***Author: This file is the accepted version of your manuscript, and it shows any changes made by the Editor-in-Chief and the Deputy Editor since you submitted your last revision. This is the version that is being sent to Manuscript Editing for further editing in accordance with NEJM style. You will receive proofs of the edited manuscript, by e-mail. The proofs will contain queries from the manuscript editor, as well as any queries that may be present in this file. The proof stage will be your next opportunity to make changes; in the meantime, please do not make any changes or send any new material to us.***

**Genetic Diversity and Protective Efficacy of the RTS,S/AS01 Malaria Vaccine**

Daniel E. Neafseya,1, Ph.D., Michal Juraskaa,2, Ph.D., Trevor Bedfordb,2, Ph.D., David Benkeserb,3, M.P.H., Clarissa Valimb,1,4, M.D., Sc.D., Allison Griggs1, M.Sc, Marc Lievens5, M.Sc., Salim Abdulla6, M.D., Ph.D., Samuel Adjei7, M.B., Ch.B., D.T.M., Tsiri Agbenyega7, M.B., Ch.B., Ph.D., Selidji T. Agnandji8,9, M.D., Pedro Aide10,11, M.D., Ph.D., Scott Anderson1, B.S., Daniel Ansong8, M.B., Ch.B., F.W.A.C.P., John J. Aponte10,11, M.D., Ph.D., Kwaku Poku Asante12, M.D., Ph.D., Philip Bejon13, Ph.D., Ashley J. Birkett14, Ph.D., Myriam Bruls5, M.Sc., Kristen M. Connolly1, B.S., Umberto D’Alessandro15, M.D., Ph.D., Carlota Dobaño10,11, Ph.D., Samwel Gesase16, M.D., Brian Greenwood12,17, M.D., Jonna Grimsby1, Ph.D., Tinto Halidou18, Ph.D., Mary J. Hamel19, M.D., D.T.M.&H., Irving Hoffman20, P.A., M.P.H, Portia Kamthunzi21, M.D., Simon Kariuki22, Ph.D., Peter G. Kremsner9,8, M.D., Amanda Leach5, M.R.C.P.C.H., Bertrand Lell9,8, M.D., Niall J. Lennon1, Ph.D., John Lusingu16, M.D., Ph.D., Kevin Marsh13, M.D., Francis Martinson21, M.D., Ph.D, Jackson T. Molel6, M.Sc., Eli L. Moss1, B.S., Patricia Njuguna13, M.B., Ch.B., M.Med, Christian F. Ockenhouse14, M.D., Ph.D., Bernhards Ogutu Ragama23, Ph.D, Walter Otieno23, M.B., Ch.B., M.Med., Lucas Otieno23, M.B., Ch.B., M.P.H., Kephas Otieno22, M.P.H., Seth Owusu-Agyei12, Ph.D, Daniel J. Park1, Ph.D., Karell Pellé1,4, Ph.D., Dana Robbins1, B.S., Carsten Russ1, Ph.D., Elizabeth M. Ryan1, B.S., Jahit Sacarlal10, M.D., Ph.D., Brian Sogoloff1, B.S., Hermann Sorgho18, Ph.D., Marcel Tanner24, Ph.D., Thor Theander25, M.D., D.Sc., Innocent Valea18, Pharm. D., Ph.D., Sarah K. Volkman1,4,26, Sc.D., Qing Yu1, M.Sc., Didier Lapierre5, M.D., Bruce W. Birren1, Ph.D., Peter B. Gilbert\*,2,3, Ph.D., Dyann F. Wirth\*1,4, Ph.D.

a These authors contributed equally to this manuscript.

b These authors contributed equally to this manuscript.

1 Broad Institute of MIT and Harvard, Cambridge, MA

2 Fred Hutchinson Cancer Research Center, Seattle, MA

3 University of Washington, Seattle, WA

4 Harvard T.H. Chan School of Public Health, Boston, MA

5 GlaxoSmithKline Vaccines, Rixensart, Belgium

6 Ifakara Health Institute, Bagamoyo, Tanzania

7 School of Medical Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

8 Albert Schweitzer Hospital, Lambaréné, Gabon

9 Institute of Tropical Medicine, University of Tübingen, Germany

10 Centro de Investigaçåo em Saúde de Manhiça, Manhiça, Mozambique

11 ISGlobal, Barcelona Ctr. Int. Health Research (CRESIB), Hospital Clinic – Universitat de Barcelona, Barcelona, Spain

12 Kintampo Health Research Centre, Kintampo, Ghana

13 KEMRI-Wellcome Trust Research Program, Kilifi, Kenya

14 PATH Malaria Vaccine Initiative, Washington DC

15 Medical Research Council Unit, Banjul, The Gambia

16 National Institute for Medical Research, Korogwe, Tanzania

17 London School of Hygiene and Tropical Medicine, London, UK

18 Institut de Recherche en Sciences de la Santé, Nanoro, Burkina Faso

19 Centers for Disease Control and Prevention, Atlanta, GA

20 University of North Carolina, Chapel Hill, NC

21 UNC Project, Lilongwe, Malawi

22 KEMRI/CDC Research and Public Health Collaboration, Kisumu, Kenya

23 KEMRI-Walter Reed Project, Kombewa, Kenya

24 Swiss Tropical and Public Health Institute, Basel, Switzerland

25 University of Copenhagen, Copenhagen, Denmark

26 Simmons College School of Nursing and Health Sciences, Boston, MA

\* To whom reprint requests should be addressed:

pgilbert@fredhutch.org

dfwirth@hsph.harvard.edu

**Corresponding author:**

Dyann F. Wirth, Ph.D.

Department of Immunology and Infectious Diseases

Harvard T.H. Chan School of Public Health

665 Huntington Avenue

Boston, MA 02115

T: 617-432-1563

dfwirth@hsph.harvard.edu

**ABSTRACT**

**Background**:The RTS,S/AS01 vaccine targets the circumsporozoite (CS) protein of *Plasmodium falciparum* and confers partial protective efficacy against clinical and severe malaria disease in infants and children (NCT00866619). We investigated whether vaccine efficacy (VE) was specific to parasite genotypes at CS.

**Methods**: We employed PCR-based next-generation sequencing of DNA extracted from 4985 participant samples to survey polymorphisms in CS. We evaluated the impact of polymorphic CS positions and haplotypic regions on VE against first episodes of clinical malaria within a year of vaccination.

**Results**: In the 5‑17 month old per-protocol category of 4557 RTS,S/AS01 vaccinated and 2328 control vaccinated participants, one-year cumulative VE against clinical malaria matching the vaccine in the entire CS C‑terminus was 50.3% (95% confidence interval [CI], 34.6 to 62.3; 139 infections), compared to 33.4% (95% CI, 29.3 to 37.2; 1951 infections) against mismatched malaria (P = 0.04 for differential VE). Hazard ratio VE was 62.7% (95% CI, 51.6 to 71.3) vs. 54.2% (95% CI, 49.9 to 58.1), P = 0.06. In the 6-12 week old category there was no evidence of differential VE.

**Conclusions**: These results suggest that the VE of the RTS,S vaccine in 5-17 month old children has increased activity against the matched allele. The overall VE will depend on the proportion of matched alleles; in this trial, less than 10% of parasites had matched alleles. (ClinicalTrials.gov number, NCT00866619)

Malaria induces substantial worldwide morbidity and mortality,1 and has proven a challenge for vaccine development efforts. The recently renewed effort to control, eliminate, and hopefully eradicate malaria will enjoy greater likelihood of success if a vaccine can be combined with other intervention methods such as drug administration campaigns and insect vector control.2,3 The most advanced candidate vaccine for protection against *Plasmodium falciparum* malaria infection, RTS,S/AS01, is a monovalent recombinant protein vaccine targeting a fragment of the circumsporozoite (CS) parasite antigen. RTS,S/AS01 underwent a large randomized, controlled phase 3 trial at 11 study sites in Africa between 2009 and 2013 to evaluate efficacy, safety, and immunogenicity in over 15000 children. The vaccine confers moderate protective efficacy against clinical disease and severe malaria that wanes over time,4–7 concordant with findings from multiple phase 2 trials.8–13 Higher protection was observed in young children (5-17 months old at first vaccination) relative to infants (6-12 weeks old at first vaccination).4–7

The mechanism by which the vaccine confers protection is incompletely understood, and differing hypotheses exist regarding the relative importance of B cell versus T cell mediated immunity to CS and other antigens.14–17 The CS protein is expressed on the surface of the infective sporozoite stage of the parasite and contains a conserved ‘NANP/NVDP’ tandem repeat exhibiting length polymorphism ranging from 37 to 44 repeat units18, which is thought to represent the dominant B cell epitope.19 The C-terminus region of CS exhibits abundant polymorphism,18,20–22 some of which resides within described T cell epitopes (Th2R, Th3R),23 but which may also function as B cell epitopes24. While the immune response provoked by RTS,S/AS01 may be distinct from the natural response to CS,25,26 the partial protective efficacy of the vaccine may be due in part to allelic specificity. Evidence for naturally acquired allele‑specific immune protection has been observed for the MSP-1 antigen27 (but not CS28) in a prospective cohort study, and allele‑specific protection has been reported in a field trial for a vaccine based on the AMA1 antigen.29

Previous genetic analyses of parasite samples from three different RTS,S phase 2 studies failed to detect an association of protective efficacy with genetic similarity to 3D7 (the vaccine construct parasite line).30–32 This study has both larger sample size and improved sequencing technology. Next generation (Illumina MiSeq, PacBio) sequencing of polymerase chain reaction (PCR) amplicons from malaria infected participant samples offers more sensitive genetic investigations of allele‑specific protection and allows more immunologically relevant analyses of multi-variant haplotypes. We used a sieve analysis methodology (**Figure S1**) previously applied to detect allele-specific HIV-1 vaccine efficacy33,34 to analyze the CS data for differential vaccine efficacy at three levels: the entire C-terminus amplicon haplotype (95 amino acids); defined haplotypic regions of the C-terminus (10-17 amino acids); and, individual polymorphic positions for 2 defined trial endpoints in 2 age categories. We also investigated the relationship between vaccine efficacy and NANP/NVDP repeat count. In addition, as a control we investigated the SERA2 locus, which was not in the vaccine and hence no differential efficacy was expected.

**METHODS**

**STUDY DESIGN**

The RTS,S phase 3 trial design has been described previously.4–7 Participant samples representing two defined endpoints in the per-protocol (received all three immunizations at Months 0, 1, 2) cohort were sequenced: 1) primary clinical malaria (first or only episode of clinical malaria with > 5000 parasites per cubic ml occurring between 14 and 385 days following the third immunization) and 2) parasite positive (symptomatic or asymptomatic) samples collected from all participants at 18 months post-vaccination with > 0 parasites/ml3. **Figure 1** and **Figure S2** summarize sample size and follow-up information for the per‑protocol category of 5-17 month olds, where genetic data were measured from 1181 RTS,S/AS01 vaccine recipients and 909 rabies control vaccine recipients who experienced the primary clinical malaria endpoint, and from 284 RTS,S/AS01 vaccine recipients and 208 control vaccine recipients who experienced the parasite positive endpoint. **Figure S3 and Figure S4** show this information for the per-protocol category of 6-12 week olds, and **Figures S5-S8** report on samples studied for both endpoints and age categories for the NANP/NVDP repeat amplicon**.** Samples were analyzed from all 11 trial sites, spanning 6 countries in Africa (**Figure 2A**).

**STUDY OVERSIGHT**

The trial (NCT00866619) was sponsored by GlaxoSmithKline Biologicals (GSK), the vaccine developer and manufacturer, and funded by both GSK and the PATH Malaria Vaccine Initiative, which received a grant from the Bill and Melinda Gates Foundation. The trial protocol was approved by all relevant ethical review boards and signed or thumbprint informed consent was obtained from the children’s parents or guardians.4–7

**SEQUENCE DATA GENERATION**

Participant samples were received as dried blood spots on Whatman FTA sample cards. Methods for DNA extraction, PCR amplification, and sequencing are described in the Supplementary Appendix. The CS C-terminus and SERA2 amplicons were sequenced on an Illumina MiSeq platform. NANP/NVDP amplicon was sequenced on a PacBio platform due to its greater length. **Figure 2B** illustrates the respective locations of the NANP/NVDP repeat region and CS C-terminus amplicon within the CS protein. The analyses included MiSeq data for 4421 and 4499 CS C-terminus and SERA2 samples, respectively, and PacBio data for 3137 CS NANP/NVDP samples [**Tables S1-S4** and SAP (Supplementary Appendix)]. All MiSeq and PacBio amplicon sequence data were submitted to the NCBI Sequence Read Archive35 (BioProject PRJNA235895).

**STATISTICAL ANALYSIS**

The analyses pre-specified for this study are described in the SAP (Supplementary Appendix). In brief, we performed sieve analyses on translated amino acid sequences to assess differential VE against first or only episodes of clinical malaria over 12 months follow up post dose 3 and parasite positivity (detected by blood slide) at 18 months post dose 3 defined by perfect vaccine haplotype match vs. mismatch (CS C-terminus) or by the number of NANP/NVDP repeats. We refer to differential VE by some parasite feature as a “sieve effect.” For the primary analysis of clinical malaria, two haplotype-specific VE parameters were assessed: cumulative VE that measures one minus the ratio (RTS,S/AS01 vs. control) of cumulative incidences of the first or only episode of clinical malaria with a specific haplotype by a given number of days *t* beyond 14 days after the third vaccination; and hazard ratio VE that measures one minus the ratio (RTS,S/AS01 vs. control) of instantaneous incidences of the endpoint that assumes proportional incidences (RTS,S/AS01 vs. control) over time. Aalen-Johansen nonparametric maximum likelihood estimation (MLE) with stratification by study site was used to estimate cumulative VE against vaccine matched and mismatched malaria, with Wald tests for non-zero VE and for a sieve effect of differential VE. Targeted MLE36 was used for addressing the same objectives adjusting for all relevant baseline subject covariates (listed in **Table S5**). Hazard ratio VE was estimated using cause-specific Cox models stratified by study site using score tests for non-zero VE and Wald tests for sieve effects.37 The analysis methods were selected to be in close alignment with those used for assessing overall VE in the original published analyses.4–7 For the parasite positive endpoint, similar analysis methods as the cumulative VE analysis of the primary clinical malaria endpoint were performed.

Many samples exhibited complex infections, generated via multiple parasite founder genotypes. Consequently, sieve analyses were performed on datasets composed of one founder haplotype randomly selected from each subject, with multiple outputation38 employed to aggregate results (details in the SAP). We also performed sieve analyses on datasets in which participants with one or more founder haplotypes were classified as having “any match” to 3D7 or “no match” to 3D7.

In addition to investigating the previously described Th2R and Th3R epitopes,23 we analyzed haplotype frequencies in a previously undefined genomic region we designate as ‘DV10’ (representing 10 amino acid positions 293-302) bounded by amino acids aspartate (D) and valine (V), and an ‘LD’ haplotype based on the six positions (314, 317, 352, 354, 356, 357) we found to exhibit linkage disequilibrium (LD; defined for sites with minor allele frequency $\geq $ 3% as r2 $\geq $ 0.1 with $\geq $ 2 other positions at the five largest study sites; **Figure S9**).

All analyses were performed separately for each age category. Multiplicity adjustment of sieve effect P values across the epitope haplotypes and positions was applied separately to the 2 age categories, the 2 studied proteins (CS and SERA2), and the VE parameter type (cumulative vs. hazard ratio). Family-wise error rate (FWER; Holm-Bonferroni39) and false discovery rate (FDR or Q-values; Benjamini-Hochberg40) error adjustment was applied. Results with Q ≤ 0.20 for all multiply compared loci or the unadjusted P ≤ 0.05 for the full CS C-terminus amplicon were considered to be significant, with FWER P ≤ 0.05 indicating more stringent significance. All P-values and Q-values are 2-sided.

**RESULTS**

**COMPLEXITY OF INFECTION**

The majority of samples from participants with primary clinical malaria in both age categories were complex, defined as being founded by 2 or more distinct parasite lineages (infants: 68%; children: 65%). In the older age category, the distribution of complexity of infection (COI) was shifted toward fewer parasite lineages in RTS,S/AS01 vaccine recipients than in control vaccine recipients (**Figure 2C**; RTS,S/AS01: 61%; control: 71%; Wald test, P < 0.001), whereas, in the younger age category, there was no evidence for a different COI distribution between the study groups (**Figure S10**; RTS,S/AS01: 67%; control: 70%; P = 0.43). This observation for the older age category is concordant with findings in two phase 2 trials of the related RTS,S/AS02 vaccine,31,32 and there are fewer 3D7 matching haplotypes in high-COI infections in the RTS,S/AS01 vaccinated group relative to the control group (**Figure S11).**

**POPULATION VARIATION PROFILE**

We searched for a sieve effect based on perfect vaccine match/mismatch in the C‑terminus of CS at three scales: the full amplicon haplotype (95 amino acids), 4 described epitopes/polymorphism cluster haplotypes (10-17 amino acids apiece), and 25 individual polymorphic positions. Descriptively, the frequency of haplotypes with an exact match to the 3D7 vaccine strain across all polymorphic positions varied considerably among study sites (**Figure 2D**). In addition, there was a lower frequency of 3D7 haplotypes in RTS,S/AS01 vs. control vaccine recipients especially at geographic sites exhibiting at least 5% frequency of 3D7 in the control arm.Similar differences were evident for the epitope haplotype frequencies (**Figure 2E**). The frequency of alleles at individual polymorphic positions matching 3D7 was variable (**Figure 2F**). **Figure S10** shows similar population frequencies in the study groups for the 6-12 week old category.

**C-TERMINUS REGION SIEVE EFFECTS**

There were 139 clinical malaria cases with a perfect full-amplicon 3D7 match (**Figure 3A**)and 1951 mismatched cases (**Figure 3B**) detected through one year following vaccination. Cumulative VE through one year after vaccination against clinical malaria with a perfect full‑amplicon 3D7 match was 50.3% (95% CI, 34.6 to 62.3) and against mismatched clinical malaria was 33.4% (95% CI, 29.3 to 37.2), with VE significantly higher against matched malaria **(Figure 3D;** **Figure 4A;** Wald sieve effect P = 0.04). The covariate-adjusted analysis gave almost identical results (**Table S5**). Cumulative VE was higher against matched than mismatched malaria throughout the follow‑up period, for example through month 6, when VE against matched malaria was 70.2% (95% CI, 56.1 to 79.8) and against mismatched malaria was 56.3% (95% CI, 51.1 to 60.9; **Figure 3C**) (sieve effect P = 0.05). Cumulative VE and sieve effects for the CS C-terminus also varied in magnitude among study sites when they were analyzed individually in the older age category (**Table S6**).Hazard ratio VE aggregating over the 12 months of post-vaccination follow-up was also higher against matched (62.7%, 95% CI, 51.6 to 71.3) than mismatched (54.2%, 95% CI 49.9 to 58.1) malaria (**Figure 4B**; **Table S7**) (sieve effect P = 0.06). Overall VE was similar to that observed against mismatched malaria, because greater than 90% of the infections are mismatched (**Figure 3D**). Post-hoc analysis defining a malaria case’s haplotypes as ‘any match’ or ‘no match’ to 3D7 also identified a sieve effect for the CS C‑terminus (cumulative VE sieve effect P <0.001; hazard ratio VE sieve effect P = 0.002; **Tables S8-S9**).In contrast, cumulative and hazard ratio VE was similar against full amplicon vaccine matched vs. mismatched malaria in the 6-12 week old category (sieve effect P = 0.58; **Tables S10-S11; Figure S12**), and for both age categories at the SERA2 locus (sieve effect P-values > 0.30, **Tables S12-S15; Figures S13-S14**).

In the 5-17 month old category there were also significant cumulative VE sieve effects (Q‑value ≤ 0.2) for the Th2R and Th3R epitopes, the Th2R/Th3R ‘LD’ haplotype, and the DV10 region (**Figure 4A**). For individual amino acid positions, there were significant cumulative VE sieve effects (Q-value ≤ 0.2) at positions 299, 301, 317, 354, 356, 359, and 361 (**Table 1**). Cumulative VE tended to decrease with the number of 3D7 mismatches at these seven positions (**Figure S15**).Hazard ratio analyses of epitopes/regions (**Figure 4B)** and individual amino acid positions (**Table S7**) yielded differential VE results consistent with those from the cumulative VE analysis, at reduced levels of significance. Vaccine efficacy was similar against matched vs. mismatched CS C‑terminus individual positions for the younger age category (all Q-values > 0.20, **Tables S10-S11**). No evidence of sieve effects was found for individual positions in either age category for the SERA2 locus (**Table S12-S15**).

For the parasite positive endpoint at 18 months post dose 3, in the older age category the VE estimates tended to be higher for CS C-terminus vaccine matched vs. mismatched malaria (e.g., VE = 53% vs. 30%, P = 0.19 for the full-amplicon), although none of the differences were statistically significant (**Table S16**). In contrast, there was no evidence at all of a sieve effect in the younger age category at the CS C-terminus for this endpoint (**Table S17**), nor at SERA2 in either age category (**Tables S18-S19**).

**NANP/NVDP REPEAT REGION**

In 3137 participant samples representing the clinical malaria endpoint with evaluable sequence data from the B cell epitope repeat region, NANP/NVDP repeat count was observed to range from 37 to 44, with a mode of 40 repeats. There was a non-significant trend of declining cumulative VE with NANP/NVDP repeat count in the older age category (P = 0.072, **Figure S16**) and no significant differential VE by repeat count in the younger age category (P = 0.89, **Figure S17**). We did not assess how VE depends on NANP/NVDP repeat amino acid sequences because the vaccine construct contains a truncated repeat region (18.5 NANP/NVDP repeats).

**DISCUSSION**

The discovery of higher RTS,S/AS01 vaccine efficacy against clinical malaria with infections matching vs. mismatching the 3D7 vaccine construct at epitope haplotypes and amino acid positions is not entirely unexpected, given the polymorphism at the C‑terminus of the CS antigenic locus and previous observations of allele-specific immune responses to other parasite proteins.27,29 This differential cumulative VE result could be a false positive, given the full CS C-terminus result was borderline significant (P = 0.04) and the analysis of 4 epitopes and 24 amino acid sites gave no Holm-Bonferroni-adjusted P-values below 0.05. However, 11 of the 28 tests yielded Q-values below or equal to 0.20, such that we expect at least 80% of these 11 results to be true positives. The present study had greater power to detect allele-specific protection than previous RTS,S studies based on phase 2 trials due to three factors: 1) larger sample size, 2) inclusion of study sites harboring a higher frequency of 3D7-matching haplotypes, and 3) use of PCR-based next generation sequencing to resolve haplotypes comprising mixed infections. Our main result of significant sieve effects for the primary clinical malaria endpoint in the 5-17 month old category was based on a large sample size, with n=2145 total clinical malaria cases with measured genetic data. In contrast, the sieve analysis of the parasite positive endpoint in 5-17 month olds, which yielded non-significant results, had four times fewer cases (n=507 total cases with genetic data). However, in terms of estimates of vaccine efficacy the sieve effects were slightly stronger for the parasite positive endpoint, suggesting that the lack of statistical significance could be due to lower statistical power than to a true lack of differential protection. Post-hoc power calculations showed only 30% power to detect a difference between VE of 53% vs. 30% for CS-C-terminus match vs. mismatch malaria, compared to 51% power for the clinical malaria endpoint. Therefore the selective vaccine protection may have operated for both clinical malaria and parasite positivity. In contrast, there is no evidence for allele specific vaccine efficacy in the 6-12 week old category, despite significant overall protection in this group. This result implies a qualitative difference in the vaccine response, in addition to the previously reported quantitative difference in anti‑CS antibody titers in the younger age category.4 The biological mechanisms remain to be elucidated but could include a role of maternal antibody, interactions with other vaccine responses or differences in immune response capacity between the infant and child age categories. These results create motivation for further immunological analysis to discern the mechanisms of selective vs. non-selective vaccine-induced immunity.

Genetic surveillance of CS sequences in parasite populations could inform future vaccine candidates targeting polymorphic malaria parasite proteins. The genotype-specific VE results we report here complement previous estimates of RTS,S/AS01 efficacy in this phase 3 trial; the previously reported 12 month hazard ratio VE of 55.8% against clinical malaria in 5-17 month olds4 can now be interpreted as a multiplicative weighted average of hazard ratio VEs of 62.7% and 54.2% against matched (n = 139) and mismatched (n=1951) parasites, respectively (**Figure 4B**). The observed variation among study sites in infections with a perfect vaccine match in the CS C‑terminus (**Figure 2D, Figures S18-S19)** may help to explain previously reported variation in overall VE among study sites, though the magnitude of this contribution is expected to be low due to the overall rarity of the 3D7 haplotype.6 Broader deployment of the vaccine could potentially result in increased selection on the 3D7 haplotype or its component epitopes and amino acid alleles. Sieve analysis of next generation sequencing data constitutes an approach for understanding partial vaccine efficacy.

**Disclosure:**

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

**Acknowledgements:**

Supported by the Bill and Melinda Gates Foundation, the PATH Malaria Vaccine Initiative, and with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No.: HHSN272200900018C. This manuscript was reviewed by Joe Cohen, Ripley Ballou, and Lode Schuerman at GlaxoSmithKline plc prior to submission. We acknowledge the support and assistance of the RTS,S Clinical Trials Partnership Committee. We wish to thank the Broad Institute Genomics Platform for data generation, James Bochicchio for project management, and Ted Holzman (FHCRC) for assistance with computational analysis.

**References**

1. WHO | World Malaria Report 2014 [Internet]. WHO. Available from: http://www.who.int/malaria/publications/world\_malaria\_report\_2014/report/en/

2. Hall BF, Fauci AS. Malaria control, elimination, and eradication: the role of the evolving biomedical research agenda. J Infect Dis 2009;200(11):1639–43.

3. Plowe CV, Alonso P, Hoffman SL. The potential role of vaccines in the elimination of falciparum malaria and the eventual eradication of malaria. J Infect Dis 2009;200(11):1646–9.

4. RTS,S Clinical Trials Partnership. First Results of Phase 3 Trial of RTS,S/AS01 Malaria Vaccine in African Children. New England Journal of Medicine 2011;365(20):1863–75.

5. RTS,S Clinical Trials Partnership. A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. N Engl J Med 2012;367(24):2284–95.

6. RTS,S Clinical Trials Partnership. Efficacy and safety of the RTS,S/AS01 malaria vaccine during 18 months after vaccination: a phase 3 randomized, controlled trial in children and young infants at 11 African sites. PLoS Med 2014;11(7):e1001685.

7. RTS,S Clinical Trials Partnership. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. Lancet 2015;

8. Alonso PL, Sacarlal J, Aponte JJ, et al. Efficacy of the RTS,S/AS02A vaccine against Plasmodium falciparum infection and disease in young African children: randomised controlled trial. Lancet 2004;364(9443):1411–20.

9. Bejon P, Lusingu J, Olotu A, et al. Efficacy of RTS,S/AS01E vaccine against malaria in children 5 to 17 months of age. N Engl J Med 2008;359(24):2521–32.

10. Asante KP, Abdulla S, Agnandji S, et al. Safety and efficacy of the RTS,S/AS01E candidate malaria vaccine given with expanded-programme-on-immunisation vaccines: 19 month follow-up of a randomised, open-label, phase 2 trial. Lancet Infect Dis 2011;11(10):741–9.

11. Abdulla S, Oberholzer R, Juma O, et al. Safety and immunogenicity of RTS,S/AS02D malaria vaccine in infants. N Engl J Med 2008;359(24):2533–44.

12. Aponte JJ, Aide P, Renom M, et al. Safety of the RTS,S/AS02D candidate malaria vaccine in infants living in a highly endemic area of Mozambique: a double blind randomised controlled phase I/IIb trial. Lancet 2007;370(9598):1543–51.

13. Agnandji ST, Asante KP, Lyimo J, et al. Evaluation of the safety and immunogenicity of the RTS,S/AS01E malaria candidate vaccine when integrated in the expanded program of immunization. J Infect Dis 2010;202(7):1076–87.

14. Lalvani A, Moris P, Voss G, et al. Potent induction of focused Th1-type cellular and humoral immune responses by RTS,S/SBAS2, a recombinant Plasmodium falciparum malaria vaccine. J Infect Dis 1999;180(5):1656–64.

15. Sun P, Schwenk R, White K, et al. Protective immunity induced with malaria vaccine, RTS,S, is linked to Plasmodium falciparum circumsporozoite protein-specific CD4+ and CD8+ T cells producing IFN-gamma. J Immunol 2003;171(12):6961–7.

16. Kester KE, Cummings JF, Ofori-Anyinam O, et al. Randomized, double-blind, phase 2a trial of falciparum malaria vaccines RTS,S/AS01B and RTS,S/AS02A in malaria-naive adults: safety, efficacy, and immunologic associates of protection. J Infect Dis 2009;200(3):337–46.

17. Campo JJ, Aponte JJ, Skinner J, et al. RTS,S Vaccination Is Associated With Serologic Evidence of Decreased Exposure to Plasmodium falciparum Liver- and Blood-Stage Parasites. Mol Cell Proteomics 2015;14(3):519–31.

18. Weedall GD, Preston BMJ, Thomas AW, Sutherland CJ, Conway DJ. Differential evidence of natural selection on two leading sporozoite stage malaria vaccine candidate antigens. Int J Parasitol 2007;37(1):77–85.

19. Esposito F, Lombardi S, Modiano D, et al. Prevalence and levels of antibodies to the circumsporozoite protein of Plasmodium falciparum in an endemic area and their relationship to resistance against malaria infection. Trans R Soc Trop Med Hyg 1988;82(6):827–32.

20. Gandhi K, Thera MA, Coulibaly D, et al. Next generation sequencing to detect variation in the Plasmodium falciparum circumsporozoite protein. Am J Trop Med Hyg 2012;86(5):775–81.

21. Bailey JA, Mvalo T, Aragam N, et al. Use of Massively Parallel Pyrosequencing to Evaluate the Diversity of and Selection on Plasmodium falciparum csp T-Cell Epitopes in Lilongwe, Malawi. J Infect Dis 2012;206(4):580–7.

22. Aragam NR, Thayer KM, Nge N, et al. Diversity of T cell epitopes in Plasmodium falciparum circumsporozoite protein likely due to protein-protein interactions. PLoS ONE 2013;8(5):e62427.

23. Good MF, Pombo D, Quakyi IA, et al. Human T-cell recognition of the circumsporozoite protein of Plasmodium falciparum: immunodominant T-cell domains map to the polymorphic regions of the molecule. Proc Natl Acad Sci USA 1988;85(4):1199–203.

24. Calle JM, Nardin EH, Clavijo P, et al. Recognition of different domains of the Plasmodium falciparum CS protein by the sera of naturally infected individuals compared with those of sporozoite-immunized volunteers. J Immunol 1992;149(8):2695–701.

25. Reece WHH, Pinder M, Gothard PK, et al. A CD4(+) T-cell immune response to a conserved epitope in the circumsporozoite protein correlates with protection from natural Plasmodium falciparum infection and disease. Nat Med 2004;10(4):406–10.

26. Aide P, Dobaño C, Sacarlal J, et al. Four year immunogenicity of the RTS,S/AS02(A) malaria vaccine in Mozambican children during a phase IIb trial. Vaccine 2011;29(35):6059–67.

27. Takala SL, Coulibaly D, Thera MA, et al. Dynamics of polymorphism in a malaria vaccine antigen at a vaccine-testing site in Mali. PLoS Med 2007;4(3):e93.

28. Gandhi K, Thera MA, Coulibaly D, et al. Variation in the circumsporozoite protein of Plasmodium falciparum: vaccine development implications. PLoS ONE 2014;9(7):e101783.

29. Thera MA, Doumbo OK, Coulibaly D, et al. A field trial to assess a blood-stage malaria vaccine. N Engl J Med 2011;365(11):1004–13.

30. Alloueche A, Milligan P, Conway DJ, et al. Protective efficacy of the RTS,S/AS02 Plasmodium falciparum malaria vaccine is not strain specific. Am J Trop Med Hyg 2003;68(1):97–101.

31. Enosse S, Dobaño C, Quelhas D, et al. RTS,S/AS02A malaria vaccine does not induce parasite CSP T cell epitope selection and reduces multiplicity of infection. PLoS Clin Trials 2006;1(1):e5.

32. Waitumbi JN, Anyona SB, Hunja CW, et al. Impact of RTS,S/AS02(A) and RTS,S/AS01(B) on genotypes of P. falciparum in adults participating in a malaria vaccine clinical trial. PLoS ONE 2009;4(11):e7849.

33. Gilbert PB. Comparison of competing risks failure time methods and time-independent methods for assessing strain variations in vaccine protection. Stat Med 2000;19(22):3065–86.

34. Edlefsen PT, Rolland M, Hertz T, et al. Comprehensive Sieve Analysis of Breakthrough HIV-1 Sequences in the RV144 Vaccine Efficacy Trial. PLOS Computational Biology 2015;11(2):e1003973.

35. NCBI Sequence Read Archive [Internet]. Available from: http://www.ncbi.nlm.nih.gov/sra

36. Van Der Laan MJ, Rubin D. Targeted Maximum Likelihood Learning. The International Journal of Biostatistics 2007;3(1):Article 4.

37. Lunn M, McNeil D. Applying Cox regression to competing risks. Biometrics 1995;51(2):524–32.

38. Follmann D, Proschan M, Leifer E. Multiple outputation: inference for complex clustered data by averaging analyses from independent data. Biometrics 2003;59(2):420–9.

39. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics 1979;6(2):65–70.

40. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;57(1):289–300.

**Figure Legends**

**Figure 1. Data Generation and Sample/Data Filtration in Children Aged 5-17 Months for the Primary Clinical Malaria Endpoint.**

**Figure 2. Description of Study Sites, Genomic Units and Number of Clones per Infection (Complexity of Infection or COI) in Children Aged 5−17 Months with the Primary Clinical Malaria Endpoint.**

Study sites (Panel A); Diagram of the Circumsporozoite (CS) protein, analyzed genomic units, polymorphic positions (LD positions indicated in red), and segment of CS in the vaccine construct (Panel B); Distributions of COI for the RTS,S/AS01 vaccine and control vaccine groups (Panel C); Frequencies of full CS C-terminus 3D7 match by study site (Panel D) and for all sites by malaria genomic unit defined by the full CS C-terminus, epitopes, polymorphic region DV10, and haplotype combining Th2R and Th3R positions in linkage disequilibrium (LD) (Panel E); polymorphic CS C-terminus amino acid positions with frequency between 1% and 99% (Panel F). The denominator of the frequencies in Panels D through F is the number of primary clinical malaria endpoints with sequence data.

Ag = Agogo; Ba = Bagamoyo; Kil = Kilifi; Kin = Kintampo; Kom = Kombewa; Kor = Korogwe; La = Lambaréne; Li = Lilongwe; Si = Siaya; Na = Nanoro.

**Figure 3. Cumulative Incidences and Vaccine Efficacies (VEs) Against the Primary Clinical Malaria Endpoint in Children Aged 5-17 Months with Parasites Matched and Mismatched to the 3D7 Full CS C-Terminus Haplotype.**

The cumulative incidence during 12 months of post-vaccination follow-up in RTS,S/AS01 and control vaccine recipients in 3D7 matched cases (Panel A), and 3D7 mismatched cases (Panel B). Panel C shows the cumulative VE against 3D7 matched and 3D7 mismatched malaria over the entire post-vaccination follow-up period, and Panel D shows the cumulative and hazard ratio VE against 3D7 matched and 3D7 mismatched malaria at 12 months post-vaccination.

**Figure 4. Forest Plot of Vaccine Efficacy (VE) against the Primary Clinical Malaria Endpoint in Children Aged 5-17 Months with parasites Matched and Mismatched to the 3D7 Full CS C-Terminus and Epitope/Region Haplotypes.**

Cumulative VE (Panel A); Hazard Ratio VE (Panel B). Estimates were stratified by study site.

**Table 1: Cumulative vaccine efficacy (VE) against 3D7 matched and 3D7 mismatched primary clinical malaria through 12 months after vaccination by circumsporozoite (CS) C-terminus amino acid position in children aged 5-17 months. Estimates were stratified by study site.**

|  |  |  |  |
| --- | --- | --- | --- |
| Amino Acid | Haplotype-Matched Efficacy† | Haplotype-Mismatched Efficacy† |  Differential Efficacy |
| Position‡ | VE (%) | 95% CI | P-value | VE (%) | 95% CI | P-value | P-value | FWER P-value | Q-value |
| DV10  |
|  294 | 34.8 | (30.8, 38.6) | < 0.001 | 31.3 | (-6.4, 55.6) | 0.09 | 0.83 | 1.00 | 0.89 |
|  295 | 34.9 | (31.0, 38.6) | < 0.001 | 12.6 | (-78.8, 57.2) | 0.71 | 0.44 | 1.00 | 0.71 |
|  296 | 34.8 | (30.9, 38.5) | < 0.001 | 7.1 | (-310.9, 79.0) | 0.92 | 0.67 | 1.00 | 0.89 |
|  298 | 33.5 | (29.1, 37.7) | < 0.001 | 41.7 | (30.1, 51.4) | < 0.001 | 0.20 | 1.00 | 0.41 |
|  299\* | 35.4 | (31.5, 39.1) | < 0.001 | -62.8 | (-186.0, 7.3) | 0.09 | 0.003 | 0.08 | 0.08 |
|  301\* | 49.2 | (35.3, 60.1) | < 0.001 | 33.1 | (28.9, 37.0) | < 0.001 | 0.03 | 0.81 | 0.15 |
|  302 | 34.8 | (30.9, 38.5) | < 0.001 | 20.0 | (-108.2, 69.3) | 0.64 | 0.69 | 1.00 | 0.89 |
|  303 | 34.6 | (30.7, 38.3) | < 0.001 | 68.4 | (-100.9, 95.0) | 0.22 | 0.45 | 1.00 | 0.71 |
| Th2R |
|  314LD | 32.7 | (26.2, 38.2) | < 0.001 | 37.3 | (30.9, 43.0) | < 0.001 | 0.33 | 1.00 | 0.61 |
|  317LD\* | 45.9 | (33.3, 56.2) | < 0.001 | 33.1 | (28.9, 37.1) | < 0.001 | 0.06 | 1.00 | 0.19 |
|  318 | 36.9 | (25.8, 46.4) | < 0.001 | 34.2 | (29.8, 38.4) | < 0.001 | 0.65 | 1.00 | 0.89 |
|  320 | 34.4 | (30.4, 38.1) | < 0.001 | 64.0 | (21.0, 83.6) | 0.01 | 0.14 | 1.00 | 0.32 |
|  321 | 35.7 | (24.3, 45.3) | < 0.001 | 34.5 | (30.1, 38.7) | < 0.001 | 0.85 | 1.00 | 0.89 |
|  322 | 34.8 | (27.2, 41.5) | < 0.001 | 34.7 | (29.4, 39.6) | < 0.001 | 0.99 | 1.00 | 0.99 |
|  324 | 37.3 | (32.0, 42.2) | < 0.001 | 30.7 | (23.3, 37.4) | < 0.001 | 0.16 | 1.00 | 0.35 |
|  327 | 35.5 | (31.2, 39.5) | < 0.001 | 30.1 | (16.5, 41.4) | < 0.001 | 0.42 | 1.00 | 0.71 |
| Th3R |
|  349 | 34.8 | (30.9, 38.5) | < 0.001 | 25.4 | (-84.1, 69.8) | 0.53 | 0.81 | 1.00 | 0.89 |
|  352 LD | 35.1 | (30.7, 39.2) | < 0.001 | 32.8 | (20.0, 43.5) | < 0.001 | 0.73 | 1.00 | 0.89 |
|  354 LD\* | 36.0 | (32.0, 39.8) | < 0.001 | 10.8 | (-22.6, 35.1) | 0.48 | 0.05 | 1.00 | 0.17 |
|  355 | 34.7 | (30.8, 38.4) | < 0.001 | 54.1 | (-286.1, 94.5) | 0.47 | 0.71 | 1.00 | 0.89 |
|  356 LD\* | 36.2 | (32.2, 40.1) | < 0.001 | 15.8 | (-7.8, 34.3) | 0.17 | 0.04 | 0.85 | 0.15 |
|  357 LD | 35.3 | (27.3, 42.5) | < 0.001 | 34.4 | (29.3, 39.2) | < 0.001 | 0.86 | 1.00 | 0.89 |
|  359\* | 36.1 | (32.0, 40.1) | < 0.001 | 22.2 | (5.2, 36.2) | 0.01 | 0.07 | 1.00 | 0.20 |
|  361\* | 39.3 | (33.6, 44.4) | < 0.001 | 29.3 | (22.4, 35.5) | < 0.001 | 0.03 | 0.81 | 0.15 |

CI denotes confidence interval, and FWER p-value adjusted39 p-value controlling the familywise error rate.

† For each amino acid position, haplotype-matched (mismatched) VE was computed only including clinical malaria endpoint events with a 3D7 matched (mismatched) amino acid at the given position.

‡ Only amino acid positions with sufficiently high minor allele frequency were included in the analysis.

LD Linkage disequilibrium haplotype includes these Th2R and Th3R amino acid positions.

\* Statistically significant differential efficacy was defined as q-value ≤ 0.2 for all 28 multiply compared haplotype loci (all epitope regions and amino acid positions with sufficiently high minor allele frequency).