
**Symposium on Natural and Applied Sciences**
Hosted Online from London, United Kingdom
Date: 5th March, 2026
Website: https://econferencia.com

# POLYGENIC RISK STRATIFICATION OF CYTOMEGALOVIRUS INFECTION USING A FIVE-LOCUS SCORING MODEL: DIAGNOSTIC ACCURACY AND CLINICAL APPLICABILITY

Djuraev M. G.
Azimova S. B.
Vasilevskiy E. A.
Dadabaev O. T.
Andijan State Medical Institute, Andijan, Uzbekistan

## Background

Cytomegalovirus infection (CMVI) remains one of the most prevalent infectious diseases worldwide, with seroprevalence reaching 85–95% in developing regions including Central Asia (1, 8). The clinical spectrum ranges from asymptomatic latency to severe multi-organ involvement in neonates, pregnant women, and immunocompromised patients (2, 6). Predicting which patients will progress from latent to manifest persistent infection remains a key clinical challenge, as traditional laboratory markers such as viral load and C-reactive protein reflect the current disease state rather than the inherent susceptibility of the host (3, 6). Single nucleotide polymorphisms (SNPs) in genes regulating immune response (TNF-α, IL-10, TLR4), oxidative stress defense (SOD2), and endothelial function (eNOS) have been individually associated with susceptibility to CMVI; however, no single SNP has demonstrated sufficient discriminative power for clinical application, with individual AUC values typically below 0.70 (4, 5, 7). The polygenic nature of susceptibility to chronic viral infections suggests that a composite approach integrating multiple genetic variants into a single quantitative measure may yield superior predictive accuracy. Such cumulative genetic risk scores have shown clinical utility in cardiovascular medicine and

oncology (3, 6), but have not been applied to CMVI in Central Asian populations. The present study aimed to develop and evaluate a five-locus cumulative genetic risk score for predicting the manifest persistent form of CMVI and to assess the added value of combining genetic data with viral load measurements.

**Objective:**

To develop a cumulative genetic risk score based on five SNPs (TNF-α, IL-10, TLR4, eNOS, SOD2) and evaluate its diagnostic performance for predicting manifest persistent CMVI, alone and in combination with viral load data.

**Materials and methods**

The study was conducted at medical institutions of the Fergana Valley (Zam-Zam Clinic, Andijan; Andijan, Namangan, Fergana Regional Infectious Disease Hospital's) during 2022–2025. Genotyping for TNF-α (G308A, rs1800629), IL-10 (G>A, rs1800896), TLR4 (A896G, rs4986790), eNOS (G894T, rs1799983), and SOD2 (Ala16Val, rs4880) was performed in 100 CMVI patients (Group I, n=42, latent form; Group II, n=58, manifest persistent form) and 80 healthy controls using real-time PCR with TaqMan probes (CFX96, Bio-Rad). Genomic DNA was extracted from venous blood (QIAamp DNA Blood Mini Kit, Qiagen). Hardy-Weinberg equilibrium was verified in the control group. The cumulative genetic risk score (0–10 points) was calculated by assigning 0 points for each protective homozygous genotype, 1 for heterozygous, and 2 for risk homozygous per locus. Risk categories: low (0–2), intermediate (3–5), high (6–8), very high (9–10). The combined model integrated the genetic score (≥6) with viral load (>$10^3$ copies/mL). Statistical analysis was performed in SPSS 26.0 and R 4.2 with Pearson $\chi^2$, Fisher exact test, OR and RR with 95% CI. ROC analysis was used

to evaluate diagnostic performance (AUC, sensitivity, specificity, PPV, NPV). Significance was set at $p<0.05$.

**Results and Discussion**

HWE was confirmed for all five loci in the control group ($p>0.05$ for all). Individual polymorphism analysis demonstrated significant associations with the manifest persistent form for all five SNPs, with risk allele frequencies consistently higher in Group II. However, no single polymorphism achieved an AUC above 0.70: IL-10 = 0.68, SOD2 = 0.66, TNF-α = 0.64, eNOS = 0.63, TLR4 = 0.61. This confirmed the necessity of a composite approach.

The distribution of cumulative genetic risk scores differed profoundly between groups. A high genetic score (≥6) was found in 65.5% of patients with the manifest form and only 16.7% with the latent form ($\chi^2=24.12$; $p<0.001$; OR=9.64; 95% CI 3.58–25.9; RR=3.93; 95% CI 1.91–8.08). In the low-risk category (0–2 points), 38.1% belonged to Group I versus only 6.9% to Group II (OR=0.12; $p<0.001$), confirming the protective effect of a favorable genetic profile. The very high category (9–10) was almost exclusively represented in Group II (22.4% vs 2.4%; OR=11.9; $p=0.004$), indicating that maximal genetic burden carries the most pronounced clinical consequences.

At the threshold of ≥6, the genetic score demonstrated a sensitivity of 65.5% (95% CI 51.9–77.5), specificity of 83.3% (95% CI 68.6–93.0), PPV of 79.2%, NPV of 71.4%, and AUC of 0.78 (95% CI 0.69–0.87; $p<0.001$). This represents a substantial improvement over any individual SNP and exceeds the threshold of 0.75 generally considered acceptable for clinical screening tools (3, 6). The rationale for the composite approach is grounded in the non-overlapping biological functions of the five SNPs: TLR4 mediates innate immune recognition, TNF-α and IL-10 regulate the pro/anti-inflammatory balance, SOD2 controls

mitochondrial antioxidant defense, and eNOS maintains NO-dependent antiviral activity (4, 5, 7). By capturing multiple levels of host defense simultaneously, the score overcomes the inherent limitation of single-marker approaches.

Combining the genetic score with a viral load criterion (>$10^3$ copies/mL) further improved diagnostic performance: AUC increased to 0.85 (95% CI 0.77–0.93; $p<0.001$), sensitivity was 67.2%, specificity reached 90.5%, PPV rose to 90.7%, and NPV was 66.7%. The superiority of the combined model reflects the complementary nature of its components: the genetic score captures the stable, lifelong predisposition of the host, while viral load reflects the current state of viral replication (3, 6). The moderate correlation between the two parameters ($r=0.44$; $p<0.001$) confirms that they capture partially overlapping but largely independent aspects of infection pathogenesis, which is why their integration yields additive diagnostic value.

A key advantage of the genetic score is its stability. Unlike viral load, CRP, or transaminase levels, the genetic profile is determined once from a standard blood sample and remains valid throughout the patient's lifetime. This property makes it uniquely suited for long-term prognostication and for identifying patients requiring enhanced surveillance even during apparent clinical remission. Goodrum F. et al. (2021) highlighted the unpredictable nature of CMV reactivation and the need for host-based markers independent of the current disease state (3), which aligns precisely with the characteristics of our score. The finding that 81.8% of children with congenital CMVI in Group II had scores of ≥6 raises the possibility that maternal genetic profiling could inform vertical transmission risk assessment (2, 5).

The population-specific relevance of these findings should be emphasized. The Fergana Valley, with high population density, extensive household contact patterns, and near-universal CMV seroprevalence (1, 8), represents a setting

where the clinical outcome is determined not by whether infection occurs, but by how the host responds. Genetic profiling therefore becomes a particularly powerful prognostic tool in this context. Several limitations should be acknowledged: the sample size limits subgroup analyses; the scoring assigns equal weight to all loci, whereas weighted scoring or machine learning approaches may improve accuracy; and multicenter validation across Central Asia is needed before widespread implementation. Expansion of the panel to include interferon response genes (IFN-γ, IFNL3) and KIR receptors may further enhance predictive power.

**Conclusions**

1. The five-locus cumulative genetic risk score enables effective stratification of CMVI patients: a score of ≥6 is associated with a 9.64-fold increased odds of manifest persistent infection ($p<0.001$), with AUC of 0.78 substantially exceeding any individual SNP (AUC range 0.61–0.68).
2. Combining the genetic score with viral load (>$10^3$ copies/mL) increases AUC to 0.85, specificity to 90.5%, and PPV to 90.7%, providing a robust tool for identifying patients requiring intensive monitoring and early antiviral intervention.
3. The genetic score is a stable, one-time measurement independent of the current clinical state, making it uniquely suited for lifelong risk assessment and personalized management of CMVI in populations with high seroprevalence.

**References**

1. Cannon M.J., Schmid D.S., Hyde T.B. Review of cytomegalovirus seroprevalence // Rev. Med. Virol. – 2010. – Vol.20. – P.202–213.
2. Boppana S.B., Ross S.A., Fowler K.B. Congenital cytomegalovirus infection: clinical outcome // Clin. Infect. Dis. – 2021. – Vol.57(Suppl 4). – P.S178–S181.




3. Goodrum F., Caviness K., Zagallo P. Human cytomegalovirus persistence // Cell. Microbiol. – 2021. – Vol.14(5). – P.644–655.
4. Jabłońska A., Kocięcka B., Adrych K. TLR4 polymorphisms and viral infections // Immunobiology. – 2022. – Vol.227(1). – Art.152156.
5. Lanzieri T.M., Dollard S.C. et al. Systematic review of congenital CMV // J. Clin. Virol. – 2017. – Vol.68. – P.6–14.
6. Griffiths P., Reeves M. Pathogenesis of human cytomegalovirus // Nat. Rev. Microbiol. – 2021. – Vol.19. – P.759–773.
7. Wilson A.G., Symons J.A. et al. Polymorphism in human TNF-alpha promoter // Proc. Natl. Acad. Sci. USA. – 2018. – Vol.94. – P.3195–3199.
8. Zuhair M., Smit G.S.A., Wallis G. Systematic review of global CMV seroprevalence // BMC Public Health. – 2022. – Vol.22. – P.1659.