essential for our analysis, and acknowledge the funding by the Ministry of Science,  
370 Technological Development and Innovation (MSTDI) of the Republic of Serbia. The funder did  
371 not play a role in the design of the study; the collection, analysis, and interpretation of the data;  
372 the writing of the manuscript; and the decision to submit the manuscript for publication.

373

374 **Authors' contributions**

375 Conceptualization: MJ, AK, TJ, MT-B.

376 Funding acquisition: TJ, BM, MT-B.

377 Writing — original draft preparation: MJ, AK, ID, OM.

378 Writing — review and editing: MJ, TJ, ID, OM, BM.

379 Investigation and Methodology: MJ, OM, AK.

380 Project administration, Supervision, Resources: TJ, AK, BM, and MT-B.

381 All authors read, critically revised, and approved the submission of the final version of the  
382 manuscript.

383

384 **Funding**

385 This work was supported by the Ministry of Science, Technological Development and  
386 Innovation (MSTDI) of the Republic of Serbia, grant number 200110. Any opinions, findings,  
387 conclusions or recommendations expressed in this material are those of the authors and do not  
388 necessarily reflect the views of MSTDI.

389

390 **Data availability statement**

391 All Data are available from the corresponding author upon reasonable request.

392

393 **DECLARATIONS**

394

395 **Ethics approval and consent to participate**

396 Not applicable.

397

398 **Conflict of interest and disclosures**

399 The authors declare that research was conducted in the absence of any commercial or financial

400 relationships that could be construed as a potential conflict of interest.

401

402

403

404

405

406

407

408

409

410

411

412

413 **References:**

414 1. Mocarski E, Shenk T, Griffiths P, Pass RF. Cytomegaloviruses. In: Knipe DM, Howley PM,  
415 Eds. *Field's Virology*, 6<sup>th</sup> ed. Philadelphia, PA, USA. Wolters Kluwer Lippincott Williams &  
416 Wilkins, **2013**:1960-2014.

417

418 2. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus prevalence and demographic  
419 characteristics associated with infection. *Rev Med Virol.* **2010**; 20:202-213. doi:  
420 10:1002/rmv.655. PMID: 20564615.

421

422 3. Pass RF, Zhang C, Evans A, et al. Vaccine prevention of maternal cytomegalovirus infection.  
423 *N Engl J Med.* **2009**; 360:1191-1199. doi: 10.1056/NEJMoa0804749. PMID: **19297572**.

424

425 4. Lasry S, Dény P, Asselot C, et al. Interstrain variation in the cytomegalovirus (CMV)  
426 glycoprotein B gene sequence among CMV- infected children attending six day care centers.  
427 *J Infect Dis.* **1996**; 174:606-609. doi: 10.1093/infdis/174.3.606. PMID: 8769620.

428

429 5. Staras SAS, Dollard SC, Radford KW, et al. Seroprevalence of cytomegalovirus infection in  
430 the United States, 1988-1994. *Clin Infect Dis.* **2006**; 43:1143-1151. doi:  
431 10.1086/508173. PMID: 17029132.

432 6. Kumar A, Coquard L, Pasquereau S, et al. Tumor control by human cytomegalovirus in a  
433 murine model of hepatocellular carcinoma. *Mol Ther Oncolytics.* **2016**; 3:16012. doi:  
434 10.1038/mt.2016.12. PMID: 27626063.



435

436 7. Herbein G, Nehme Zeina. Tumor control by cytomegalovirus: A door open for oncolytic  
437 virotherapy? *Mol Ther Oncolytics*. **2020**; 17:1-8. doi: [org/10.1016/J.omto.2020.03.004](https://doi.org/10.1016/J.omto.2020.03.004).  
438 PMID: 32300639.

439

440 8. Massara L, Khairallah C, Yared N, et al. Uncovering the anticancer potential of murine  
441 cytomegalovirus and human colon cancer cells. *Mol Ther Oncolytics*. **2020**; 16:250-261. doi:  
442 [org/10.1016/j.omto.2020.01.007](https://doi.org/10.1016/j.omto.2020.01.007). PMID:32140563.

443

444 9. Éboli LPdCB, Tannuri ACA, Tannuri U. Seropositivity for cytomegalovirus and PCR-EBV  
445 monitoring: Protective factors for posttransplant lymphoproliferative disorder in pediatric  
446 liver transplant. *Pediatr Transplant*. **2022**; 26:e14226. doi: [10.1111/petr.14226](https://doi.org/10.1111/petr.14226). PMID:  
447 35037358.

448

449 10. Litjens NHR, Wagen Lvd, Kuball J, et al. Potential beneficial effects of cytomegalovirus  
450 infection after transplantation. *Front Immunol*. **2018**; 1;9:389. doi:  
451 [10.3389/fimmu.2018.00389](https://doi.org/10.3389/fimmu.2018.00389). PMID: 29545802.

452

453 11. Yu S-C, Ko K-Y, Teng S-C, et al. A clinicopathological study of cytomegalovirus  
454 lymphadenitis and tonsillitis and their association with Epstein-Barr virus. *Infect Dis Ther*.  
455 **2021**; 10:2661-2675. doi: [org. 10.1007/s40121-021-00528-1](https://doi.org/10.1007/s40121-021-00528-1). PMID: 34623624.

456

- 457 12. Bigley AB, Baker FL, Simpson RJ. Cytomegalovirus: an unlikely ally in the fight against  
458 blood cancers? *Clin Exp Immunol.* **2018**; 193:265-274. doi: 10.1111/cei.13152. PMID:  
459 29737525.
- 460
- 461 13. Janković M, Knežević A, Todorović M, et al. Cytomegalovirus infection may be  
462 oncoprotective against neoplasms of B-lymphocyte lineage: single-institution experience and  
463 survey of global evidence. *Virologia.* **2022**; 19:155. doi: 10.1186/s12985-022-01884-1. PMID:  
464 36171605.
- 465
- 466 14. Kann L, Kinchen SA, Williams BI, et al. Youth risk behavior surveillance – United States,  
467 1977. *MMWR CDC Surveill Summ.* **1998**; 47:1-89. PMID: 9719790.
- 468
- 469 15. Halpern CT, Hallfors D, Bauer DJ, et al. Implications of racial and gender differences in  
470 patterns of adolescent risk behavior for HIV and other sexually transmitted diseases.  
471 *Perspect Sex Reprod Health.* **2004**; 36:239-247. doi: 10.1363/psrh.36.239.04. PMID:  
472 15687082.
- 473
- 474 16. Markides K, Coreil J. The health of Hispanics in the southwestern United States: an  
475 epidemiologic paradox. *Public Health Rep.* **1986**; 101:253-265. PMID: 3086917.
- 476
- 477 17. Chen H, Wu AH, Wang S, et al. Cancer Mortality Patterns by Birthplace and Generation  
478 Status of Mexican Latinos: The Multiethnic Cohort. *J Natl Cancer Inst.* **2022**; 114:959-968.  
479 doi: 10.1093/jnci/djac078. PMID: 35404450.

- 480
- 481 18. Graham C, Metayer C, Morimoto LM, et al. Hispanic Ethnicity Differences in Birth  
482 Characteristics, Maternal Birthplace, and Risk of Early-Onset Hodgkin Lymphoma: A  
483 Population-Based Case-Control Study. *Cancer Epidemiol Biomarkers Prev.* **2022**; 31:1788-  
484 1795. doi: 10.1158-9965.EPI-22-0335. PMID: 35709749.
- 485
- 486 19. Oluwasegun A, Abodurnin F, Nasef KE, et al. IBCL-462 Racial-Ethnicity Gap in Survival  
487 among Non-Hodgkin Lymphoma Survivors. *Clin Lymphoma Myeloma Leuk.* **2022**; Suppl. 2:  
488 S393. doi: 10.1016/S2152-2650(22)01566-X. PMID: 36164118.
- 489
- 490 20. Diaz Ayllon KE, Mowitz M. The Hispanic Paradox among Texas Newborns. *Pediatrics.*  
491 **2019**; 144 (2\_Meeting Abstract): 690. doi: 10.1542/peds.144.2MA7.690.
- 492
- 493 21. Blansky D, Fazzari M, Mantzaris I, et al. Racial and ethnic differences in all-cause mortality  
494 among Hispanics diagnosed with follicular lymphoma and chronic lymphocytic leukemia in  
495 Bronx, NY. *Cancer Causes Control.* **2022**; 33:137-147. doi: 10.1007/s10552-021-01507-0.  
496 PMID: 34677741.
- 497
- 498 22. Pinheiro PS, Sherman RL, Trapido EJ, et al. Cancer incidence in first generation U.S.  
499 Hispanics: Cubans, Mexicans, Puerto Ricans, and new Latinos. *Cancer Epidemiol*  
500 *Biomarkers Prev.* **2009**; 18:2162-2169. doi: 10.1158/1055-9965.EPI-09-0329. PMID:  
501 19661072.
- 502

- 503 23. Nicolle LE, Minuk GY, Postl B, et al. Cross-sectional seroepidemiologic study of the  
504 prevalence of cytomegalovirus and herpes simplex virus infection in a Canadian Inuit  
505 (Eskimo) community. *Scand J Infect Dis.* **1986**; 18:19-23. doi:  
506 10.3109/00365548609032301. PMID: 3008308.
- 507
- 508 24. Preiksaitis JK, Larke RP, Froese GJ. Comparative seroepidemiology of cytomegalovirus  
509 infection in the Canadian Arctic and an urban center. *J Med Virol.* **1988**; 24:299-307. doi:  
510 10.1002/jmv.1890240307. PMID: 2835426.
- 511
- 512 25. Geris JM, Spector LG, Pfeiffer RM, et al. Cancer risk associated with cytomegalovirus  
513 infection among solid organ transplant recipients in the United States. *Cancer.* **2022**;  
514 128:3985-3994. doi: 10.1002/cncr.34462. PMID: 36126024.
- 515
- 516 26. SEER 12 (2000 Census Data). U.S. Bureau of Census **2020**, DEC Registrating data (PL 94-  
517 171), Tables P1, P2.
- 518
- 519 27. Incidence - North American Association of Central Cancer Registries, ©**2020** American  
520 Cancer Society, Inc. Surveillance Research.
- 521
- 522 28. Zuhair M, Smit GSA, Wallis G, et al. Estimation of the worldwide seroprevalence of  
523 cytomegalovirus: a systematic review and meta-analysis. *Rev Med Virol.* **2019**; 29(3):e2034.  
524 doi: 10.1002/rmv.2034. PMID: 30706584.
- 525

- 526 29. Simanek AM, Dowd JB, Pawelec G, et al. Seropositivity to cytomegalovirus, inflammation,  
527 all-cause and cardiovascular disease-related mortality in the United States. *PLoS One*. **2011**;  
528 6(2):e16103. doi: [10.1371/journal.pone.0016103](https://doi.org/10.1371/journal.pone.0016103). PMID: 21379581.
- 529
- 530 30. Schwartz HN, Fried PL, Xue, Q-L, et al. Chronic cytomegalovirus infection and  
531 inflammation are associated with prevalent frailty in community-dwelling older women. *J*  
532 *Am Geriatr Soc*. **2005**; 53:747-754. doi: 10.1111/j1532-5415.2005.53250.x. PMID: 15877548.
- 533
- 534 31. Noppert GA, AielloAE, O’Rand AM, et al. Race/ethnic and educational disparities in the  
535 association between pathogen and educational disparities in the association between  
536 pathogen burden and laboratory-based cumulative deficit index. *J Racial Ethn Health*  
537 *Disparities*. **2020**; 7:99-108. doi: 10.1007/s40615-019-00638-0. PMID: 31642044.
- 538
- 539 32. Fowler K, Mucha J, Neuman M, et al. A systematic literature review of the global  
540 seroprevalence of cytomegalovirus: possible implications for treatment, screening, and  
541 vaccine development. *BMC Public Health*. **2022**; 22:1659 (15pp). doi: 10.1186/s12889-022-  
542 13971-7. PMID: 36050659.
- 543
- 544 33. Bate SL, Dollard SC, Cannon MJ. Cytomegalovirus seroprevalence in the United States: The  
545 national Health and Nutritional Examination Surveys, 1988-2004. *Clin Infect Dis*. **2010**; 50:  
546 1439-1447. doi: 10.1086/652438. PMID: 20426575.
- 547

- 548 34. National Center for Health Statistics, Plan and Operation of the Third National Health and  
549 Nutrition Examination Survey 1988-94. Series 1: programs and collection procedures. Vital  
550 Health Stat I **1994**; 32:1-407. Available at:  
551 [http://www.cdc.gov/nchs/data/series/sr\\_01/sr01\\_032.pdf](http://www.cdc.gov/nchs/data/series/sr_01/sr01_032.pdf).  
552
- 553 35. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics  
554 (NCHS). National Health and Nutrition Examination Survey: Documentation **2009**,  
555 [http://www.cdc.gov/nchs/nhanes/vardemo\\_b.htm](http://www.cdc.gov/nchs/nhanes/vardemo_b.htm). Accessed 12 April 2010 by Bate *et al.*  
556 **2010** [33].  
557
- 558 36. U. S. Census Bureau. Census 2000 Summary File 2, Table PCT27, United States,  
559 Washington, DC: U.S. Census Bureau, **2008**.  
560
- 561 37. Centers for Disease Control and Prevention, Analytic and Reporting Guidelines. The Third  
562 National Health and Nutrition Examination Survey. NHANES III (1988-1994). Hyattsville,  
563 MD: National Center for Health Statistics, **1996**. Available at:  
564 <https://www.cdc.gov/nchs/data/nhanes/nhanes3/nh3gui.pdf>.  
565
- 566 38. Surveillance, Epidemiology, and End Results (SEER) Program. *SEER\*Stat Database:*  
567 *Incidence-SEER 18 Registries Research Data + Hurricane Katrina Impacted Louisiana*  
568 *Cases, November 2020 Submission (2000-2018), Katrina/Rita Population Adjustment.*  
569 *Linked to County Attributes-Total U.S., 1969-2018 Counties*. National Cancer Institute,

570 Division of Cancer Control and Population Sciences, Surveillance Research Program,  
571 Surveillance System Branch; **2021**.

572

573 39. Surveillance Research Program. *SEER\*Explorer: an interactive website for SEER cancer*  
574 *statistics*. National Cancer Institute; **2021**. Accessed April 15, 2021.

575 [seer.cancer.gov/explorer/](https://seer.cancer.gov/explorer/)

576

577 40. Surveillance, Epidemiology, and End Results (SEER) Program. *SEER\*Stat Database:*

578 *Incidence-SEER Research Limited-Field Data with Delay-Adjustment, 21 Registries,*

579 *Malignant Only, November 2020 Submission (2000-2018)-Linked to County Attributes-Time*

580 *Dependent (1990-2018) Income/Rurality; 1969-2019 Counties*. National Cancer Institute,

581 Division of Cancer Control and population Sciences, Surveillance Research Program; **2021**.

582

583

584 41. Surveillance, Epidemiology, and End results (SEER) Program. *SEER\*Stat Database: North*

585 *American Association of Central Cancer Registries (NAACCR) Incidence Data-Cancer in*

586 *North America Analytic File, 1995-2018, With Race/Ethnicity; Custom File With County;*

587 *American Cancer Society Facts and Figures Projection Project (which includes data from*

588 *the Center for Disease Control and Prevention's national Program of Cancer Registries, the*

589 *Canadian Council of Cancer Registries' Provincial and territorial Registries, and the*

590 *national Cancer Institute's SEER Registries, certified by the NAACCR as meeting high-*

591 *quality incidence data standards for the specified time periods*). National Cancer Institute,

592 Division of Cancer Control and Population Sciences, Surveillance Research Program; **2021**.

- 593
- 594 42. Sherman R, Firth R, Charlton M, *et al.* Eds. *Cancer in North America: 2014-2018. Volume*  
595 *Two: Registry-Specific Cancer Incidence in the United States and Canada.* North American  
596 Association of Central Cancer registries, Inc; **2021**.
- 597
- 598 43. Zhu J, Davidson M, Leinoven M, *et al.* Prevalence and persistence of antibodies to herpes  
599 viruses, Chlamydia pneumonia, and Helicobacter pylori in Alaskan Eskimos: the  
600 GOCADAN Study. *Clin Microbiol Infect* **2006**; 12:118-122. doi: 10.1111/j.1469-  
601 0691.2005.01319.x. PMID: 16441448.
- 602
- 603 44. Henley SJ, Ward EM, Scott S, *et al.* Annual Report to the Nation on the status of Cancer,  
604 *Part I: national Cancer Statistics.* *Cancer* **2020**; 126:2225-2249. doi: 10.1002/cncr.32802.  
605 PMID: 32162336.
- 606
- 607 45. Cronin KA, Scott S, Firth AU, *et al.* Annual report to the nation on the status of cancer, part  
608 1: National cancer statistics. *Cancer* **2022**; 128:4251-4284. doi: 10.1002/cncr.34479. PMID:  
609 36301149.
- 610
- 611 46. Miller BA, Chu KC, Hankey BF, *et al.* Cancer incidence and mortality patterns among  
612 specific Asian and pacific Islander populations in the U. S. [Appnxs. 1&2, Tabs. 2-4],  
613 *Cancer Causes Control* **2008**; 19:227-256. doi: 10.1007/s10662-007-9088-3. PMID:  
614 18066673.
- 615



- 616 47. U. S. Department of Health and Human Services (HHS) (2000). Healthy people **2010**, 2<sup>nd</sup> ed.  
617 With understanding and improving health and objectives for improving health, 2 vols. U. S.  
618 Government Printing Office, Washington.  
619
- 620 48. Rubicz R, Zhu J, Laston S, et al. Statistical genetic analysis of serological measures of  
621 common, chronic, infections in Alaska Native participants in the GOCADAN study. *Genet*  
622 *Epidemiol* **2013**; 37:751-757. doi: 10.1002/gepi.21745. PMID: 23798484.  
623
- 624 49. U. S. Department of Health and Human Services (HHS), National Institutes of Health **2007**.  
625 Strategic research plan to reduce and ultimately eliminate health disparities, vol. 1. Available  
626 from [http://ncmhd.nih.gov/our\\_programs/strategic/index.asp](http://ncmhd.nih.gov/our_programs/strategic/index.asp). Accessed Sept, 2007.  
627
- 628 50. National Cancer Institute SEER Limited-Use Data. Available from URL:  
629 <http://www.seer.cancer.gov/data/>. Tables I-20 and I-21, Figures I-13 and I-14.  
630
- 631 51. Sung H, Ferley J, Siegel, RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of  
632 Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J Clin*  
633 **2021**; 71:209-249. doi: 10.3322/caac.21660. PMID: 33538338.  
634
- 635 52. Conde-Glez C, Lazcano-Ponce E, Rojas R, et al. Seroprevalences of varicella-zoster virus,  
636 herpes simplex virus and cytomegalovirus in a cross-sectional study in Mexico. *Vaccine*  
637 **2013**; 31:5067-5074. doi: 10.1016/j.vaccine.2013.08.077. PMID: 24021305.  
638

- 639 53. Furman D, Jojic V, Sharma S, et al. Cytomegalovirus infection enhances the immune  
640 response to influenza. *Sci. Transl. Med.* **2015**; 7:281ra43. doi: 10.1126/scitranslmed.aaa2293.  
641 PMID: 25834109.  
642
- 643 54. Lantos PM, Hoffman K, Permar SR, et al. Geographic Disparities in Cytomegalovirus  
644 Infection During Pregnancy. *Racial Ethn Health Disparities.* **2018**; 5:782-786. doi:  
645 10.1007/s40616-017-0423-4. PMID: 28840519.  
646
- 647 55. Rook GAW. Evolution, the immune system, and the health consequences of socioeconomic  
648 inequality. *mSystems.* **2022**; 7: e0143821. doi: 10.1128/msystems.01438-21. PMID:  
649 35285679.  
650
- 651 56. Dowd JB, Palermo TM, Aiello AE. Family poverty is associated with cytomegalovirus  
652 antibody titers in U.S. children. *Health Psychol.* **2012**; 31:5-10. doi: 10.1037/a0025337.  
653 PMID: 21895.  
654
- 655 57. Dowd JB, Aiello AE, Alley DE. Socioeconomic disparities in the seroprevalence of  
656 cytomegalovirus infection in the U.S. population: NHANES III. *Epidemiol Infect.* **2009**;  
657 137:58-65. doi: 10.1017/S0950268808000551. PMID: 18413004.  
658
- 659 58. Lin L, Noone A-M, Gomez SL, et al. Cancer incidence trends among native Hawaiians and  
660 other pacific Islanders in the United States, 1990-2008. *J Natl Cancer Inst.* **2013**; 105:1086-  
661 1095. doi:10.1093/jnci/djt156. PMID: 23878354.

- 662
- 663 59. Gomez SL, Noone A-M, Lichtensztajn DY, et al. Cancer incidence trends among Asian-
- 664 American population in the United States, 1990-2008. *J Natl Cancer Inst.* **2013**; 105:1096-
- 665 1110. doi: 10.1093/jnci/djt157. PMID: 23878350.
- 666
- 667 60. Kolonel L, Wilkens LR, Migrant Studies. In: Eds. Schottenfeld D, Fraumeni JF Jr. *Cancer*
- 668 *Epidemiology and Prevention*, 3rd ed. New York: Oxford University Press. **2006**; 189-201.
- 669
- 670 61. Bosetti C, Levi F, Ferlay J, et al. Incidence and mortality from non-Hodgkin Lymphoma in
- 671 Europe: The end of epidemics? *Int J Cancer.* **2008**; 123:1917-1923. doi:
- 672 [doi.org/10.1002/ijc.23722](https://doi.org/10.1002/ijc.23722). PMID: 18688859.
- 673
- 674 62. Hoshiba T, Asamoto A, Yabuki Y. Decreasing seropositivity of cytomegalovirus of pregnant
- 675 women in Japan. [Article in Japanese] *Nihon Rinsho.* **1998**; 56:193-196. PMID: 9465689.
- 676
- 677 63. Plotkin SA. Preventing infection by human cytomegalovirus. *J Infect Dis.* (Suppl. 1 Article).
- 678 **2020**; 221:S 123-127. doi:10.1093/infdis/jiz448. PMID: 32134484.
- 679
- 680 64. Griffiths P, Plotkin S, Mocarski E, et al. Desirability and feasibility of a vaccine against
- 681 cytomegalovirus. *Vaccine.* **2013**; 31 (Suppl. 2): B197-B203. doi:
- 682 10.1016/j.vaccine.2012.10.074. PMID: 23598482.
- 683

- 684 65. Espey DK, Wu X-C, Swan J, et al. Annual report to the nation on the status of Cancer, 1975-  
685 2004, Featuring Cancer in American Indians and Alaska Natives. *Cancer* **2007**; 110:2119-  
686 2152. doi:10.1002/cncr.23044. PMID: 17939129.
- 687
- 688 66. Nikitin Yu P, Boichenko NS, Astakhova TI, et al. Cancer in Russian Inuit. *Acta Oncologica*.  
689 **1996**; 35:617-619. doi:10.3109/02841869609096995. PMID: 8813070.
- 690
- 691 67. Jenkins AL, Gyorkos TW, Culman KN, et al. An overview of factors influencing the health  
692 of Canadian Inuit infants, *Int J Circumpolar Health*. **2003**; 62: 17-39,  
693 doi:10.3402/ijch.v62i1.17526. PMID: 12725339.
- 694
- 695 68. Goplani SV, Janiz AE, Martinez SA, et al. Trends in cancer incidence among American  
696 Indians and Non-Hispanic Whites in the United States, 1999-2015. *Epidemiology*. **2020**;  
697 31:205-215. doi: 10.1097/EDE.0000000000001140. PMID: 31764279.
- 698
- 699 69. Langer-Gould A, Wu J, Lucas R, et al. Epstein-Barr virus, cytomegalovirus, and multiple  
700 sclerosis susceptibility. A Multiethnic Study. *Neurology*. **2017**; 89:1330-1337. doi:  
701 10.1212/WNL.0000000000004412. PMID: 28855411.
- 702
- 703 70. Melkonian SC, Jim MA, Haverkamp D, et al. Disparities in cancer incidence and trends  
704 among American Indians and Alaska Natives in the United States, 2010-2015. *Cancer*  
705 *Epidemiol Biomarkers Prev*. **2019**; 28:1604-1611. doi: 10.1058/1055-9965.EPI-19-0288.  
706 PMID: 31575554.

- 707
- 708 71. Leung J, Kennedy JL, Haberling DL, et al. Congenital CMV-Coded Diagnosis Among  
709 American Indian and Alaska Native Infants in the United States, 2000-2017. *J Immigr Minor*  
710 *Health.* **2020**;22(5):1101-1104. doi:10.1007/s10903-020-01024-3
- 711
- 712 72. Lanzieri TM, Kruszon-Moran D, Amin MM, et al. Seroprevalence of cytomegalovirus  
713 among children 1 to 5 years of age in the United States from the National Health and  
714 Nutrition Examination Survey of 2022 to 2012. *Clin Vaccine Immunol.* **2015**; 22:245-247.  
715 doi: 10.1128/CVI.00697-14. PMID: 25520150.
- 716
- 717 73. Kelly JJ, Lanier AP, Alberts S, et al. Differences in cancer incidence among Indians in  
718 Alaska and New Mexico and U.S. Whites, 1993-2002. *Cancer Epidemiol Biomarkers Prev.*  
719 **2006**; 15:1515-1519. doi: 10.1058/1055-9965.EPI-05-0454. PMID: 16896042.
- 720
- 721 74. US Census Bureau. *The American community – Pacific Islanders 2004*. American  
722 community survey reports. <http://www.census.gov/prod/2007pubs/acs-06.pdf>. Accessed  
723 September 1, **2012**.
- 724
- 725 75. National Cancer Institute, Surveillance Epidemiology and End Result.  
726 <http://seer.cancer.gov/about>. Accessed September 1, **2012**.
- 727
- 728 76. Haverkamp D, Melkonian SC, Jim MA. Growing disparity in the incidence of colorectal  
729 cancer among non-Hispanic American Indian and Alaska native populations – United States,

730 2013-2017. *Cancer Epidemiol Biomarkers Prev.* **2021**; 30:1799-1806. doi: 10.1158/1055-

731 9965.EPI-21-0343. PMID:34341050.

732

733 77. Wiggins CL, Espey DK, Wingo PA, et al. Cancer among American Indians and Alaska

734 Natives in the United States, 1999-2004. *Cancer Supplement.* **2008**; 113:1142-1152. doi:

735 10.1002/cncr.23734. PMID: 18720375.

736

737 78. Badami KG, McQuilkan Bickerwstaffe S, Wells JE, et al. Cytomegalovirus seroprevalence

738 and 'cytomegalovirus-safe' seropositive blood donors. *Epidemiol Infect* **2009**; 137:1776-

739 1780. doi: 10.1017/90950268809990094. PMID: 19480727.

740

741 79. Zamora SM, Pinheiro PS, Lin Gomez S, et al. Disaggregating Hispanic American cancer

742 mortality burden by detailed ethnicity. *Cancer Epidemiol Biomarkers Prev.* **2019**; 28:1353-

743 1363. doi: 10.1158/1055-9965. PMID: 31147314.

744

745 80. Swan J, Edwards BK. Cancer rates among American Indians and Alaska Natives. *Cancer.*

746 **2003**; 98:1262-1277. doi: 10.1002/cncr.11633. PMID: 12973851.

747

748 81. Crocombe RG (2007). *Asia in the Pacific Islands: Replacing the West.* (p.13) Suva IPS

749 Publications, University of the South Pacific. Institute of Pacific Studies. [ISBN 978 982 02](#)

750 [0388 4.](#)

751

- 752 82. Goodrum F. Human cytomegalovirus latency: approaching the Gordian Knot. *Annu Rev*  
753 *Virol.* **2016**; 3:333-337. doi:10.1146/annurev-virology-110615-042422. PMID: 27501258.  
754
- 755 83. Griffiths P, Babunian C, Ashby D. The demographic characteristics of pregnant women  
756 infected with cytomegalovirus. *Int J Epidemiol.* **1985**; 14:447-452. doi: 10.1093/ije/14.3.447.  
757 PMID: 2997052.  
758
- 759 84. Becker TM, Magder L, Harrison HR, et al. The epidemiology of infection with the human  
760 herpesviruses in Navajo children. *Am J Epidemiol.* **1988**; 127:1071-1078. doi:  
761 10.1093/oxfordjournals.aje.a114883. PMID: 2833854.  
762
- 763 85. Sorem KA. Cancer incidence in Zuni Indians of New Mexico. *Yale J Biol Med.* **1985**; 58:  
764 489-496. PMID: 4090535.  
765
- 766 86. Nagel B, Frankel L, Ardeljan A, Cardeiro M, Rashid S, Takabe K, Rashid OM. The  
767 Association of Human Cytomegalovirus Infection and Colorectal Cancer: A Clinical  
768 Analysis. *World J Oncol.* **2023**;14(2):119-124. doi: 10.14740/wjon1565. Epub 2023 Mar 24.  
769 PMID: 37188037.  
770
- 771 87. Rashid S, Ardeljan A, Frankel LR, Cardeiro M, Kim E, Nagel BM, Takabe K, Rashid O.  
772 Human Cytomegalovirus (CMV) Infection Associated With Decreased Risk of Bronchogenic  
773 Carcinoma: Understanding How a Previous CMV Infection Leads to an Enhanced Immune  
774 Response Against Malignancy. *Cureus.* **2023**;15(4):e37265. doi: 10.7759/cureus.37265. PMID:  
775 37162767.

776 **Legends to tables and figures:**

777

778 **Table 1.** Race/ethnicity-specific epidemiology of new cancer cases [SEER 12 (2015-2019) +  
779 NPCR age-adjusted participants <sup>a</sup>] and cytomegalovirus seroprevalence [NHANES III]  
780 (accessible data are selected for comparative purposes).

781

782 **Table 2.** Age-standardized, delay-adjusted incidence rates <sup>a</sup> for the most common cancers <sup>b</sup> (all  
783 anatomical sites combined, both genders, all ages) among the racial/ethnic groups <sup>c</sup> (2012-2016)<sup>d</sup>  
784 in the United States compared to their CMV seropositivity (%).

785

786 **Table 3.** Correlation between worldwide incidence rate (100,000 population/year) of individual  
787 malignant tumors across 73 countries [WHO's GLOBOCAN database (51)] and the country  
788 specific prevalence of CMV seropositivity rates (28).

789

790 **Figure 1.** New cancer (all invasive types combined, all sites)/10<sup>5</sup> country specific population  
791 plotted against CMV seropositivity rates [28]. Inverse relationship over 73 countries [51]  
792 (Spearman's  $\rho = -0.732$ ;  $p < 0.001$ ), signals a globally pervasive protection against tumorigenesis  
793 driven by CMV.

794

795 **Figure 2.** Malignant melanoma/10<sup>5</sup> population vs. country specific prevalence of CMV  
796 seropositivity [28] correlate strongly inversely (Spearman's  $\rho = -0.763$ ;  $p < 0.001$ ) across 73  
797 countries [51].

798



799 **Figure 3.** Kidney cancer/ $10^5$  population and the country specific CMV seroprevalence [28] are  
800 strongly and inversely connected across 73 countries [51] (Spearman's  $\rho = -0.754$ ;  $p < 0.001$ ).

801

802 **Figure 4.** Incidence of breast cancer/ $10^5$  population and the country specific CMV seropositivity  
803 rate [28] spanning 73 continent-wide countries [51]. An inverse (protective) association is  
804 suggested by Spearman's  $\rho = -0.719$ ;  $p < 0.001$ .

805

806 **Figure 5.** Incidence of Kaposi's sarcoma (mostly in HIV+ people)/ $10^5$  population and the  
807 specific country-level rate of CMV seropositivity [28] covering 73 countries [51] are not  
808 correlated (Spearman's  $\rho = -0.007$ ;  $p = 0.953$ ).

809

810

811

812

813

814

815

816

817

818

819

820

821

**Table 1.** Race/ethnicity-specific epidemiology of new cancer cases [SEER 12 (2015-2019) + NPCR age-adjusted participants<sup>a</sup>] and cytomegalovirus seroprevalence [NHANES III].

	New Cancer Rates (all invasive types and stages of cancers)				Cytomegalovirus prevalence			
Characteristics	Cancers Combined (95% CI)	Study date (period)	N <sup>b</sup>	Citations (registries and first authors)	% (95% CI)	Study date (period)	N	Citations (registries, first author)
<i>Demographic division (by gender, race/ethnicity)</i>								
<b>Men, all ages</b>								
All races/ethnicities (incl. Hispanic)	476.5 (475.0-478.0)	2019	5 541	SEER+NPCR, ACS, Inc. SR	54 <sup>c</sup> (52-56)	1988- 1994	6 843	NHANES III, Bate (2010) [33]
	487.9 (487.4-488.4)	2020	99% popul.	Henley (2020) [44]				
<i>Demographic stratification</i>								
White (Non-Hispanic) <sup>d</sup>	469.3 (468.0-502.7)	2019	8 412	ACS, Inc. SR, SEER+ NPCR	51.2 (49.2-53.2)	1988- 1994	8 212	NHANES III, Staras (2006) [5]
	587	1998-		Miller (2008) [46]				

	(585.6-588.5)	2002						
	444.5 (502.4-503.6)	2012- 2016	99% popul.	Henley (2020) [44]				
Black (non-Hispanic)	412.7 (449.6-455.8)	2019	6 784	SEER+NPCR ACS, Inc. SR	75.8 (74.4-76.9)	1988- 1994	6 228	NHANES III Staras (2006) [5]
	547.6 (545.8-549.3)	2012- 2016	99% popul.	Henley (2020) [44]				
Asian/Pacific Islander, API (Non-Hispanic)	285.4	2008- 2017	25 693	USCS	93.3 (90.1-95.7)	2003- 2006	342	Badami (2009) [78] <sup>e</sup>
	303.7 (301.0-306.5)	1990- 2014	16 062	13 SEER, Miller (2008) [46]				
	296.5 (294.7-298.3)	2012- 2016	99% popul.	Henley (2020) [44]				
American Indian/Alaska Native, AI/AN (Non- Hispanic)	412.3 (395.4-429.7)	2019	1 584	SEER+NPCRACS, Inc. SR	89 (85-93)	2005- 2006	274	Zhu (2006) [43]
	259			Swan (2003) [80]	>80	~12-yr apart	338	Rubicz (2013) [48]
	420.0 412.6-427.5)	2012- 2016	99% popul.	Henley (2020) [44]				
Hispanic (any race) <sup>f</sup>	350.0 (347.9-355.2)	2019	6 296	SEER+NPCR ACS, Inc. SR	77.6 (75.8-79.4)	1988- 1994	4 921	NHANES III, Bate (2010) [33]
	377.8 (376.4-379.2)	2012- 2016	99% popul.	Henley (2020) [44]				

United Mexican States <sup>g</sup> (Mexico)	139.7	2020	37 412	GLOBOCAN IARC database [WHO] (2021) [51]	83 (mean, %)  91.3  90.7 (88.4-93.0)	1980- 2016  2000- 2020  2010	262 epid. studies  29 epid. studies  20 000 <sup>b</sup>	Zuhair (2019) [28]  Fowler (2022) [32]  Conde-Glez (2013) [52]
<b>Women, all ages</b>								
All races/ethnicities (incl. Hispanic)	416.4 (415.0-417.7)	2000- 2019	7 237	SEER 2015- 2019+NPCR, ACS, Inc. SR	64 (61-66)	1988- 1994	7 695	NHANES III, Bate (2010) [33]
	421.4 (421.1-421.8)	2012- 2016	99% popul.	Henley (2020) [44]				
<i>Demographic stratification</i>								
White (Non-Hispanic)	445.0 (443.2-446.8)	2000- 2019	8 212	SEER 2015- 2019+NPCR, ACS, Inc. SR	51 (49-53)	1988- 1994	6 296	NHANES III, Staras (2006) [5]
	444.5 (444.0-445.0)	2012- 2016	99% popul.	Henley (2020) [44]				
Black (Non-Hispanic)	405.9 (402.0-409.8)	2000- 2019	2.9x10 <sup>5</sup>	SEER 2015- 2019+NPCR, ACS, Inc. SR	70.6 (68.5-72.8)	1988- 1994	6 228	NHANES III Staras (2006) [5]
	412.8	2012- 2016	99%	Henley (2020) [44]				

	(411.6-414.1)	2016	popul.					
API (Non-Hispanic)	312.6 (308.9-316.4)	2000-2003	1.2x10 <sup>5</sup>	SEER 2015-2019+NPCR, ACS, Inc. SR, Miller (2008) [46]	NR	2003-2006	NR	Badami (2009) [65]
	295.7 (294.1-297.3)	2012-2016	99% popul.	Henley (2020) [44]				
AI/AN (Non-Hispanic)	414.9 (392.0-438.8)	2000-2019	6 260	SEER 2000-2003+NPCR, ACS, Inc. SR	93.5 (3.0±1.1)	2005-2006	336	Zhu (2006) [43]
four leading AI/AN cancers	230	1999-2015	61 461	CDC, Goplani (2020) [68]	>80	~12-yr apart	463	Rubicz (2013) [48]
AI/AN (non-Hispanic)	391.9 (385.8-398.2)	2012-2016	99% popul.	Henley (2020) [44]				
Hispanic (any race)	346.3 (343.4-349.3)	2000-2019	>5x10 <sup>4</sup>	SEER 2000-2003+NPCR, ACS, Inc. SR	72.7	2011-2014	136	Langer-Gould (2017) [69]
	333.6 (332.5-334.8)	2012-2016	99% popul.	Henley (2020) [44]	77.6 (75.8-79.4)	1988-1994	4 921	Bate (2010) [33]
United Mexican States <sup>g</sup> (Mexico)	142.4	2020	43 281	GLOBOCAN, IARC database [WHO] (2021) [51]	83 (mean, %)	1980-2016	262 epid. studies	Zuhair (2019) [28]
					90.1	2006	20 000	Conde-Glez (2013) [52]

All cancers combined (both genders, all ages)								
All races/ethnicities (incl. Hispanic)	440.1 (439.1-441.0)	2000- 2019	12 778	SEER 2000- 2003+NPCR, ACS, Inc. SR	48	2000- 2020	29 epid. studies	Fowler (2022) [32]
	447.9 (447.6-448.2)	2012- 2016	99% popul.	Henley (2020) [44]	58.9 (57.1-60.4)	1988- 1994	21 639	NHANES III, Staras (2006) [5]
<i>Demographic stratification</i>								
White (Non-Hispanic) (referent)	469.3 (468.0-470.6)	2000- 2019	2.7x10 <sup>6</sup>	SEER 2000- 2003+NPCR, ACS, Inc. SR	46.35	1988- 1994	8 212	NHANES III Staras (2006) [5]
	467.5 (447.6-448.2)	2012- 2016	99% popul.	Henley (2020) [44]				
Black (Non-Hispanic)	452.7 (449.6-455.8)	2000- 2019	4.2x10 <sup>5</sup>	SEER 2000- 2003+NPCR, ACS, Inc. SR	75.8 (74.4-76.9)	1988- 1994	6 228	NHANES III Staras (2006) [5]
	466.9 (465.8-467.9)	2012- 2016	99% popul.	Henley (2020) [44]				
API (Non-Hispanic)	303.7 (301.0-306.5)	2000- 2019	2.7x10 <sup>5</sup>	SEER 2000- 2003+NPCR, ACS, Inc. SR	93.3 (90.1-95.7)	2003- 2008	342	Badami (2009) [78]
	293.6 (292.5-294.8)	2012- 2016	99% popul.	Henley (2020) [44]				

AI/AN (Non-Hispanic)	412.3 (395.4-429.7)	2000-2019	11 696	SEER 2000-2003+NPCR, ACS, Inc. SR	92	2005-2006	673	Zhu (2006) [43]
Four leading AI/AN cancers	239	2012-2016		Swan (2003) [80]	>80	~15-yr apart	737	Rubicz (2013) [48]
AI/AN (Non-Hispanic)	472	2012-2016		Mekonian (2019) [70] <sup>h</sup>				
	401.4 (396.7-406.1)		99% popul.	Henley (2020) [44]				
Hispanic (any race)	350.0 (347.9- 352.2)	2000-2019	5.0x10 <sup>5</sup>	SEER 2000-2003+NPCR, ACS, Inc. SR	77.6 (75.8-79.4)	1988-1994	4 921	NHANES III, Bate (2010) [33]
	348.9 (348.0-349.7)	2012-2016	99% popul.	Henley (2020) [44]				
United Mexican States <sup>g</sup> (Mexico)	140.4	2020	1.9x10 <sup>5</sup>	GLOBOCAN, IARC database [WHO] (2021) [51]	89.2 83 (mean, %)	2006 1980-2016	20 000 262 studies	Conde-Glez (2013) [52] Zuhair (2019) [28]

822

823 <sup>a</sup> *Sources of data:* Surveillance, Epidemiology, and End Result SEER\*Stat Database (2001-2019) – United States Department of  
824 Health and Human Services, Centers for Disease Control and Prevention, and National Cancer Institute and National Program of  
825 Cancer Registries (NPCR). Based on the 2021 submission. The U.S. Population Data File information is used in meticulous  
826 calculation of integrated SEER and NPCR incidence of cancer by authors of the publications cited herein. Rates are annually averaged

827 per 10<sup>5</sup> population for all anatomic sites and age-standardized to the year 2000 U.S. standard (Census Bureau by the direct method)  
828 population (19 age groups - Census P25-1103). Puerto Rico data were excluded. Created by state cancerprofiles.cancer.gov (2022).

829  
830 <sup>b</sup>*N*, sample size, study reports or counts.

831  
832 <sup>c</sup> Significance was set at  $p < 0.05$  (associated with the Student *t* test comparing pairwise difference between CMV prevalence (%) and  
833 annual incident cancer rate). Confidence intervals (CI) were available only for some prevalence rates. For significance see **Tab. 2** and  
834 *Statistical Analysis* section.

835  
836 <sup>d</sup> Data based on NHIA. Individuals from Hawaii, Seattle, and Alaska Native Registries are excluded (SEER 2003).

837  
838 <sup>e</sup> We utilized CMV prevalence estimates for New Zealand ‘CMV safe’ Pacific Islander blood donors [78]. This was done because,  
839 being ancestrally and biologically linked and sharing similar heritage and migration history over the last 5,000 years (as supported by  
840 linguistics and anthropology [81]), CMV serology status of the American/Asian Pacific Islanders may be fairly comparable to that of  
841 New Zealand/Pacific Islander population. CMV seropositivity of New Zealand Asian blood donors was 77.6% (95% CI: 79.4-81.8).

842  
843 <sup>f</sup> Persons of Hispanic ethnicity can be of any race or subpopulation race combinations. ‘Hispanic’ and ‘Latina/o’ are used  
844 interchangeably to refer to Latin Americans and/or Spanish-speakers. Analysis of Hispanics should be disaggregated across Hispanic  
845 ethnic groups to better understand CMV burden among diverse Hispanic populaces.

846  
847 <sup>g</sup> United Mexican States (Mexico) data was introduced as a convenient contrast, a single point excluded from calculations. The most  
848 sizeable portion of Hispanics in U.S. is of Mexican descent. The cancer incidence rate per 10<sup>5</sup> persons and age-adjusted to the IARC  
849 world standard population is ~140 (95%CI not reported). Central and South American (Latin American) populations also have the  
850 lowest overall age-adjusted cancer mortality rates per 10<sup>5</sup> population (man: 91.19 [95% CI: 89.09-93.29]; women: 69.72 [95% CI:



851 68.42-71.01]) [69]. These populations also have the highest prevalence of CMV seropositivity (80-95%) [33] which underscores  
852 inverse association, shielding from cancer by CMV seropositivity.

853

854 <sup>h</sup> Additional references used in this work, including mostly API and AI/AN: [72-76,64-66], are included for readers who might wish to  
855 undertake more precise calculations.

856

857 *Notes:* Gender by race/ethnicity was modelled by adjusting for age, birthplace, household income level, insurance, household  
858 education, and crowding index by Bate *et al.* (2010) [33]. Cancer incidence rates and seroprevalence of CMV data must be viewed  
859 cautiously, because of broad geographic cover of the study populations and racial misclassification which differed somewhat over  
860 time. Hence quite an amount of biases exists in numerical figures standing for the crude rate estimates chosen. Other weaknesses of  
861 the estimates of racial/ethnic disease in data herein used have been previously recognized by Miller *et al.* [46] and other researches  
862 (see refs. 42-47 cited in [46]). We relied on their work in the current exploration.

863

864

865

<b>Table 2.</b> Age-standardized, delay-adjusted incidence rates <sup>a</sup> for common cancers <sup>b</sup> (all anatomical sites combined, both genders, all ages) among the racial/ethnic groups <sup>c</sup> (2012-2016) <sup>d</sup> in the United States compared to population CMV seropositivities (%).			
<b>Cancer incidence rate (95% CI) by race/ethnicity</b>		<b>Prevalence of CMV seropositivity by race/ethnicity</b>	
		(%)	<i>N</i>
All races	447.9 (447.6,448.2)	59	12 778
<i>Demographic stratification</i> (in descending rank order)			

866

Non-Hispanic White	467.5 (467.1, 467.8)	46	8 212
Non-Hispanic Black	466.9 (465.8, 467.9)	76	6 228
Non-Hispanic AI/AN (PRCDA)	401.4 (396.7, 406.1)	92	1 410
Hispanic	348.9 (348.0, 349.7)	78	4 921
Non-Hispanic API	293.6 (292.5, 294.8)	93	342
United States of Mexico <sup>e</sup> (Mexico)	140.4 (NR)	89	20 000

867

868

869 *Abbreviations:* AI/AN, American Indian/Alaska Native; API, Asian/Pacific Islander; CI,  
870 confidence interval; PRCDA, Indian Health Services Purchased/Referred Care Delivery Area *n*,  
871 sample size (unweighted), NR, not reported.

872 *Sources of data:* Henley *et al.* (2020) [44], Cronin *et al.* 2022 [45], GLOBOCAN, 2020  
873 (Estimated incidence rate of cancer for Mexico came from national mortality estimates by  
874 modelling, using mortality/incidence ratios derived from cancer registry data in neighboring  
875 countries) but see **Table 1**; Cannon *et al.* (2010) [2], Staras *et al.* (2006), Bate *et al.* (2010) [33],  
876 and Conde-Glez *et al.* [52] but please see **Table 1**.

877 <sup>a</sup>Rates are per 100 000 individuals, adjusted for potential delays in reporting and age-adjusted to  
878 the 2000 U.S. standard population (19 age groups; U.S. Bureau of the Census. Current  
879 Population Reports, Publication 25-1130. U.S. Government Printing Office, 2000 [Census P25-  
880 1130]).

881 <sup>b</sup>Cancer incidence rates are age-adjusted for 2012 to 2016 for all racial/ethnic groups combined.  
882 More than 15 cancers may appear for men and women to include the top cancers in each  
883 racial/ethnic group.

884 <sup>c</sup>Racial/ethnic groups are mutually exclusive. Data for non-Hispanic AI/AN are restricted to  
885 PRCDA counties (including Minnesota).

886 <sup>d</sup>Areas in the United States with high-quality incidence data are included and listed in Table 1 in  
887 Henley *et al.* (2020) [44].

888 °Mexico, provided for the contrast only, was excluded from statistical evaluation of association  
889 between incidence of cancer and CMV seroprevalence.

890 *Note:* there is a highly significant inverse association between cancer rates and the corresponding  
891 prevalence of CMV seropositivity among the race/ethnic groups in the U.S. ( $p < 0.001$ ; see,  
892 *Statistical Analysis*). Annualized cancer incidence rates (men only, all sites, all ages) among  
893 AIs/ANs vary widely (almost 4-fold in the 1980-ties). They were dramatically lower in the  
894 Southwest, New Mexico Indians (Navajo, Apache, and Pueblo), 276.6;  $N=1,103$ , and Alaska  
895 Indians, 668.1;  $N=547$  [73], and also much lower than in Non-Hispanic Whites of New Mexico  
896 [85].

897  
898  
899  
900  
901  
902  
903  
904

Tumor/localization	Spearman's $\rho$	$p$ -value	Inverse correlation (+/--)
Melanoma (skin)	-0.763	0.001	+
Kidney	-0.754	0.001	+
All cancers	-0.732	0.001	+
All cancers (excluding non-melanoma skin cancer)	-0.726	0.001	+
Breast	-0.719	0.001	+

Testis	-0.711	0.001	+
Non-melanoma (skin)	-0.692	0.001	+
Colorectum	-0.671	0.001	+
Vulva	-0.665	0.001	+
Prostate	-0.663	0.001	+
Corpus uteri	-0.656	0.001	+
Oropharynx	-0.651	0.001	+
Pancreas	-0.633	0.001	+
Multiple myeloma	-0.633	0.001	+
Leukemia	-0.632	0.001	+
Hodgkin lymphoma	-0.618	0.001	+
Non-Hodgkin lymphoma	-0.617	0.001	+
Mesothelioma	-0.574	0.001	+
Lip, oral cavity	-0.551	0.001	+
Lung	-0.548	0.001	+
Brain/CNS	-0.541	0.001	+
Thyroid	-0.532	0.001	+
Bladder	-0.519	0.001	+
Ovary	-0.461	0.001	+
Penis	-0.432	0.001	+
Hypopharynx	-0.377	0.001	+
Salivary glands	-0.35	0.002	+
Gallbladder	0.316	0.006	--
Nasopharynx	0.266	0.023	--
Vagina	-0.224	0.056	+
Larynx	-0.165	0.164	+
Esophagus	-0.149	0.208	+
Cervix uteri	0.118	0.319	--
Stomach	-0.085	0.473	+
Kaposi's sarcoma	-0.007	0.953	+

Liver	0.007	0.951	--
-------	-------	-------	----

905

906 *Notes:* Significant inverse correlation, implying a possible prevention of tumorigenesis due to  
907 primary CMV infection the world over, is marked by +; the apparent lack of this association (i.e.  
908 absence of CMV-mediated anti-cancer protection) is denoted by --.

909 Most recent clinical studies confirm negative (i.e., protective) correlation between CMV  
910 infection and both colorectal [86] and bronchogenic cancer [87].

911 For global correlations, between an aggregated cancer incidence per distinct country and  
912 the country's specific CMV seropositivity rate, please see **Figure 1**.

913

914