Combining machine learning and mathematical models of disease dynamics to guide development of novel disease interventions

Authors: Monica Golumbeanu[1,2]†, Guojing Yang[1,2]†, Flavia Camponovo[1,2,3], Erin M. Stuckey[4], Nicholas Hamon[5], Mathias Mondy[5], Sarah Rees[5], Nakul Chitnis[1,2], Ewan Cameron[6,7,8], and Melissa A. Penny[1,2]*

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[3]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA
[4]The Bill and Melinda Gates Foundation, Seattle, WA, USA
[5]Innovative Vector Control Consortium, Liverpool, United Kingdom
[6]Malaria Atlas Project, Big Data Institute, University of Oxford, Oxford, UK
[7]Curtin University, Perth, Australia
[8]Telethon Kids Institute, Perth Children's Hospital, Perth, Australia
†These authors contributed equally.
Email: *Correspondence to: melissa.penny@unibas.ch

One Sentence Summary: Defining quantitative profiles of novel disease interventions by combining machine learning with mathematical models of disease transmission

Abstract:

The development of novel interventions against a disease entails optimising their specifications to achieve desired health goals such as disease reduction. As testing is limited early in development, it is difficult to predefine these optimal specifications, prioritize or continue investment in candidate interventions. Mathematical models of disease can provide quantitative evidence as they can simulate deployment and predict impact of a new intervention considering deployment, health-system, population and disease characteristics. However, due to large uncertainty early in development, as well as model complexity, testing all possible combinations of interventions and deployments becomes infeasible. As a result, mathematical models have been only marginally used during intervention development to date. Here, we present a new approach where machine learning enables the use of detailed disease models to identify optimal properties of candidate interventions to reach a desired health goal and guide development. We demonstrate the power of our approach by application to five novel malaria interventions under development. For various targeted reductions of malaria prevalence, we quantify and rank intervention characteristics which are key determinants of health impact. Furthermore, we identify minimal requirements and tradeoffs between operational factors, intervention efficacy and duration to achieve different levels of impact and show how these vary across disease transmission settings. When single interventions cannot achieve significant impact, our method allows finding optimal combinations of interventions fulfilling the desired health goals. By enabling efficient use of disease models, our approach supports decision-making and resource investment in the development of new interventions for infectious diseases.

Significance Statement

During development of novel disease interventions (e.g. vaccines), a target product profile (TPP) document defines intervention characteristics required to meet health goals. As clinical trials are limited early in development, mathematical models simulating disease dynamics can help define TPPs. However, testing all combinations of intervention, delivery and environment characteristics is infeasible and so complex mathematical models have not been used until now. We introduce a new approach to define TPPs, combining models of disease with machine learning. We examined several novel malaria interventions, identifying key characteristics, minimum efficacy and duration of effect that ensure significant reductions in malaria prevalence. This approach therefore enabled mathematical models of disease to support intervention development, by identifying intervention requirements that ensure public health impact.

Introduction

Target Product Profiles (TPPs) are dynamic documents commonly used during the development of a cutting-edge medical product, defining its required characteristics to fulfill an unmet health need (*1*). By offering a comprehensive snapshot of the development process at any given point in time, a TPP constitutes a vital reference for dialogue between various stakeholders to guide decisions on the development direction to be pursued (*1-5*). A well-constructed TPP is thus essential for efficient resource allocation and success during the development phase (*1, 5*). However, the process of establishing TPPs relies on minimal clinical or quantitative evidence. They are often set by expert opinion and consensus based on limited quantitative consideration of the complex dynamics of disease or predictions of the likely intervention impact while achieving the identified unmet health need (*6*). Furthermore, few TPPs consider operational aspects such as deployment coverage in addition to product-specific characteristics such as efficacy or half-life. This has implications for the appropriate definition of intervention effectiveness characteristics according to local health systems and health targets (*1, 6, 7*).

Mathematical models of disease transmission dynamics can be used to bridge this gap, as they quantitatively estimate the impact of interventions while including considerable evidence of disease progression and transmission, host immunity, as well as environmental or health system dynamics and their interaction with interventions (*8, 9*) (Fig. 1). However, models have been mainly used at late stages during the development of a new intervention; for example, to predict likely impact or cost-effectiveness from data collected in Phase 3 clinical trials (*10-15*). Model investigations are usually informed by scenario analysis accounting for the delivery and target age groups, as well as properties of the new intervention pre-defined or informed by late clinical trials (*16-19*). In these constrained scenarios, high model and parameter complexity tend to obscure the complex relationships between intervention parameters, operational factors, health outcomes and public health impact (*20*). Exhaustive scenario analyses are highly computationally expensive, rendering the full exploration of all possible interventions for a disease, in conjunction with all possible delivery scenarios, combinatorically infeasible.

Here we propose a new ethos where epidemiological models guide the development of novel disease interventions designed to achieve quantified health goals. To do this efficiently, we use

machine learning combined with mathematical models to perform a directed search of the entire space of intervention profiles, to define properties of new interventions (sometimes referred to as "tools" or "products") that will achieve the desired health goal. Placing the end goal of public health impact at the center of decision making is increasingly important to direct Research & Development (R&D) efforts in the face of finite resources. The use of mathematical models enables translation of R&D efforts into potential impact. In this paper, we show how modelling can support this process, and introduce a framework that quantitatively defines product characteristics within TPPs.

Previous approaches using disease models to inform TPPs have tackled the combinatorically-complex parameter space by only exploring a discrete, constrained set of parameters (*21-23*). These approaches have provided insightful knowledge and emphasized the importance of using disease models for defining TPPs. Nevertheless, they have provided a concomitantly constrained view of intervention specifications. Our framework tackles and moves beyond these challenges. On one hand, it allows us to rigorously define TPPs by efficiently exploring highly complex parameter spaces of mathematical disease models, and on the other hand it identifies the determinants of desired public health impact to inform tradeoffs between product characteristics and use-cases. Furthermore, as the ultimate health goal guides decisions on interventions and for optimal use of the supportive framework presented here, an engaging, iterative exchange with stakeholders to define desired outcomes and the likely delivery use-cases of the new interventions is essential.

Our framework utilizes a machine learning approach using Gaussian processes (GPs) (*24*) to generate computationally light emulators of detailed mathematical models of disease dynamics (Fig. 2A). These emulators constitute an interface that easily links properties of deployed interventions and operational factors to health goals. Furthermore, the emulators capture not just the mean tendency of complex disease models dynamics, but also the inherent variance caused by the stochasticity in the models (*25*). Disease model emulators allowed us to efficiently perform sensitivity analyses of intervention and health system parameters on predicted public health impacts at low computational cost. Furthermore, by coupling emulators with nonlinear optimization techniques, we constructed a predictive framework that identifies key determinants of intervention impact as well as the minimal intervention profiles required for achieving a given

114 health goal (Fig. 2). The framework consists of (i) a comprehensive disease progression and
115 transmission simulation model applied on a discrete, uniformly sampled set of input parameters;
116 (ii) training of an emulator on the sampled set of parameters and corresponding impact
117 outcomes; (iii) using sensitivity analysis to understand drivers of intervention impact; and (iv)
118 applying a non-linear constrained optimization algorithm to explore intervention operational and
119 effectiveness characteristics meeting various targets and deployment use-cases specified
120 following iterative consultation with product development experts. A detailed description of the
121 components of the developed framework can be found in Fig. 2A and the Materials and Methods
122 section.

123 We apply our framework to assess and optimize new interventions for preventing malaria
124 transmission. Strategic investment in new interventions is becoming crucial for malaria control
125 and elimination programs, as existing interventions are currently challenged by increasing
126 resistance (*26-28*). Mathematical models of malaria transmission (Fig. 1) have been used
127 extensively to estimate the impact of malaria interventions and to optimize intervention packages
128 for specific geographies (*10, 29-32*). As yet, these malaria models have not been systematically
129 applied in directing the design of new interventions, nor in understanding how intervention-
130 specific, epidemiological and systems factors jointly contribute to impact. Following
131 consultation with malaria product development experts, we used our new framework to define
132 the required profiles in terms of coverage, efficacy and duration characteristics in TPPs of new
133 putative malaria interventions to reach desired public health goals such as prevalence reduction
134 contingent on operational constraints (Fig. 2).

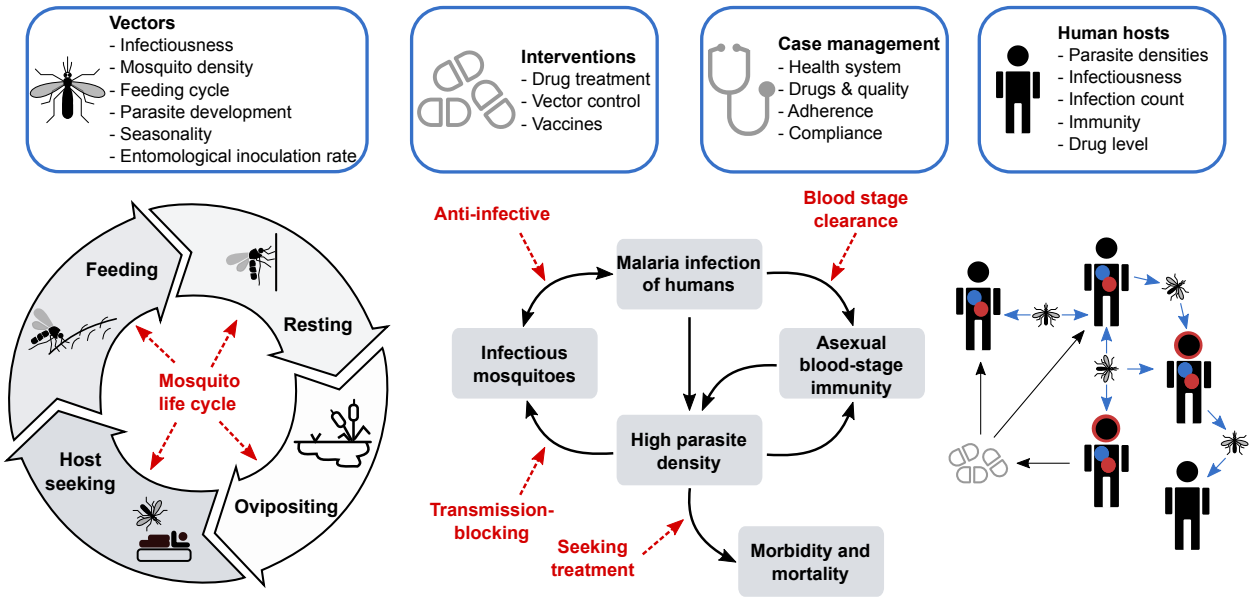


Fig. 1 Schematic of a stochastic simulation platform of malaria transmission dynamics. *OpenMalaria (33) is an open source, stochastic, individual-based model which simulates malaria epidemiology and transmission dynamics across humans and mosquitoes. The pattern of yearly malaria infection in the absence of interventions is determined by the entomological inoculation rate (EIR), which is a model input. Each infected human host in the simulated population has an associated parasite density and duration of infection, where each infection is also modelled individually, and follows a modelled transmission cycle (central diagram) which captures effects such as immunity, infectiousness to mosquitoes, morbidity or mortality. During the simulation, a wide range of human and vector interventions can be applied, affecting the transmission cycle at various stages (red arrows). Setting-specific characteristics such as population demographics, mosquito species entomological characteristics or seasonality are explicitly modelled. Various health outcomes are monitored over time including patent infections, uncomplicated clinical disease, severe disease in and out of hospital and malaria mortality (a detailed description of the simulation model and its features is provided in Materials and Methods and Table S1.1).*

Results

A disease model and machine learning approach to quantitatively define malaria interventions

Our analysis workflow (Fig. 2) starts with sensibly-informed TPP scenarios; the definition of targeted health goals corresponding to unmet health needs; and possible use cases following continuous consultation with product development experts. The health goals in the present analysis are reductions of malaria prevalence for all ages (*Pf*$PR_{0\text{-}99}$) and prevention of resurgence. Next, within the *"Disease model"* component, malaria transmission is modeled by the means of the established, stochastic, individual-based model OpenMalaria (*33*) (Fig. 1, Table S1.1). A comprehensive set of simulated scenarios is built by uniformly sampling the parameter space (defined by the parameters emphasized in bold under *"Tool specifications"* and *"Setting"* components in Fig. 2 and detailed in Table S2.1). The scenarios are simulated with the disease model yielding an extensive database of disease outcomes. In the machine learning part of the approach, the database of simulated scenarios and corresponding outcomes is used to train a predictive model, in this case a Heteroskedastic Gaussian process model (see detailed training procedure in Materials and Methods). The predictive model acts as an emulator of the complex individual-based mathematical model. Specifically, the emulator can predict the disease outcome for the given health goal and any set of input parameters. For this reason, the trained emulator can be efficiently and promptly used in downstream analyses to design TPPs of new malaria interventions, and to identify their quantitative properties to meet the health goal previously defined. More precisely, sensitivity analysis allows searching for key determinants of intervention impact, while constrained optimization analyses yield the optimal required intervention properties that meet specified impactful health goals.

We used our validated individual-based model, "OpenMalaria" (*29, 33*) (detailed description in the Materials and Methods section, Fig. 1 and Table S1.1) to simulate malaria epidemiology and transmission dynamics within various transmission settings. These settings cover a broad spectrum of transmission and mosquito biting behavior archetypes relevant for attaining general guiding principles in the early development phase of new interventions. Within the simulations, we quantitatively examined several malaria interventions that are currently under development or developed within the last ten years, including monoclonal antibodies, drugs, vaccines as well as novel vector control interventions. To be able to investigate a wide range of interventions,

instead of considering their characteristics explicitly, we more generally modelled their action on parasite or vector targets (Fig. 1). Accordingly, each intervention was modeled through its deployment coverage, efficacy, half-life or duration of effect on the given parasite or vector stage in the transmission cycle (see Materials and Methods, Fig. S2.3, Table S2.1 for detailed intervention specifications). For simplification, the words 'half-life and 'duration' are used interchangeably to describe the longevity of the intervention effect (further details and definitions in Materials and Methods).

Intervention impact in the current study was assessed assuming a single health goal of malaria prevalence reduction and thus through predicted reduction in *Plasmodium falciparum* malaria prevalence across all ages, $Pf\mathrm{PR}_{0\text{-}99}$, corresponding to true infection prevalence and not patent (detected with a diagnostic such as rapid diagnostic test (RDT), or polymerase chain reaction (PCR), Fig. S2.2, S3.1-S3.4). We learned simplified predictive emulators for the OpenMalaria simulation results by training GP models on a limited set of simulated scenarios (Fig. 2B). We show that the trained GP models accurately capture the dependencies between the disease model input parameters and the output intervention impact, and are able to reliably predict the reduction in $Pf\mathrm{PR}_{0\text{-}99}$ attributable to any input intervention characteristics (Fig. 2B, S4.1-S4.3, Table S4.1). Our work thus builds on recent applications of GPs in disease modelling and burden prediction for malaria (*34*). Using the trained GP emulator, through global sensitivity analysis, we evaluated the key determinants of intervention impact (Fig. 2C). In addition, we performed a constrained search for intervention and delivery profiles (TPPs) that maximize impact under a particular health goal, given concrete, expert-informed, operational constraints such as possible deployment coverage, or feasible intervention properties such as efficacy or duration of protection (Fig. 2D).

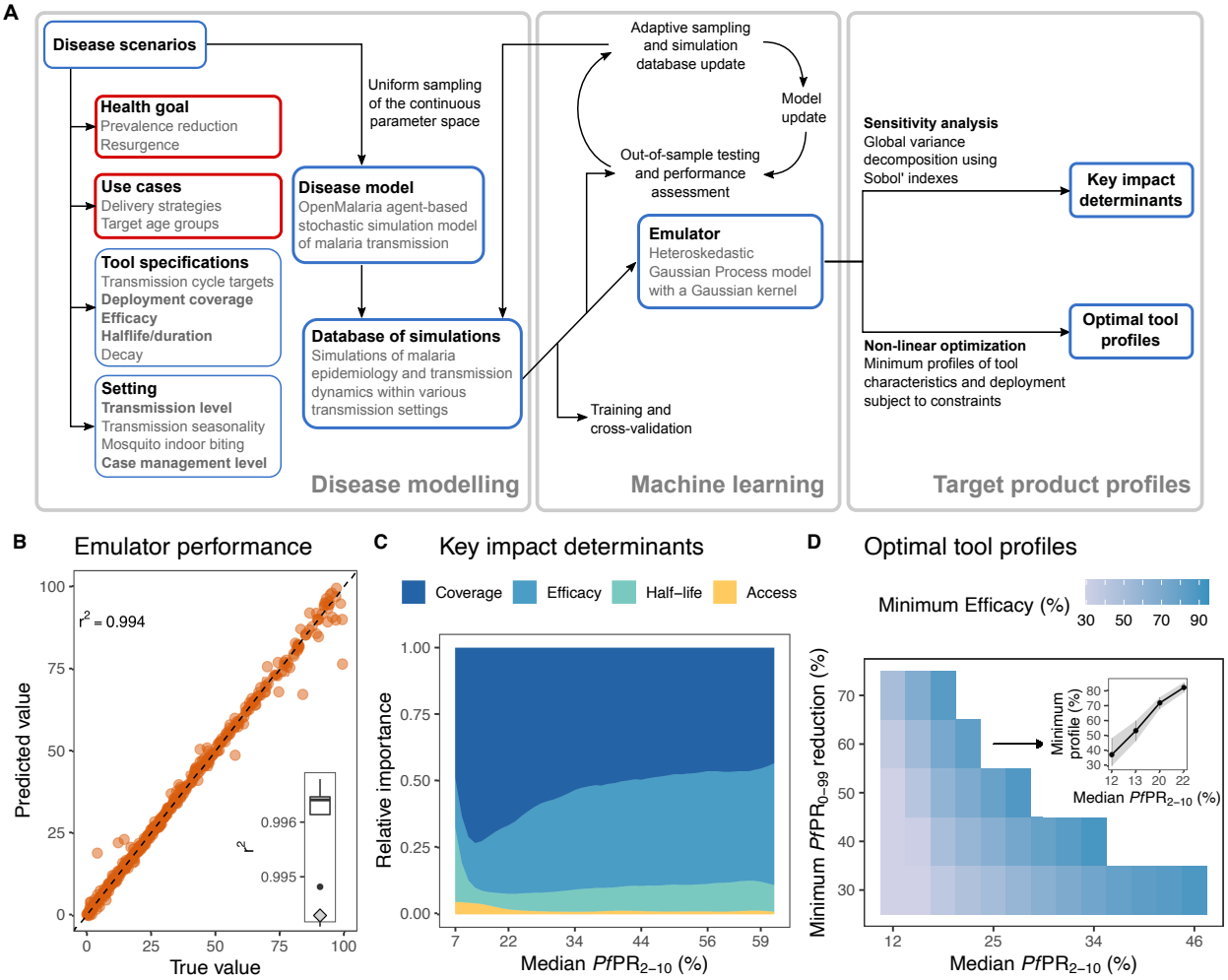


Fig. 2 Quantitatively defining TPPs of novel disease interventions.
(A) Detailed schematic representation of the proposed quantitative framework to support product development (full specifications in Materials and Methods). Figures ***(B)-(D)*** *present the results of applying the framework for an anti-infective malaria vaccine (mass administration, seasonal transmission with high indoor mosquito biting):* ***(B)*** *Correlation between simulated true (x axis) and emulator predicted (y axis) $PfPR_{0\text{-}99}$ reduction with a GP emulator trained in a cross-validation scheme (Pearson correlation coefficient r^2 distribution shown in boxplot) and validated on an out-of-sample test set (r^2 left upper corner and grey diamond on the boxplot in the right lower corner).* ***(C)*** *Example vaccine impact determinants: the colors represent proportions of the emulator output variance (relative importance) attributable to intervention specifications, as well as health system access.* ***(D)*** *Example feasible landscape of optimal vaccine efficacy profiles for various health goals (minimum targeted $PfPR_{0\text{-}99}$ reductions, y axis). For each health goal, the heatmap displays the minimum required efficacy when applied at a coverage of 60% and with a half-life of 7 months, assuming an access to care level of 25% (example in the insert plot for a target reduction of at least 60%). Results in figures (C) and (D) are displayed for a range of median simulated true $PfPR_{2\text{-}10}$ (before intervention deployment, rounded values, x axis).*

227

228 *Intervention impact and the importance of their characteristics*

229 With guidance from different groups of experts and partners (see definition of the various
230 stakeholders involved in Materials and Methods), we conducted an extensive analysis in the
231 malaria development space, covering a diverse spectrum of interventions pertaining to 1) anti-
232 infective monoclonal antibodies 2) anti-infective vaccines, 3) transmission-blocking vaccines, 4)
233 outdoor attractive targeted sugar baits, and 5) eave tubes. Following simulation with
234 OpenMalaria of deployment of each of these interventions through mass administration
235 campaigns over several years (see Materials and Methods), we first analyzed the predicted
236 distributions of reduction in true *Pf*$PR_{0\text{-}99}$ (Fig. 3A, S3.2-S3.4). We found that, in general, when
237 aiming for substantial, prompt reductions in prevalence for this particular health target, vector
238 control was by far the most impactful intervention across all settings. Monoclonal antibodies,
239 anti-infective and transmission-blocking vaccines had a more pronounced impact in low-
240 transmission settings compared to endemic settings (Fig. 3A, S3.1-S3.4, Table S5.1).

241 Sensitivity analysis indicated that the impacts of these interventions on malaria prevalence were
242 driven by different characteristics of their efficacy profiles, deployment strategies, or access to
243 care for treatment of clinical cases, for either short and long impact follow-up (Fig. 3B-E, S6.1 –
244 6.2). Across a large proportion of the simulated scenarios, over all parasite and vector targets and
245 interventions, coverage of the deployed intervention was overwhelmingly the primary driver of
246 impact especially in low transmission settings (Fig. 3B-E, S6.1-S6.2). For therapeutic
247 interventions, the impact of short-term passive immunizations such as monoclonal antibodies
248 relied on their deployment coverage and the health system (Fig. 3B, S6.1). In contrast, for long-
249 acting interventions such as vaccines, impact was driven by deployment coverage and efficacy
250 (Fig. 3C, S6.1). Highly-efficient vector control interventions such as attractive targeted sugar
251 baits had a strong effect on prevalence (Fig. 3A), and their duration of effect was the most
252 important determinant (Fig. 3D, S6.2). The immediate impact of long-term vector control
253 interventions such as eave tubes was driven by deployment coverage, while their half-life was a
254 key determinant for preventing resurgence (Fig. 3E, S6.2).
255

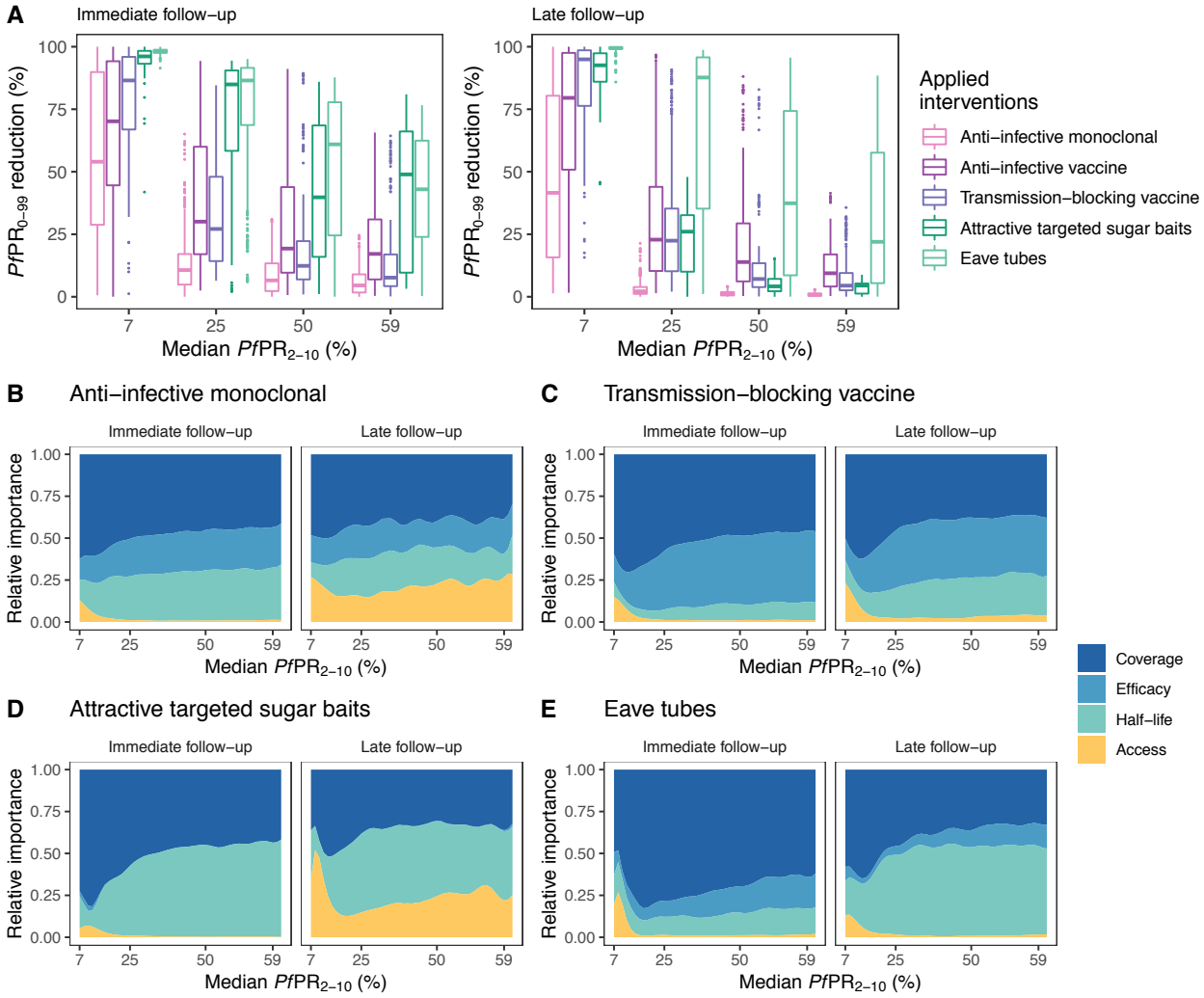


Fig. 3: Effects of novel malaria interventions on disease prevalence and their key drivers of impact.

(A) *Distribution of obtained reduction in $PfPR_{0\text{-}99}$ following deployment of various malaria interventions under development (shown with different colors) for a range of simulated transmission settings (specified by median true $PfPR_{2\text{-}10}$ rounded values, x axis). Each boxplot displays the interquartile range (box), the median value (horizontal line), the largest and smallest values within 1.5 times the interquartile range (whiskers), and the remaining outside values (points) of the $PfPR_{0\text{-}99}$ reduction values obtained across all the simulations for each given setting. The remaining plots of the figure present the results of sensitivity analysis showing, across the same simulated $PfPR_{2\text{-}10}$ settings, the determinants of intervention impact on $PfPR_{0\text{-}99}$ reduction for anti-infective monoclonal antibodies* ***(B)****, transmission-blocking vaccines* ***(C)****, attractive targeted sugar baits* ***(D)*** *and eave tubes* ***(E)****. Determinants of impact are shown for both immediate and late follow-up, when interventions are applied once per year for three years in a seasonal transmission setting with high indoor mosquito biting (full intervention specifications provided in Materials and Methods and results for other settings and interventions shown in Fig. S6.1-6.2 and Table S5.1).*

Minimal requirements of novel malaria interventions to achieve a defined health goal

For the five aforementioned malaria interventions, we explored their optimal profiles for a broad set of target *Pf*$PR_{0\text{-}99}$ reduction levels, creating landscapes of intervention profiles according to their minimal characteristics across various transmission settings (Fig. 4-5, S7.1-S8.5). These landscapes provide a broad and comprehensive overview of the intervention potential capabilities and limitations in achieving a desired health goal. For example, as opposed to an anti-infective monoclonal antibody which requires high efficacy and duration to achieve large *Pf*$PR_{0\text{-}99}$ reduction in only a limited number of settings (Fig. 4A-B, S7.1-S7.2), attractive targeted sugar baits that kill mosquitoes achieve a wider range of target *Pf*$PR_{0\text{-}99}$ reductions in high-transmission settings as well (Fig. 4C-D, S7.5). Similarly, while in settings with lower transmission (*Pf*$PR_{2\text{-}5}$ <30%), anti-infective and transmission-blocking vaccines had comparable requirements in achieving similar *Pf*$PR_{0\text{-}99}$ reduction targets, anti-infective vaccines showed a higher potential and reached additional targets in high-transmission, endemic settings (Fig. 5).

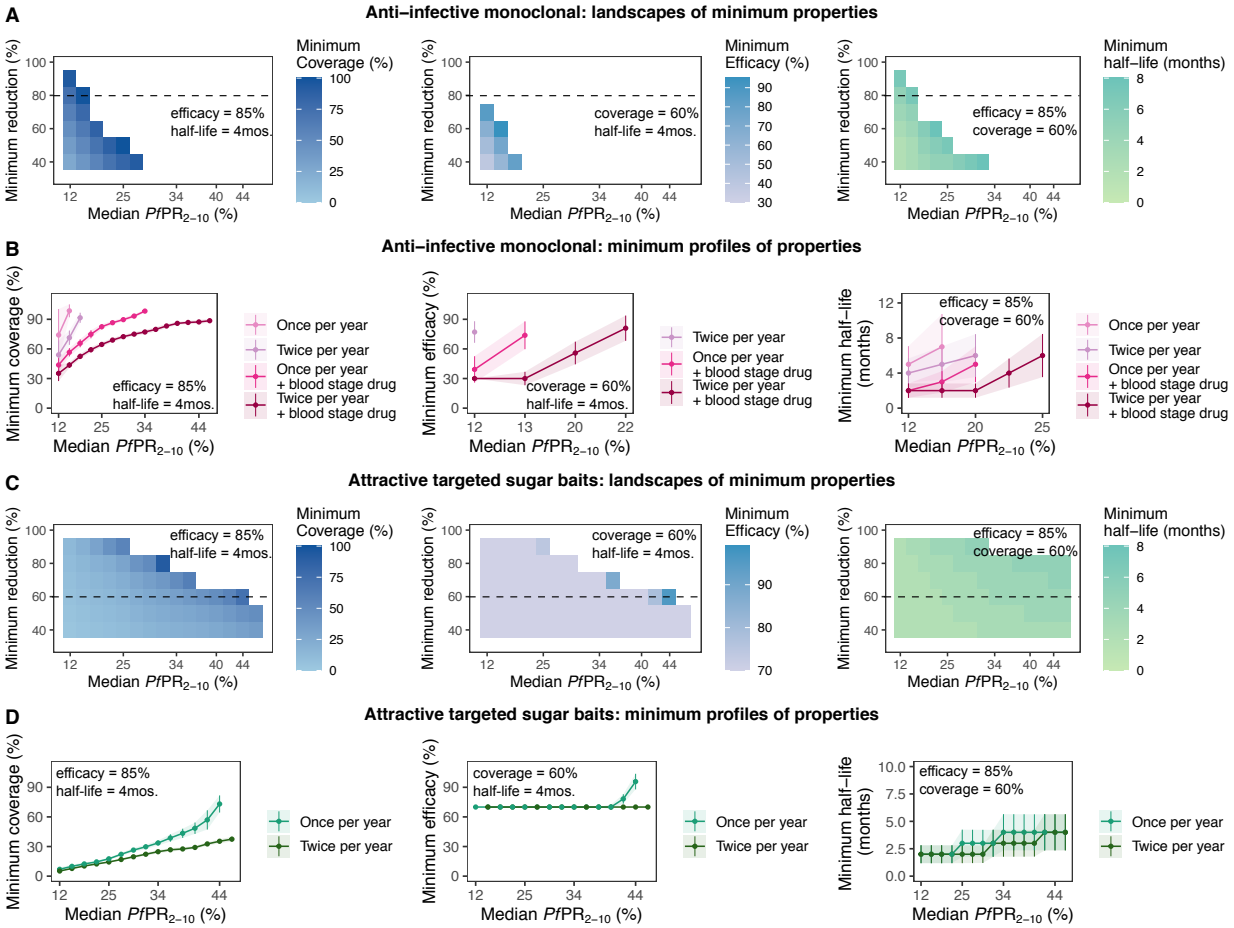


Fig. 4: Estimated optimal intervention and delivery profiles (TPPs) for monoclonal antibodies and attractive targeted sugar baits.

*The heatmaps in figures **(A)** and **(C)** represent landscapes of optimal, constrained intervention specifications (coverage, efficacy, and half-life) required to achieve a broad range of targeted minimal reductions in $PfPR_{0\text{-}99}$ (y axis) across different simulated true $PfPR_{2\text{-}10}$ settings (rounded values, x axis). Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (values marked on each figure). Results are shown for an anti-infective monoclonal antibody **(A)** and attractive targeted sugar baits **(C)**. For a defined health goal of reduction in $PfPR_{0\text{-}99}$ (dashed horizontal lines on figures **(A)** and **(C)**), the corresponding minimum product profile requirements are shown for an anti-infective monoclonal antibody in **(B)** and attractive targeted sugar baits in **(D)**. Both figures **(B)** and **(D)** show how these requirements change when these interventions are delivered at various frequencies (once or twice per year), and when the anti-infective monoclonal antibody is delivered in combination with a blood-stage drug. The simulated health system access was 25%. Descriptions of all intervention properties for identification of minimal profiles are detailed in Table S2.2, while additional results for other settings and interventions are provided in Fig. S7.1-S7.6 and Table S5.1.*

307 For a detailed overview of landscapes of intervention profiles for all simulated settings and
308 interventions see Fig. S7.1-S7.6. These landscapes together with the results of the sensitivity
309 analysis offer an evidence-based prioritization of resources during the product development
310 process. For example, we found that while both efficacy and half-life are important for
311 immediate prevalence reductions with monoclonal antibodies, their effect is limited in preventing
312 resurgence and is only supported by high case-management levels (Fig. 3, 4, S6.1, S7.1-S7.2).
313 Conversely, the efficacy of anti-infective vaccines drives immediate impact, whereas half-life of
314 effect has greater importance for achieving and maintaining *Pf*PR_{0-99} reductions (Fig. 3, 5, S6.1,
315 S7.3-S7.4). These results suggest that if vaccines and monoclonal antibodies are to support
316 preventing resurgence, then R&D efforts should focus on increasing and establishing antibody
317 longevity.

318 Our analysis shows that coverage is the primary driver of impact (Fig. 3B-E, S6.1-S6.2). This
319 result has important implications for interventions requiring multiple applications to achieve high
320 efficacy, indicating that it is of crucial importance to target both vulnerable populations and the
321 proportion of the population missed by the intervention. While for some interventions high
322 coverage deployment might be very difficult or impossible to achieve, our analysis shows that
323 this can be alleviated by increasing the deployment frequency or through deploying
324 combinations of interventions, which may have cost implications (Fig. 4B, 4D, 5B, 5D, S7.1-
325 S7.5, S8.1-8.4).

326 We found that combining several interventions targeting different stages in the transmission
327 cycle can strongly affect the minimum requirements of a putative new intervention, potentially
328 increasing the impact of an otherwise weaker intervention. For example, for an anti-infective
329 monoclonal antibody with an initial half-life of 7 months and deployed at a coverage of 60%
330 reflecting completion of multiple doses, achieving a prevalence reduction of 80% was impossible
331 when deployed once yearly for three years (Fig. 4A, S7.1). Furthermore, achieving the
332 aforementioned health goal required an efficacy of 80% when the intervention was deployed
333 twice per year for three years (Fig. 4B, Fig. S7.2). However, when deployment of the
334 monoclonal antibody was coupled with a short half-life blood-stage parasite treatment such as
335 dihydroartemisinin-piperaquine or artemether-lumefantrine, its minimum required efficacy was
336 considerably reduced for both delivery frequencies (Fig. 4B, S7.1-S7.2, S8.1). Conversely, if we

assume an initial efficacy of 85% for the monoclonal antibody, we find that its minimal required half-life can be reduced if we deploy this intervention in combination with the blood-stage parasite clearing drug (Fig. 4B, S7.1-S7.2, S8.1). These results partly motivated the current development of anti-infective monoclonal antibodies; use-cases will likely include deployment with existing or new antimalarial treatment.

We also showed that a modified deployment schedule could reduce requirements for properties of some interventions. For example, for highly-efficacious attractive targeted sugar baits, higher coverage and half-life were required when implemented once per year for three years compared with an accelerated delivery schedule of twice per year for three years (Fig. 4C-D). Except for high transmission settings ($Pf\mathrm{PR}_{2\text{-}10} > 41\%$), a minimum required efficacy of 70% was sufficient to attain the desired health goal for the majority of settings and for both delivery schedules (Fig. 4C-D, Fig. S7.5, Fig. S8.4). This result is also reflected in the sensitivity analysis (Fig. 3D). Accordingly, the variation in intervention efficacy in the ranges investigated has little importance in driving the intervention impact and suggests that, once a vector control intervention such as attractive targeted sugar baits achieves a high killing efficacy (here greater than 70%), a next step of optimizing other intervention characteristics such as deployment coverage or duration leads to a higher impact. These results demonstrate the strength of our analysis in identifying the intervention characteristics to be prioritized for R&D.

When coupled with a short half-life blood-stage parasite treatment, requirements of coverage, efficacy and half-life were reduced also for anti-infective and transmission blocking vaccines to achieve the targeted reductions of $Pf\mathrm{PR}_{0\text{-}99}$ (Fig. 5, S7.3-7.4, S8.2-S8.3). In particular for high-transmission settings ($Pf\mathrm{PR}_{2\text{-}5}>25\%$), given an RTS,S-like half-life of 7 months, both anti-infective and transmission-blocking vaccines could not achieve any of the defined prevalence reduction goals if deployed singly. This was the case for any deployment coverage given an initial efficacy of 85% as well as for any efficacy given a deployment coverage of 60%. Combining vaccine deployment with a blood-stage drug not only significantly expanded the achievable health targets also to high-transmission settings, but also reduced vaccine properties requirements. Our analysis reveals that anti-infective vaccines had a higher potential than transmission-blocking vaccines, requiring less performance and achieving higher prevalence reductions targets also in higher transmission settings. When combined with blood-stage parasite

367 treatment, the coverage, efficacy and half-life requirements of anti-infective vaccines were lower
368 compared to those of transmission-blocking vaccines for the same prevalence reduction targets
369 (Fig. 5, S7.3, S7.4, S8.2, S8.3).

370 Our comprehensive analysis was applied to explore determinants of impact and required profiles
371 of interventions across two seasonal settings (seasonal and perennial) and three types of
372 mosquito biting patterns (low, medium and high indoor biting). A detailed overview of impact
373 determinants and optimal intervention profiles is presented in the Supplementary Materials (Fig.
374 S6.1-S8.5, and additional key results summarized in Table S5.1).
375

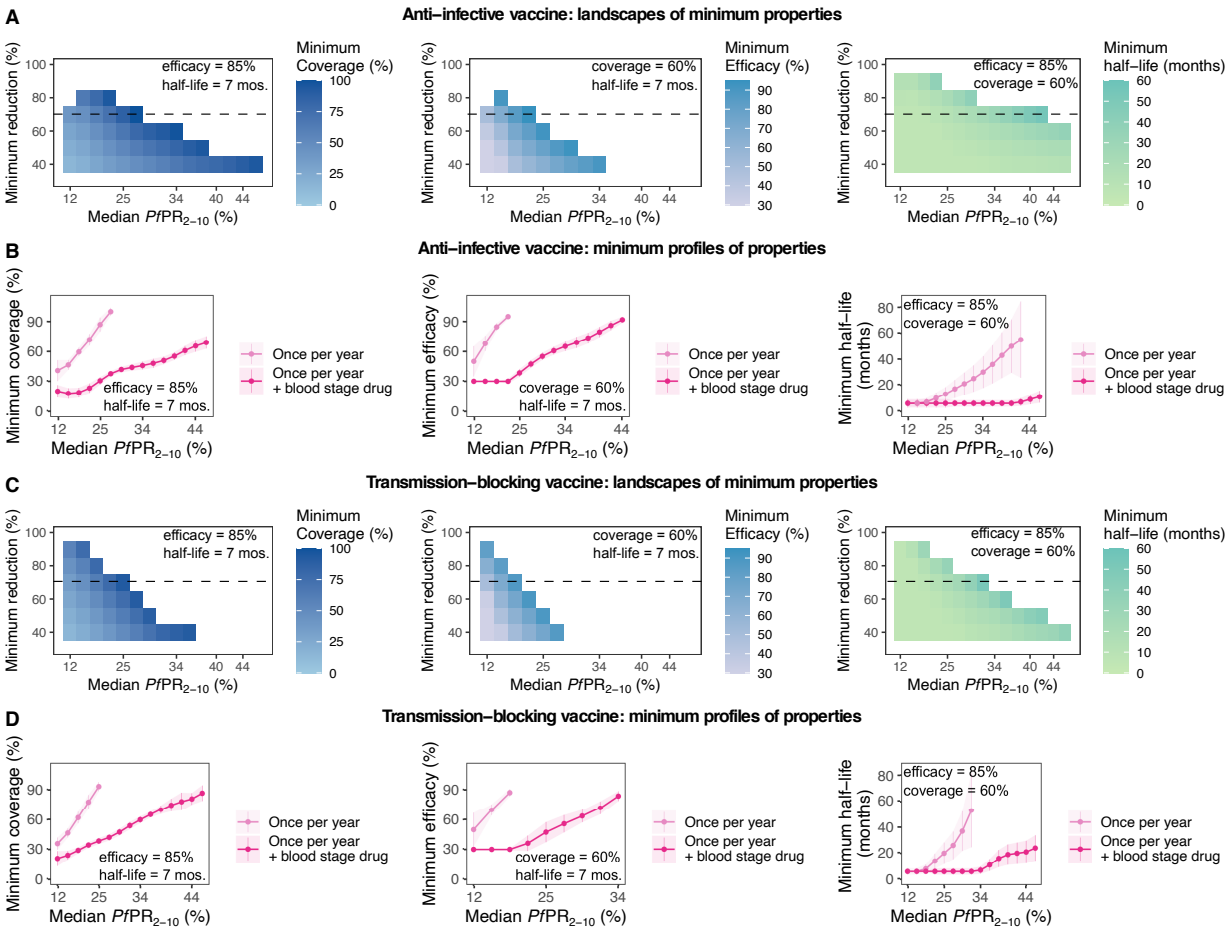


Fig. 5. Estimated optimal intervention and delivery profiles (TPPs) for anti-infective and transmission blocking vaccines deployed once per year with or without a blood stage clearing drug.

*The heatmaps in figures **(A)** and **(C)** represent landscapes of optimal, constrained intervention characteristic profiles (coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in $PfPR_{0\text{-}99}$, y axis) across different simulated true $PfPR_{2\text{-}10}$ settings (rounded values, x axis). Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (values marked on each figure). Results are shown for an anti-infective vaccine **(A)** and a transmission blocking vaccine **(C)**. Given a defined health goal of reduction in $PfPR_{0\text{-}99}$ (minimum 70% reduction, dashed horizontal lines on figures **(A)** and **(C)**), the corresponding minimum product profile requirements are shown for the same two interventions, i.e., for an anti-infective vaccine in **(B)** and transmission blocking vaccine in **(D)**. Both figures **(B)** and **(D)** show how the minimum required profiles change when these interventions are delivered in combination with a blood-stage drug. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%. The descriptions of all intervention properties for identification of minimal profiles are detailed in Table S2.2.*

Discussion

In this study, we introduced a new modeling and machine-learning framework that for the first time enables quantitative differentiation between operational, setting, and intervention parameters as determinants of intervention impact, using detailed simulation models of disease. Our framework can be used for any disease where a valid model of disease progression or natural history of disease is available. We provided mathematical tools for efficiently and quantitatively defining the minimum profiles of malaria interventions as well as delivery approaches required to reach a desired health goal. Furthermore, our methodology provides a means to refine the identified optimal efficacy and duration characteristics as additional information becomes available. As a result, we can apply fully-detailed disease models to direct the design of novel interventions and understand how intervention-specific, epidemiological and systems factors jointly contribute to impact and thus inform TPP guidance. Most immediately, the approach is highly relevant to define successful interventions against emerging diseases such as SARS-CoV-2, and to support efficient, fast development of operational strategies. As uncertainties in disease progression and epidemiology can be incorporated in our approach, it also provides a way to systematically sort through large complex landscapes of unknowns and thus refine properties of interventions or clinical trials as more knowledge is available.

The value of our approach is realized through iterative collaboration with product development experts, to perform model-based guidance throughout the development process, and refine feedback on model predictions as interventions progress through development. For malaria, where multiple interventions are in development, it also offers an approach for product developers from diverse fields (such as therapeutics and insecticide development) to collaborate and incorporate knowledge of other interventions into their TPP development. Although in our analysis we used reduction of $Pf\mathrm{PR}_{0\text{-}99}$ as a health goal, our method can be applied to other disease burden statistics as required. The same rationale applies for investigation of other deployment strategies, required doses of interventions or further intervention combinations.

While also bringing valuable quantitative insights to guide product development, our analysis of novel malaria interventions reproduces previous findings concerning intervention characteristics which are key drivers of impact. Previous studies have shown that intervention coverage is a major determinant of impact in the context of mass drug administration (*17*), of vaccines (*35*) as

426 well as vector control (*36*). Furthermore, our analysis reaffirms previous work showing the
427 ability of vector control interventions to achieve substantial reductions in malaria burden (*37*).

428 Our approach constitutes a powerful tool to help address the challenges of current malaria
429 strategies and develop new interventions to progress towards malaria elimination. While
430 currently promising interventions such as insecticide treated nets, seasonal malaria
431 chemoprevention (SMC) and intermittent preventative treatment (IPT) have been very successful
432 at reducing malaria incidence and saving lives, their improved burden reduction and future
433 success is currently challenged by limited adherence, resource and time constraints to increase
434 coverage and usage in underserved populations, as well as resistance (*38*). Furthermore, for
435 settings where SMC has not been implemented or not recommended (for example in East Africa
436 or in perennial settings), there remains a gap in available interventions to protect vulnerable
437 populations who experience the highest burden of malaria. Similarly, for settings with outdoor
438 biting mosquitoes, the development and rollout of novel vector control interventions is needed.
439 New therapeutics and immune therapies suitable for seasonal delivery such as long-acting
440 injectables or monoclonal antibodies are currently being developed that may close one of these
441 gaps (*39, 40*). However, in order to efficiently make decisions on their development, guidance on
442 their key performance characteristics and definition of their TPP documents is needed from early
443 stages. Our quantitative framework can support the development of interventions from the
444 beginning by generating the evidence to inform and define evaluation criteria ensuring new
445 products meet relevant health targets, while considering how these products may affect disease
446 burden and epidemiology within a population. As we show here, this relies on iterative dialogue
447 with stakeholders, to first define health targets, simulated scenarios, achievable intervention
448 properties and operational settings. The modelling part of the framework incorporates all this
449 information as well as relevant disease transmission dynamics, building an *in-silico* system for
450 testing the developed intervention. Next, the sensitivity analysis part of the framework informs
451 which intervention characteristics drive impact and are thus crucial in achieving the defined
452 health goal, providing insights on the development processes to be prioritized. Finally, the
453 optimization analysis part of the framework reveals the potential of the developed intervention
454 and how its efficacy and coverage requirements change according to the defined health targets
455 and deployment setting. The landscapes of intervention profiles help product developers to gauge
456 development and investment efforts and select promising products. Furthermore, our approach

allows investigating combinations of new and existing interventions, identifying alternatives to alleviate shortcomings such as coverage limitations. To achieve a final TPP, several iterations of the analysis are required, to ensure that the optimal tradeoffs between intervention capabilities and target goals for a given setting are achieved.

As with all modelling studies, our approach is exposed to several limitations. The provided quantitative estimations in this study incorporate an increased level of uncertainly due to the additional emulation layer and are dependent on the performance of the trained emulator. We addressed this challenge with extensive adaptive sampling and testing to ensure a high level of accuracy of the trained emulators (Fig. 2, S4.1-S4.3, Table S4.1). Despite the intrinsic uncertainty, the framework is intended to provide guiding principles and an efficient means of exploring the space of intervention characteristics which otherwise would not be possible. Evidently, our analysis relies on the disease model assumptions of disease and transmission dynamics as well as expert opinion of likely intervention parameterizations in absence of clinical knowledge. Lastly, the current analysis explored a subset of use-cases, transmission settings and intervention combinations. Future work should focus on the likely settings and relevant use cases as the interventions are being developed and their TPP documents refined.

Moving beyond the work presented in this paper, our framework would allow combining simulation models with other sources of data describing geographical variation in disease, for example, modelled health systems or modelled prevalence (*41, 42*) and to incorporate interactions of interventions with novel interventions for surveillance. Clinical trials for new interventions could thereby be prioritized to geographical settings, where public health impact is likely to be maximized and where appropriate, to inform decisions on achieving non-inferiority or superiority endpoints (*43, 44*). A significant extension is incorporating economic considerations which may affect development decisions, including both costs of R&D, as well as implementation and systems costs for final deployment.

Materials and Methods

The approach introduced here combines infectious disease modeling with machine learning to understand determinants and define quantitative properties of target product profiles of new

486 malaria interventions. The building blocks and methodology of the approach are schematically
487 outlined in Fig. 2 which guides the following sections of Materials and Methods.

488 **1 Description of the disease model**

489 **1.1 Individual-based model of malaria transmission**

490 We used OpenMalaria (*33, 45*), an open source stochastic individual-based model to
491 simulate malaria epidemiology and transmission dynamics across humans and mosquitoes in
492 various settings. OpenMalaria considers the natural history of malaria in humans linked with a
493 deterministic, entomological model of the mosquito oviposition cycle and malaria transmission
494 in mosquitoes (*46, 47*) (Table S1.1). The modelled transmission cycle (Fig. 1) considers the
495 chain of processes following infection of a human host, simulating malaria infection in
496 individuals and modelling infection characteristics such as parasite density, duration of infection,
497 infectivity to mosquitoes, and health outcomes such as morbidity, mortality or anemia.
498 OpenMalaria specifically captures heterogeneity in host exposure, susceptibility and immune
499 response, taking into consideration the effects of several factors such as acquired immunity,
500 human demography structure, or seasonality (*48-51*). Furthermore, the model includes a detailed
501 representation of the health system (*52*), and of a wide range of human and vector control
502 interventions while tracking multiple health outcomes over time (Fig. 1, Table S1.1).

503 OpenMalaria has been widely documented and validated against a multitude of field
504 studies, compared to existing models and used in extensive studies to provide evidence for the
505 epidemiological effects of various interventions(*10, 29, 33, 53-55*). It comprises 14 model
506 variants based on distinct sets of assumptions on its epidemiology and transmission components
507 (*45*). For the present analysis, the "base" simulation model was used. The mathematical
508 equations of the model, its assumptions and calibrations have been thoroughly described in
509 numerous previous publications, therefore are not specified here, however, an overview of the
510 OpenMalaria modelled processes and assumptions along with the corresponding references are
511 provided in Table S1.1.

512

513 **1.2 Calibration of the disease model and description of simulation experiments**

514 OpenMalaria has been calibrated and validated in previous studies using historical
515 epidemiological data (*29, 33, 45*). The present analysis uses a previously-calibrated version of

516 the model which reflects demographics, epidemiology, entomology, health system and
517 seasonality of a health facility catchment area in Tanzania (*50, 52, 56*).
518 The simulated human population size in this analysis was 10'000 individuals, with its age
519 structure informed by data collected from a health and demographic surveillance site in Ifakara,
520 Tanzania, available through the INDEPTH network (*57*). For all simulations, we assumed there
521 were no imported infections during the whole study period.
522 Health system characteristics (Table S1.1) were defined through parameterization of a case
523 management model based on data provided by the Tanzanian National Malaria Control Program
524 (*52*). To define the simulated case management level, the probability of effective cure within two
525 weeks from the onset of fever (E_{14}) was varied within the interval [0 - 0.8] corresponding to a
526 probability of seeking care (access to treatment) within 5 days from the onset of fever (E_5) within
527 the interval [0.04 - 0.5] (*53*). During the model simulations, the case management level was
528 constant over time.
529 Mosquito entomological parameters and seasonal exposure patterns were estimated from
530 field studies conducted in the Namawala and Michenga villages located nearby Ifakara in
531 Tanzania (*58, 59*). Two archetypal seasonal settings were simulated: a seasonal exposure setting
532 with one transmission peak in September estimated from the mentioned field studies (Fig. S2.1),
533 and a perennial setting with uniform, constant exposure throughout the year. Two mosquito
534 species were present in the simulated settings: endophagic (indoor-biting, human blood index
535 equal to 0.99) and exophagic (outdoor-biting, human blood index is 0.5), respectively. The ratios
536 between the population sizes of indoor and outdoor mosquito species were classified into three
537 levels corresponding to high (indoor proportion is 0.8 out of total mosquito population), mid
538 (indoor proportion is 0.5) and low indoor biting (indoor proportion is 0.2). The extent of malaria
539 transmission in each simulation was defined by the yearly entomological inoculation rate (EIR).
540 For each simulation, EIR was sampled from the interval [1, 25] leading to a simulated range of
541 *Plasmodium falciparum* parasite rate or prevalence (*Pf*PR) distributions across the various
542 transmission settings (Figures S2.1, S2.2, Table S2.1).
543
544 **1.3 Definition of intervention profiles**
545 Adopting a holistic view, we built an agnostic, standardized representation for each
546 malaria intervention. Accordingly, a malaria intervention was characterized through the targets

547 of the transmission life cycle it affects, along with the efficacy, half-life and decay of its effect
548 (Figure 1, S2.3, Table S2.1). The efficacy of a therapeutic intervention was quantified by its
549 ability to clear parasites or prevent infection, while for mosquito-targeted interventions (vector
550 control tools) it corresponded to the ability of the intervention to kill or prevent mosquitoes from
551 biting human hosts. For each intervention, its efficacy decayed over time according to a specific
552 decay type (defined in Fig. S2.3). The coverage of interventions was quantified by the
553 percentage of the population affected by the respective intervention. Geographical setting
554 characteristics such as entomological inoculation rates (EIR), seasonality, case-management
555 coverage, as well as transmission and vector characteristics were also included in the simulation
556 specifications (Fig. 1, Table S2.1).

557 We defined the following intervention targets in the transmission cycle (Fig. 1):

558 - **Anti-infective:** acts at the liver stage and prevents occurrence of a new infection
559 - **Blood stage clearance:** clears blood-stage parasites by administration of a drug
560 - **Transmission blocking:** prevents parasite development into gametocytes
561 - **Mosquito life-cycle killing effect:** kills mosquitoes during different stages of their life cycle,
562 such as, for example, before a blood meal (pre-prandial killing) and/or after a blood meal
563 (post-prandial killing). Furthermore, mosquitoes are affected by vector control interventions
564 according to their indoor and outdoor biting patterns.

565 The length of the intervention effect was described via either half-life for exponential,
566 sigmoidal or biphasic decay profiles, or by duration for step-like decay profiles. Generally, half-
567 life refers to half-life of intervention efficacy decay, representing the time in which the initial
568 intervention efficacy has been reduced by 50% (Fig. S2.3, Table S2.1). As opposed to half-life,
569 the duration of effect is equivalent to the entire decay time. For simplicity, since only one
570 intervention had a step-like decay, we use the words half-life and duration interchangeably.

571 To define the breadth, range and profiles of simulated malaria interventions, we
572 collaborated with end users at the Bill & Melinda Gates Foundation and the product development
573 partnerships PATH's Malaria Vaccine Initiative (PATH-MVI) and Innovative Vector Control
574 Consortium (IVCC). For each new intervention in the portfolio of PATH-MVI, IVCC and others,
575 we undertook several expert discussion groups to catalogue the ranges of potential effectiveness;
576 potential delivery strategies; parasite or vector targets; the likely properties in terms of action
577 (target), efficacy, duration and decay; and use cases/delivery (age target, mass intervention,

578 yearly deployment or other). These results are summarized in Table S2.1 which presents a
579 comprehensive description of all intervention characteristics, parameter values, as well as the
580 ranges they were varied within. Setting-specific characteristics used for the different simulated
581 scenarios are also summarized in Table S2.1.

582 In our current study, each intervention or combination of interventions was applied as
583 mass intervention targeting all ages equally, along with continuous case management. In this
584 analysis, we did not examine targeting particular populations or age groups to develop our
585 approach. The deployed mass intervention packages followed a long period of model warm up
586 (150 years), and were implemented in June and/or December for three years (Fig. S3.1).
587 Coverage at deployment time refers to the percentage of the population covered by the
588 intervention's initial efficacy, irrespective of how many doses/applications are required to reach
589 that coverage, assuming that the necessary doses have previously occurred.

590

591 **1.4 Translation of input EIR to *Pf*PR$_{2\text{-}10}$ and *Pf*PR$_{0\text{-}99}$**

592 For each simulation, OpenMalaria requires the definition of the intensity and seasonality
593 of malaria exposure specified through the input EIR level and its yearly profile in the absence of
594 interventions (Fig. S2.1). EIR is an appropriate measure for reflecting transmission intensity
595 (*60*), however it is difficult to measure in the field and its interpretation in the context of
596 intervention impact is difficult to apprehend when looking at the effects of drugs and vaccines
597 (*61, 62*). For this reason, although EIR is the force of infection input to all OpenMalaria
598 simulations, we report simulation outcomes and downstream analyses at the corresponding
599 median *Pf*PR$_{2\text{-}10}$ and *Pf*PR$_{0\text{-}99}$ before the interventions are deployed. We report true infection
600 prevalence and not patent PCR or RDT-detected. To do so, we discretized the continuous EIR
601 space into discrete unit-wide intervals and the median *Pf*PR was calculated across the obtained
602 *Pf*PR for all simulations in each discrete interval (Fig. S2.2).

603

604 **1.5 Definition of impact and health goals**

605 A comprehensive set of simulated scenarios was built by sampling uniformly the
606 parameter space of setting and intervention characteristics. To estimate the impact of the
607 deployed interventions, in each simulation, we calculated the reduction in *Pf*PR$_{0\text{-}99}$ attributable to
608 the deployed intervention. *Pf*PR$_{0\text{-}99}$ reduction was calculated by comparing the initial average

prevalence in the year before any interventions were deployed to the average yearly prevalence obtained in the first year (short follow-up) and in the third year (long follow-up) after deployment of interventions (Fig. S3.1). Consequently, the defined health goals corresponded to a given minimum threshold of *Pf*PR$_{0\text{-}99}$ reduction that the deployed interventions should achieve.

Figures S3.2 – S3.4 present the distributions of obtained *Pf*PR$_{0\text{-}99}$ reduction for the OpenMalaria simulation experiments covering all the interventions and deployments investigated in the present study. In seasonal, low transmission settings (EIR < 2) a proportion of simulations reached elimination before any intervention was deployed and were removed from the analysis (Fig. S3.5). Since this happened for over 75% of simulations at EIR<2, we did not investigate optimal intervention profiles for transmission settings with EIR<2. Arguably, for these settings close to elimination, a different health goal, such as probability of elimination, would be more appropriate which is not within the scope of the present study focusing on *Pf*PR$_{0\text{-}99}$ reduction.

2 Building a disease model emulator with Gaussian processes

As it was computationally intensive and challenging to run an exhaustive number of simulations in order to explore with OpenMalaria all the parameter space for diverse combinations of interventions, settings and deployments, we applied machine learning techniques and kernel methods to leverage our analysis. Precisely, starting from a training dataset of simulations generated with OpenMalaria, we used Gaussian process (GP) models (*24*) to infer the relationship between simulation variables (e.g., intervention coverage, half-life, efficacy, etc.) and corresponding intervention impact (*Pf*PR$_{0\text{-}99}$ reduction). This approach allowed us to build a fast, simplified predictive model that could provide estimates of the disease model output for any new inputs without running new OpenMalaria model simulations.

Gaussian process models are non-parametric models which define a prior probability distribution over a collection of functions using a kernel, smooth function. Precisely, given the relationship

$$y = f(\boldsymbol{x}) + \varepsilon$$

where y in our case is the *Pf*PR$_{0\text{-}99}$ reduction and $\boldsymbol{x}$ represents the set of intervention parameters $x_1, \ldots, x_n$, the main assumption of a GP is that

$$P\big(f(x_1), f(x_2), \ldots, f(x_n)\big) \sim N(\mu, \Sigma)$$

where

$$\Sigma_{x_i,x_j} = K(x_i, x_j)$$

is the covariance matrix of the Gaussian distribution, μ is its mean and K is a kernel function (*24*). Once data is observed, the posterior probability distribution of the functions consistent with the observed data can be derived which is then used to infer outcomes at unobserved locations in the parameter space (*24*). The intuition behind a GP model is based on the "smoothness" relationship between its components. Accordingly, points which are close in the input parameter space will lead to close points in the output space.

2.1 Training data

For each intervention and setting, a training dataset was built using discrete Latin hypercube uniform sampling (*63*) across the input parameter space (defined in Table S2.1). Ten stochastic realizations (replicates) of each sampled data point were considered. OpenMalaria was run on the sampled data points and *Pf*$PR_{0\text{-}99}$ was calculated for both short and long follow-up. The size of the training set was varied between 10 and 1000 points (100 – 10000 including replicates) for several simulation experiments (Fig. S4.1) and the performance of the trained GP was assessed via the Pearson correlation coefficient r^2. The minimum training set size which led to $r^2 > 0.95$ was selected for the remaining simulation experiments.

2.2 Gaussian process emulators

For each transmission setting and intervention, a GP model with a Gaussian kernel was trained in a 5-fold cross-validation scheme using the training set with OpenMalaria simulations. For training the GP, we used the R package *HetGP* version 1.1.1 (*64, 65*). HetGP is a powerful implementation of GP models, featuring heteroskedastic GP modeling embedded in a fast and efficient maximum-likelihood-based inference scheme.

GP performance was assessed by calculating the correlation between true and predicted outputs on out-of-sample test sets as well as the mean squared error (Fig. S4.2 – 4.3, Table S4.1). Precisely, the training set was split in 5 subsets and, iteratively, 4 of these subsets were used for training the GP, while the remaining one was used as an out-of-sample test set during the cross-validation procedure. After assessing the prediction error during the cross-validation procedure, the GP was trained using the entire training set. Furthermore, since the trained GP model provides the mean and variance of each predicted output, we used this probabilistic

representation to assess the uncertainty of the trained model across the entire parameter space and to refine the GP model through adaptive sampling (*66-68*). Accordingly, we iteratively sampled new training points from high-uncertainty regions of the parameter space and updated the model with the new training samples until the correlation between true and predicted values on an out-of-sample test set reached a plateau. Finally, a separate out-of-sample test set was built to assess the overall performance of the GP (Fig. S4.2 – 4.3, Table S4.1).

3 Identifying impact determinants through sensitivity analysis

In order to estimate the contribution of each model input and its interactions with the other inputs to the variance of the model outcome, we conducted a global sensitivity analysis based on variance decomposition (*69*). This analysis shows which input parameters have higher impact on the model outcome. It relies on the decomposition of the output variance in a sum of individual input parameter conditional variances:

$$\mathrm{Var(Y)} = \sum_{i} \mathrm{V_i} + \sum_{i} \sum_{j>i} \mathrm{V_{ij}} + \cdots + \mathrm{V_{12\ldots d}}$$

where Y is the model outcome (in our case, *Pf*$PR_{0\text{-}99}$ reduction), d corresponds to the number of model inputs, and the conditional variances are defined as:

$$\mathrm{V_i} = \mathrm{Var(E(Y|x_i))}$$

$$\mathrm{V_{ij}} = \mathrm{Var}\left(\mathrm{E}\left(\mathrm{Y}\middle|\mathrm{x_i}, \mathrm{x_j}\right)\right) - \mathrm{V_i} - \mathrm{V_j}$$

$$\mathrm{V_{ijk}} = \mathrm{Var}\left(\mathrm{E}\left(\mathrm{Y}\middle|\mathrm{x_i}, \mathrm{x_j}, \mathrm{x_k}\right)\right) - \mathrm{V_{ij}} - \mathrm{V_{jk}} - \mathrm{V_{ik}} - \mathrm{V_i} - \mathrm{V_j} - \mathrm{V_k}$$

$$\ldots$$

with $x_1, \ldots, x_n$ representing the model input parameters.

Based on the above decomposition of output variance, the first order sensitivity index is defined as:

$$\mathrm{S_i} = \frac{\mathrm{V_i}}{\mathrm{Var(Y)}}$$

and corresponds to the proportion of output variance assigned to the main effect of X_i, i.e., regardless its interactions with other model inputs (*69, 70*).

To account for the contribution of each model input as well as the variance of its interactions with other inputs to the variability of the model output, the total effect sensitivity index is used:

$$T_i = 1 - \frac{\mathrm{Var}(\mathrm{E}(Y|x_{\sim i}))}{\mathrm{Var}(Y)}$$

where the notation ~i stands for all indices except i (*69, 70*).

In the above decomposition of model output variance, by replacing the expressions of the sensitivity indexes, the following properties can be deduced:

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \cdots + S_{12\ldots d} = 1$$

and

$$\sum_i T_i \geq 1 .$$

To compute the sensitivity indexes, we use the function "soboljansen" from the R package "sensitivity" (*71*). The function estimates the sensitivity indices through MCMC sampling, using a Monte Carlo approximation for computing conditional expectations. Within the sampling scheme, we sampled 100'000 points to estimate the sensitivity indices.

Calculating the sensitivity indices defined above, the variance of the GP emulator output was thus decomposed into proportions attributable to intervention characteristics, i.e., intervention efficacy, half-life and deployment coverage, as well as access to care. Using the main effects, we defined the relative importance r_i of each characteristic as a proxy for impact determinants as follows:

$$r_i = \frac{S_i}{\sum_{i=1}^{d} Si}$$

where d is the number of intervention characteristics and $\sum_{i=1}^{d} r_i = 1$.

4 Finding minimal intervention properties

The trained GP models for each transmission setting and intervention were used within a general-purpose optimization scheme in order to identify minimum intervention properties that reach a defined *Pf*PR$_{0\text{-}99}$ reduction goal given operational and intervention constraints.

Let

$$g(\boldsymbol{x}) = g(x_1, x_2, x_3, x_4)$$

denote the GP model predicting the mean prevalence reduction obtained after deploying an intervention with given characteristics in a transmission setting, with

$$x_1 = tool\ coverage$$

$$x_2 = tool\ half - life$$

$$x_3 = tool\ efficacy$$

$$x_4 = access\ to\ treatment.$$

For various levels of $Pf\mathrm{PR}_{0\text{-}99}$ denoted with p_k, each intervention characteristic was optimized separately, keeping the remaining characteristics as well as the level of case management fixed to pre-set levels. Precisely, the optimization procedure searches for

$$\min(x_i) |\ x_{\sim i}$$

such as

$$g(\boldsymbol{x}) \geq p_k$$

with the constraints:

$$l_i \leq\ x_i \leq\ u_i,$$

where l_i and u_i are the lower and upper bounds of x_i, respectively and the notation $\sim i$ is used to represent all the characteristics except i. A detailed description of the parameter specifications during optimization for each intervention is provided in Table S2.2.

To solve the above optimization problem, we used a general nonlinear augmented Lagrange multiplier method (*72, 73*) implemented in the R package "*Rsolnp*" (*74*). To ensure optimality of the obtained solutions and avoid local minima, 10 random restarts were chosen among 1000 uniformly-sampled input parameter sets and the optimization procedure was run separately for each restart (implemented in function "*gosolnp*" in the same R package). To capture the variance of the optimal intervention profile, since the output of a GP model is a distribution, we solved the above optimization problem for several cases and we report the distribution of the obtained minima when:

$$(i)\ \ g(\boldsymbol{x}) =\ \mu$$

$$(ii)\ \ g(\boldsymbol{x}) =\ \mu \pm\ \sigma$$

$$(iii)\ g(\boldsymbol{x}) =\ \mu \pm\ 2\sigma$$

where μ is the predicted mean of the GP model and σ is the standard deviation. Where the nonlinear optimization algorithm did not find any solutions, we performed an additional fine grid search of 10000 uniformly-sampled data points.

In seasonal settings, at low transmission (simulated EIR < 2, corresponding simulated true $Pf\mathrm{PR}_{2\text{-}10}$ < 11.7%), over 75% of simulations reached malaria elimination ($Pf\mathrm{PR}_{0\text{-}99}$ = 0) under the simulated levels of case management, before intervention deployment (Fig. S3.5). For this reason, the space of obtained prevalence reductions following intervention deployment was

759 rather sparse and the obtained optima were not reliable and often did not converge. Therefore,
760 we chose to report minimum intervention profiles for settings with true *Pf*PR_{2-10} >= 11.7% (with
761 RDTs this yields a patent *Pf*PR_{2-10} >= 5.8%).

762

763 **5 Iterative communication with stakeholders**

764 During the development of our methodological framework, we actively engaged in regular
765 communication and exchanges with different expert groups. The stakeholders involved in these
766 discussions were the Bill and Melinda Gates Foundation (BMGF), the Innovative Vector Control
767 Consortium (IVCC) and the PATH's Malaria Vaccine Initiative (PATH-MVI). Coordinated by
768 BMGF, these exchanges ensured a crucial discussion environment, aiding and guiding the
769 methodology at various levels: intervention profiling, and defining relevant intervention use
770 cases, and product characteristics. Furthermore, the framework has been presented and validated
771 in presence of the stakeholders in successive meetings. These discussions contributed towards
772 refining the investigation of various intervention profiles and led to exploration of intervention
773 combinations. Subsequently, the iterative exchanges with the stakeholders have not only shaped
774 but also proven the value of our methodological framework in its versatility to adapt addressing
775 relevant questions along the product development pathway.

776

777 **Acknowledgments:**

778 We would like to thank Lydia Burgert, Theresa Reiker, Andrew Shattock, Thomas Smith, and
779 Dylan Muir for insightful discussions and feedback on the developed methodology and the
780 manuscript. We would also like to thank Thomas van Boeckel and Amalio Telenti for providing
781 useful feedback on the manuscript. Calculations were performed at sciCORE
782 (http://scicore.unibas.ch/) scientific computing core facility at University of Basel. We would
783 further like to thank collaborators at the Innovative Vector Control Consortium (IVCC), PATH's
784 Malaria Vaccine Initiative (MVI) and Joerg Moehrle from Medicines for Malaria Venture
785 (MMV) for their insightful discussions and feedback on the model scenarios. This work has been
786 possible thanks to the Malaria Team at the Bill and Melinda Gates Foundation who facilitated
787 exchanges with the product development partners and supported model scenarios and
788 interpretations. In particular, we would like to thank Scott Miller, Jean-Luc Bodmer, Laura
789 Norris, Bruno Moonen, Dan Strickman, and Philip Welkhoff.

Funding: This work was supported by the Bill and Melinda Gates Foundation (OPP1170505 to MAP) and the Swiss National Science Foundation (PP00P3_170702 to MAP);

Author contributions: M.G., G.Y. and M.A.P conceived the study, designed the simulation experiments, developed methodology and analyzed the results. F.C., E.C., and N.C. provided methodological expertise. E.M.S., N.H., M.M., and S.R contributed with their expertise regarding product development, intervention properties and guided analysis. M.G., G.Y., M.A.P. wrote the manuscript. All authors provided continuous feedback and approved the final manuscript;

Competing interests: All authors declare no competing interests;

Data and materials availability: All the analysis code used in the paper as well as corresponding documentation, parameterizations and configuration files for the software workflow necessary to generate the simulation data with OpenMalaria and reproduce the analysis are available at https://github.com/SwissTPH/TPP_workflow.

Supplementary Materials are below following the references

.

List of Supplementary Figures and Tables

The following supplementary Figures and Tables complement the analysis and results reported in the main manuscript and are organized as follows:

- Description of the malaria disease transmission dynamics (OpenMalaria) model components and assumptions:
  - **Table S1.1.** Overview of the OpenMalaria model components.
- Simulated malaria transmission dynamics in the presented analysis:
  - **Fig. S2.1.** Illustration of the yearly malaria transmission and prevalence patterns in simulated seasonal settings.
  - **Fig. S2.2.** Simulated distributions of true and patent (detected with PCR or RDT) $Pf\mathrm{PR}_{0\text{-}99}$ and $Pf\mathrm{PR}_{2\text{-}10}$ for various input EIR levels in absence of interventions.
- Parameterizations of simulated malaria interventions and their optimization setup:
  - **Fig. S2.3.** Representation of decay and the range of efficacy and half-life against different parasite or vector targets for intervention-agnostic malaria interventions.
  - **Table S2.1.** Description and ranges of simulation variables.
  - **Table S2.2.** Specifications of the optimization procedure for TPP development.
- Simulation outputs:

- **Fig. S3.1.** Examples of OpenMalaria simulation outputs.
- **Fig. S3.2.** Distributions of prevalence reduction following yearly deployment of single interventions.
- **Fig. S3.3.** Distributions of prevalence reduction following yearly deployment of combinations of interventions.
- **Fig. S3.4.** Distributions of prevalence reduction following deployment of single and combinations of interventions twice per year.
- **Fig. S3.5.** Simulations reaching malaria elimination before intervention deployment.

- Emulator training and evaluation:
  - **Fig. S4.1.** Assessment of the performance of the trained GP depending on the training set size.
  - **Fig. S4.2.** Performance of the trained GP emulators predicting immediate intervention impact.
  - **Fig. S4.3.** Performance of the trained GP emulators predicting long-term intervention impact.
  - **Table S4.1.** Performance of the trained Gaussian Process emulators predicting immediate and long-term intervention impact.
- Summary of analysis results for all simulated transmission settings and interventions:
  - **Table S5.1.** Key findings guiding target product profiles of new malaria interventions.
- Sensitivity analyses and impact determinants of interventions:
  - **Fig. S6.1.** Key drivers of impact for therapeutic malaria interventions across different transmission settings.
  - **Fig. S6.2.** Key drivers of impact for vector control malaria interventions across different transmission settings.
- Feasible landscapes of optimal, constrained intervention profiles (TPPs) for achieving a desired health goal across different transmission settings and operational factors:
  - **Fig. S7.1** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed once per year.
  - **Fig. S7.2.** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed twice per year.
  - **Fig. S7.3.** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective vaccine deployed once per year.
  - **Fig. S7.4.** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for a transmission-blocking vaccine deployed once per year.
  - **Fig. S7.5.** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for attractive targeted sugar baits deployed once or twice per year.
  - **Fig. S7.6.** Feasible landscapes of optimal, constrained intervention profiles (TPPs) for eave tubes deployed once per year.
- Minimum profiles of interventions for achieving a desired health goal across different transmission settings and operational factors:
  - **Fig. S8.1.** Optimal intervention profiles (TPPs) for anti-infective monoclonal antibodies under various deployment regimes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%.

- **Fig. S8.2.** Optimal intervention profiles (TPPs) for anti-infective vaccines under various deployment regimes to achieve a *Pf*PR_{0-99} reduction of at least 70%.
- **Fig. S8.3.** Optimal intervention profiles (TPPs) for transmission-blocking vaccines under various deployment regimes to achieve a *Pf*PR_{0-99} reduction of at least 70%.
- **Fig. S8.4.** Optimal intervention profiles (TPPs) for attractive targeted sugar baits under various deployment regimes to achieve a *Pf*PR_{0-99} reduction of at least 70%.
- **Fig. S8.5.** Optimal intervention profiles (TPPs) for eave tubes to achieve a *Pf*PR_{0-99} reduction of at least 70%.

References and Notes:

1. P. Cocco, A. Ayaz-Shah, M. P. Messenger, R. M. West, B. Shinkins, Target Product Profiles for medical tests: a systematic review of current methods. *BMC medicine* **18**, 1-12 (2020).
2. A. Brooks, J. K. Nunes, A. Garnett, R. Biellik, D. Leboulleux, A. J. Birkett, C. Loucq, Aligning new interventions with developing country health systems: Target product profiles, presentation, and clinical trial design. *Global Public Health* **7**, 931-945 (2012).
3. R. Chin, B. Y. Lee, *Principles and practice of clinical trial medicine*. (Elsevier, 2008).
4. J. N. Burrows, S. Duparc, W. E. Gutteridge, R. H. van Huijsduijnen, W. Kaszubska, F. Macintyre, S. Mazzuri, J. J. Möhrle, T. N. Wells, New developments in anti-malarial target candidate and product profiles. *Malaria journal* **16**, 26 (2017).
5. U. F. a. D. Administration., Guidance for industry and review staff Target Product Profile — a strategic development process tool. available at https://www.fda.gov/media/102657/download. (2007).
6. B. Y. Lee, D. S. Burke, Constructing target product profiles (TPPs) to help vaccines overcome post-approval obstacles. *Vaccine* **28**, 2806-2809 (2010).
7. K. B. Ebels, C. Clerk, C. H. Crudder, S. McGray, K. Magnuson, K. Tietje, P. LaBarre, in *IEEE Global Humanitarian Technology Conference (GHTC 2014)*. (IEEE, 2014), pp. 555-560.
8. H. Heesterbeek, R. M. Anderson, V. Andreasen, S. Bansal, D. De Angelis, C. Dye, K. T. Eames, W. J. Edmunds, S. D. Frost, S. Funk, Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, (2015).
9. R. M. Anderson, B. Anderson, R. M. May, *Infectious diseases of humans: dynamics and control*. (Oxford university press, 1991).
10. M. A. Penny, R. Verity, C. A. Bever, C. Sauboin, K. Galactionova, S. Flasche, M. T. White, E. A. Wenger, N. Van de Velde, P. Pemberton-Ross, Public health impact and cost-effectiveness of the RTS, S/AS01 malaria vaccine: a systematic comparison of predictions from four mathematical models. *The Lancet* **387**, 367-375 (2016).
11. H. C. Slater, A. Ross, A. L. Ouédraogo, L. J. White, C. Nguon, P. G. Walker, P. Ngor, R. Aguas, S. P. Silal, A. M. Dondorp, Assessing the impact of next-generation rapid diagnostic tests on Plasmodium falciparum malaria elimination strategies. *Nature* **528**, S94-S101 (2015).
12. J. Gerardin, C. A. Bever, B. Hamainza, J. M. Miller, P. A. Eckhoff, E. A. Wenger, Optimal Population-Level Infection Detection Strategies for Malaria Control and

Elimination in a Spatial Model of Malaria Transmission. *PLOS Computational Biology* **12**, e1004707 (2016).

13. M. Walker, J. I. Hamley, P. Milton, F. Monnot, B. Pedrique, M.-G. Basáñez, Designing antifilarial drug trials using clinical trial simulators. *Nature Communications* **11**, 1-11 (2020).
14. H. Kimko, J. Pinheiro, Model-based clinical drug development in the past, present and future: a commentary. *Br. J. Clin. Pharmacol.* **79**, 108-116 (2015).
15. C. Vegvari, E. Cauet, C. Hadjichrysanthou, E. Lawrence, G.-J. Weverling, F. De Wolf, R. M. Anderson, Using clinical trial simulators to analyse the sources of variance in clinical trials of novel therapies for acute viral infections. *PloS one* **11**, (2016).
16. P. Selvaraj, J. Suresh, E. A. Wenger, C. A. Bever, J. Gerardin, Reducing malaria burden and accelerating elimination with long-lasting systemic insecticides: a modelling study of three potential use cases. *Malaria journal* **18**, 307-307 (2019).
17. O. J. Brady, H. C. Slater, P. Pemberton-Ross, E. Wenger, R. J. Maude, A. C. Ghani, M. A. Penny, J. Gerardin, L. J. White, N. Chitnis, Role of mass drug administration in elimination of Plasmodium falciparum malaria: a consensus modelling study. *The Lancet Global Health* **5**, e680-e687 (2017).
18. M. Kretzschmar, Disease modeling for public health: added value, challenges, and institutional constraints. *Journal of public health policy* **41**, 39-51 (2020).
19. K. Panayidou, S. Gsteiger, M. Egger, G. Kilcher, M. Carreras, O. Efthimiou, T. P. Debray, S. Trelle, N. Hummel, G. m. r. group, GetReal in mathematical modelling: a review of studies predicting drug effectiveness in the real world. *Research synthesis methods* **7**, 264-277 (2016).
20. L. Rychetnik, M. Frommer, P. Hawe, A. Shiell, Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health* **56**, 119 (2002).
21. A. B. Hogan, P. Winskill, R. Verity, J. T. Griffin, A. C. Ghani, Modelling population-level impact to inform target product profiles for childhood malaria vaccines. *BMC medicine* **16**, 1-11 (2018).
22. G. F. Killeen, N. Chitnis, S. J. Moore, F. O. Okumu, Target product profile choices for intra-domiciliary malaria vector control pesticide products: repel or kill? *Malaria journal* **10**, 207 (2011).
23. W. Pan-Ngum, T. Kinyanjui, M. Kiti, S. Taylor, J.-F. Toussaint, S. Saralamba, T. Van Effelterre, D. J. Nokes, L. J. White, Predicting the relative impacts of maternal and neonatal respiratory syncytial virus (RSV) vaccine target product profiles: A consensus modelling approach. *Vaccine* **35**, 403-409 (2017).
24. C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. (The MIT Press, 2005).
25. D. Ríos Insua, F. Ruggeri, M. P. Wiper, *Bayesian analysis of stochastic process models*. (Wiley, Chichester, 2012).
26. WHO, World Health Organization: Global report on insecticide resistance in malaria vectors: 2010–2016 accessed at https://www.who.int/malaria/areas/vector_control/insecticide_resistance/en/. (2018).
27. L. J. Espinoza, Malaria Resurgence in the Americas: An Underestimated Threat. *Pathogens* **8**, (2019).
28. C. J. M. Whitty, E. Ansah, Malaria control stalls in high incidence areas. *BMJ* **365**, l2216 (2019).

29. T. Smith, N. Maire, A. Ross, M. Penny, N. Chitnis, A. Schapira, A. Studer, B. Genton, C. Lengeler, F. Tediosi, D. De Savigny, M. Tanner, Towards a comprehensive simulation model of malaria epidemiology and control. *Parasitology* **135**, 1507-1516 (2008).
30. P. Winskill, P. G. T. Walker, J. T. Griffin, A. C. Ghani, Modelling the cost-effectiveness of introducing the RTS,S malaria vaccine relative to scaling up other malaria interventions in sub-Saharan Africa. *BMJ Global Health* **2**, e000090 (2017).
31. P. Winskill, P. G. Walker, R. E. Cibulskis, A. C. Ghani, Prioritizing the scale-up of interventions for malaria control and elimination. *Malaria Journal* **18**, 122 (2019).
32. E. Korenromp, G. Mahiané, M. Hamilton, C. Pretorius, R. Cibulskis, J. Lauer, T. A. Smith, O. J. T. Briët, Malaria intervention scale-up in Africa: effectiveness predictions for health programme planning tools, based on dynamic transmission modelling. *Malaria Journal* **15**, 417 (2016).
33. T. Smith, G. F. Killeen, N. Maire, A. Ross, L. Molineaux, F. Tediosi, G. Hutton, J. Utzinger, K. Dietz, M. Tanner, Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of Plasmodium falciparum malaria: overview. *The American Journal of Tropical Medicine and Hygiene* **75**, 1-10 (2006).
34. E. Cameron, K. E. Battle, S. Bhatt, D. J. Weiss, D. Bisanzio, B. Mappin, U. Dalrymple, S. I. Hay, D. L. Smith, J. T. Griffin, E. A. Wenger, P. A. Eckhoff, T. A. Smith, M. A. Penny, P. W. Gething, Defining the relationship between infection prevalence and clinical incidence of Plasmodium falciparum malaria. *Nature Communications* **6**, 8170 (2015).
35. C. E. Utazi, J. Thorley, V. A. Alegana, M. J. Ferrari, S. Takahashi, C. J. E. Metcalf, J. Lessler, F. T. Cutts, A. J. Tatem, Mapping vaccination coverage to explore the effects of delivery mechanisms and inform vaccination strategies. *Nature communications* **10**, 1-10 (2019).
36. A. L. Wilson, O. Courtenay, L. A. Kelly-Hope, T. W. Scott, W. Takken, S. J. Torr, S. W. Lindsay, The importance of vector control for the control and elimination of vector-borne diseases. *PLoS neglected tropical diseases* **14**, e0007831 (2020).
37. N. M. Ferguson, Challenges and opportunities in controlling mosquito-borne infections. *Nature* **559**, 490-497 (2018).
38. J. Gutman, S. Kovacs, G. Dorsey, A. Stergachis, F. O. ter Kuile, Safety, tolerability, and efficacy of repeated doses of dihydroartemisinin-piperaquine for prevention and treatment of malaria: a systematic review and meta-analysis. *The Lancet infectious diseases* **17**, 184-193 (2017).
39. F. Macintyre, H. Ramachandruni, J. N. Burrows, R. Holm, A. Thomas, J. J. Möhrle, S. Duparc, R. H. van Huijsduijnen, B. Greenwood, W. E. Gutteridge, Injectable anti-malarials revisited: discovery and development of new agents to protect against malaria. *Malaria journal* **17**, 402 (2018).
40. L. T. Wang, L. S. Pereira, Y. Flores-Garcia, J. O'Connor, B. J. Flynn, A. Schön, N. K. Hurlburt, M. Dillon, A. S. Yang, A. Fabra-García, A Potent Anti-Malarial Human Monoclonal Antibody Targets Circumsporozoite Protein Minor Repeats and Neutralizes Sporozoites in the Liver. *Immunity* **53**, 733-744. e738 (2020).
41. D. J. Weiss, T. C. Lucas, M. Nguyen, A. K. Nandi, D. Bisanzio, K. E. Battle, E. Cameron, K. A. Twohig, D. A. Pfeffer, J. A. Rozier, Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study. *The Lancet* **394**, 322-331 (2019).

42. M. Sahu, F. Tediosi, A. M. Noor, J. J. Aponte, G. Fink, Health systems and global progress towards malaria elimination, 2000–2016. *Malaria journal* **19**, 1-12 (2020).
43. R. Ristl, S. Urach, G. Rosenkranz, M. Posch, Methods for the analysis of multiple endpoints in small populations: A review. *Journal of biopharmaceutical statistics* **29**, 1-29 (2019).
44. Food, D. Administration, Multiple endpoints in clinical trials: guidance for industry. *College Park, Maryland: Food and Drug Administration*, (2017).
45. T. Smith, A. Ross, N. Maire, N. Chitnis, A. Studer, D. Hardy, A. Brooks, M. Penny, M. Tanner, Ensemble Modeling of the Likely Public Health Impact of a Pre-Erythrocytic Malaria Vaccine. *PLOS Medicine* **9**, e1001157 (2012).
46. N. Chitnis, T. Smith, R. Steketee, A mathematical model for the dynamics of malaria in mosquitoes feeding on a heterogeneous host population. *Journal of Biological Dynamics* **2**, 259-285 (2008).
47. N. Chitnis, D. Hardy, T. Smith, A Periodically-Forced Mathematical Model for the Seasonal Dynamics of Malaria in Mosquitoes. *Bulletin of Mathematical Biology* **74**, 1098-1124 (2012).
48. N. Maire, T. Smith, A. Ross, S. Owusu-Agyei, K. Dietz, L. Molineaux, A model for natural immunity to asexual blood stages of Plasmodium falciparum malaria in endemic areas. *The American journal of tropical medicine and hygiene* **75**, 19-31 (2006).
49. A. Ross, G. Killeen, T. Smith, Relationships between host infectivity to mosquitoes and asexual parasite density in Plasmodium falciparum. *The American journal of tropical medicine and hygiene* **75**, 32-37 (2006).
50. A. Ross, N. Maire, L. Molineaux, T. Smith, An epidemiologic model of severe morbidity and mortality caused by Plasmodium falciparum. *The American journal of tropical medicine and hygiene* **75**, 63-73 (2006).
51. T. Smith, A. Ross, N. Maire, C. Rogier, J.-F. Trape, L. Molineaux, An epidemiologic model of the incidence of acute illness in Plasmodium falciparum malaria. *The American journal of tropical medicine and hygiene* **75**, 56-62 (2006).
52. F. Tediosi, N. Maire, T. Smith, G. Hutton, J. Utzinger, A. Ross, M. Tanner, An approach to model the costs and effects of case management of plasmodium falciparum malaria in sub-saharan Africa. *The American Journal of Tropical Medicine and Hygiene* **75**, 90-103 (2006).
53. M. A. Penny, N. Maire, C. A. Bever, P. Pemberton-Ross, O. J. T. Briët, D. L. Smith, P. W. Gething, T. A. Smith, Distribution of malaria exposure in endemic countries in Africa considering country levels of effective treatment. *Malaria Journal* **14**, 384 (2015).
54. N. R. Smith, J. M. Trauer, M. Gambhir, J. S. Richards, R. J. Maude, J. M. Keith, J. A. Flegg, Agent-based models of malaria transmission: a systematic review. *Malaria journal* **17**, 1-16 (2018).
55. T. Reiker, N. Chitnis, T. A. Smith, Modelling reactive case detection strategies for interrupting transmission of Plasmodium falciparum malaria. *Malaria Journal* **18**, 259 (2019).
56. N. Maire, F. Tediosi, A. Ross, T. Smith, Predictions of the epidemiologic impact of introducing a pre-erythrocytic vaccine into the expanded program on immunization in sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene* **75**, 111-118 (2006).

57. A. M. Ekström, J. Clark, P. Byass, A. Lopez, D. De Savigny, C. A. Moyer, H. Campbell, A. J. Gage, P. Bocquier, C. AbouZahr, O. Sankoh, INDEPTH Network: contributing to the data revolution. *The Lancet Diabetes & Endocrinology* **4**, 97 (2016).
58. G. F. Kileen, A. Ross, T. Smith, Infectiousness of malaria-endemic human populations to vectors. *The American Journal of Tropical Medicine and Hygiene* **75**, 38-45 (2006).
59. T. C. Smith, J. D. D. Charlwood, J. Kihonda, S. Mwankusye, P. F. Billingsley, J. Meuwissen, E. B. Lyimo, W. Takken, T. Teuscher, M. Tanner, Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission. *Acta tropica* **54 1**, 55-72 (1993).
60. C. Drakeley, D. Schellenberg, J. Kihonda, C. Sousa, A. Arez, D. Lopes, J. Lines, H. Mshinda, C. Lengeler, J. A. Schellenberg, An estimation of the entomological inoculation rate for Ifakara: a semi-urban area in a region of intense malaria transmission in Tanzania. *Tropical Medicine & International Health* **8**, 767-774 (2003).
61. A. M. Shaukat, J. G. Breman, F. E. McKenzie, Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. *Malaria journal* **9**, 122 (2010).
62. L. S. Tusting, T. Bousema, D. L. Smith, C. Drakeley, in *Advances in parasitology*. (Elsevier, 2014), vol. 84, pp. 151-208.
63. M. Stein, Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics* **29**, 143-151 (1987).
64. M. Binois, R. B. Gramacy, M. Ludkovski, Practical heteroscedastic gaussian process modeling for large simulation experiments. *Journal of Computational and Graphical Statistics* **27**, 808-821 (2018).
65. M. Binois, R. Gramacy, hetGP: Heteroskedastic Gaussian Process Modeling and Design under Replication. *R package version* **1**, (2017).
66. A. M. Gopakumar, P. V. Balachandran, D. Xue, J. E. Gubernatis, T. Lookman, Multi-objective optimization for materials discovery via adaptive design. *Scientific reports* **8**, 3738 (2018).
67. R. Dehghannasiri, D. Xue, P. V. Balachandran, M. R. Yousefi, L. A. Dalton, T. Lookman, E. R. Dougherty, Optimal experimental design for materials discovery. *Computational Materials Science* **129**, 311-322 (2017).
68. M. Binois, J. Huang, R. B. Gramacy, M. Ludkovski, Replication or Exploration? Sequential Design for Stochastic Simulation Experiments. *Technometrics* **61**, 7-23 (2019).
69. I. M. Sobol′, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* **55**, 271-280 (2001).
70. A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. (Halsted Press, 2004).
71. R. C. Team. (Vienna, Austria, 2013).
72. M. R. Hestenes, Multiplier and gradient methods. *Journal of Optimization Theory and Applications* **4**, 303-320 (1969).
73. Y. Ye, Ph. D. thesis, Department of ESS, Stanford University, (1987).
74. A. Ghalanos, S. Theussl, Rsolnp: general non-linear optimization using augmented Lagrange multiplier method. *R package version* **1**, (2012).
75. T. Smith, N. Maire, K. Dietz, G. F. Killeen, P. Vounatsou, L. Molineaux, M. Tanner, Relationship between the entomologica inoculation rate and the force of infection for

Plasmodium Falciparum malaria. *The American Journal of Tropical Medicine and Hygiene* **75**, 11-18 (2006).

76. W. E. Collins, G. M. Jeffery, A retrospective examination of the patterns of recrudescence in patients infected with Plasmodium falciparum. *The American journal of tropical medicine and hygiene* **61**, 44-48 (1999).

77. G. F. Killeen, A. Ross, T. Smith, Infectiousness of malaria-endemic human populations to vectors. *The American journal of tropical medicine and hygiene* **75**, 38-45 (2006).

78. A. Ross, T. Smith, The effect of malaria transmission intensity on neonatal mortality in endemic areas. *The American journal of tropical medicine and hygiene* **75**, 74-81 (2006).

79. E. M. Stuckey, T. Smith, N. Chitnis, Seasonally Dependent Relationships between Indicators of Malaria Transmission and Disease Provided by Mathematical Model Simulations. *PLOS Computational Biology* **10**, e1003812 (2014).

Supplementary Materials:

Title: Combining machine learning and mathematical models of disease dynamics to guide development of novel disease interventions

Authors: Monica Golumbeanu[1,2]†, Guojing Yang[1,2]†, Flavia Camponovo[1,2,3], Erin M. Stuckey[4], Nicholas Hamon[5], Mathias Mondy[5], Sarah Rees[5], Nakul Chitnis[1,2], Ewan Cameron[6,7,8], and Melissa A. Penny[1,2]*

Affiliations:
[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[3]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA
[4]The Bill and Melinda Gates Foundation, Seattle, WA, USA
[5]Innovative Vector Control Consortium, Liverpool, United Kingdom
[6]Malaria Atlas Project, Big Data Institute, University of Oxford, Oxford, UK
[7]Curtin University, Perth, Australia
[8]Telethon Kids Institute, Perth Children's Hospital, Perth, Australia
*Correspondence to: melissa.penny@unibas.ch
†These authors contributed equally.

One Sentence Summary: Defining quantitative profiles of novel disease interventions by combining machine learning with mathematical models of disease transmission

Short title: Quantitatively defining new disease interventions

Supplementary Materials:

Table of contents

1 Disease model 41
Table S1.1 Overview of the OpenMalaria model components. 42
2 Disease scenarios 43
Figure S2.1 Illustration of the yearly malaria transmission and prevalence patterns in simulated seasonal settings. 43
Figure S2.2 Simulated distributions of true and patent (detected with PCR or RDT) *Pf*PR$_{0-99}$ and *Pf*PR$_{2-10}$ for various input EIR levels in absence of interventions. 45
Figure S2.3 Representation of decay and the range of efficacy and half-life against different parasite or vector targets for intervention-agnostic malaria interventions. 46
Table S2.1 Description and ranges of simulation variables. 47
Table S2.2 Specifications of the optimization procedure for TPP development. 48
3 Results: Disease model simulations 50
Figure S3.1. Examples of OpenMalaria simulation outputs. 50
Figure S3.2. Distributions of prevalence reduction following yearly deployment of single interventions. 51
Figure S3.3. Distributions of prevalence reduction following yearly deployment of combinations of interventions. 52

Figure S3.4. Distributions of prevalence reduction following deployment of single and combinations of interventions twice per year. 53
Figure S3.5. Simulations reaching malaria elimination before intervention deployment. 54
4 Results: Emulator performance..... 55
Figure S4.1. Assessment of the performance of the trained GP depending on the training set size. 55
Figure S4.2. Performance of the trained GP emulators predicting immediate intervention impact. 56
Figure S4.3. Performance of the trained GP emulators predicting long-term intervention impact. 57
Table S4.1. Performance of the trained GP emulators predicting immediate and long-term intervention impact. 58
5 Results: Summary of key intervention impact determinants, optimal intervention profiles, and vaccine results 59
Table S5.1. Key findings guiding target product profiles of new malaria interventions. 60
6 Results: Key determinants of impact 61
Figure S6.1. Key drivers of impact for therapeutic malaria interventions across different transmission settings. 61
Figure S6.2. Key drivers of impact for vector control malaria interventions across different transmission settings. 63
7 Results: Feasible landscapes of optimal, constrained intervention profiles 64
Figure S7.1. Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed once per year. 64
Figure S7.2. Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed twice per year. 65
Figure S7.3 Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective vaccine deployed once per year. 66
Figure S7.4. Feasible landscapes of optimal, constrained intervention profiles (TPPs) for a transmission-blocking vaccine deployed once per year. 67
Figure S7.5. Feasible landscapes of optimal, constrained intervention profiles (TPPs) for attractive targeted sugar baits deployed once or twice per year. 68
Figure S7.6. Feasible landscapes of optimal, constrained intervention profiles (TPPs) for eave tubes deployed once per year. 69
8 Results: Optimal intervention profiles 70
Figure S8.1 Optimal intervention profiles (TPPs) for anti-infective monoclonal antibodies under various deployment regimes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%. 70
Figure S8.2 Optimal intervention profiles (TPPs) for anti-infective vaccines under various deployment regimes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%. 71
Figure S8.3 Optimal intervention profiles (TPPs) for transmission-blocking vaccines under various deployment regimes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%. 72
Figure S8.4 Optimal intervention profiles (TPPs) for attractive targeted sugar baits under various deployment regimes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%. 73
Figure S8.5 Optimal intervention profiles (TPPs) for eave tubes to achieve a $Pf\mathrm{PR}_{0\text{-}99}$ reduction of at least 70%. 74

1203 **1 Disease model**
1204
1205

Name	Description and assumptions	References
Key modelled epidemiological processes		
Malaria infection of humans	- Determined by EIR which is a model input and affects the force of infection in the simulated setting - Exposure of humans to mosquitoes depends on age	(*33, 75*)
Infection progression in humans: asexual parasite densities and immunity	- Blood stage parasite density depends on the time since infection and is affected by naturally acquired immunity - The duration of infection follows a log-normal distribution - Immunity (both pre-erythrocytic and blood-stage) develops progressively following consequent episodes of exposure to infection and decays exponentially - Acquired immunity reduces parasite density of subsequent infections - Super-infection is possible with cumulative parasite densities	(*33, 48, 75, 76*)
Transmission from infected humans to mosquitoes	- Depends on the density of parasites present in the human with gametocyte densities following a lag from parasite densities	(*33, 49, 77*)
Clinical illness, morbidity, mortality and anemia	- Acute clinical illness depends on human host parasite densities and their pyrogenic threshold which evolves over time depending on the individual exposure history - Acute morbidity episodes can be uncomplicated or evolve to severe episodes - A proposition of the severe episodes leads to deaths	(*33, 50, 51, 78*)
Modelled characteristics of the transmission setting		
Population age structure	- Informed by health and demographic surveillance data from Tanzania	(*50, 57*)
Transmission seasonality	- Seasonally-forced, the same transmission pattern is reproduced each year in absence of interventions	(*75, 79*)
Case management	- Modelled through a comprehensive decision tree-based model which determines the corresponding treatment implications depending on the occurring clinical events such as fevers and seeking of care - Its representation includes specification of diagnostic tests, effects of treatment, case fatality, case sequelae and cure rates	(*52*)
Entomological setting	- Comprehensive simulation of the mosquito lifecycle and behavior towards human and animal hosts (biting, resting) embedded in a dynamic entomological model of the mosquito oviposition cycle - Multiple vector species can be simulated simultaneously	(*47*)

Modelled interventions			
Vector control	-	Available interventions: long-lasting insecticide-treated nets (LLINs), indoor residual spraying (IRS), house screening, baited traps, repellents, push-pull	
Drugs and Vaccines	-	Drugs and vaccines acting at various levels of the parasite life cycle (transmission blocking, anti-infective, blood-stage clearance)	
Deployment characteristics	-	Interventions can be deployed for several rounds to a targeted group of individuals and specified coverages	
Simulation regimes and model variants			
Time steps	-	Simulation outputs are tracked every 5 days	
Model variants	-	Varying assumptions in immunity decay, treatment and heterogeneity of transmission result in 14 model variants	(*45*)
Software availability and documentation			
- Source code and wiki page available on GitHub: https://github.com/SwissTPH/openmalaria/			

Table S1.1
Overview of the OpenMalaria model components.
The individual based stochastic model of malaria in humans and transmission has been described previously. This model was originally developed in 2003-2006 (*33*), with mosquito dynamics updated in 2008 (*46*) and an additional 13 structural model variants developed and parameterized in 2012 (*45*) representing different model assumptions on immunity decay, disease, comorbidities, and heterogeneity in transmission.

2 Disease scenarios

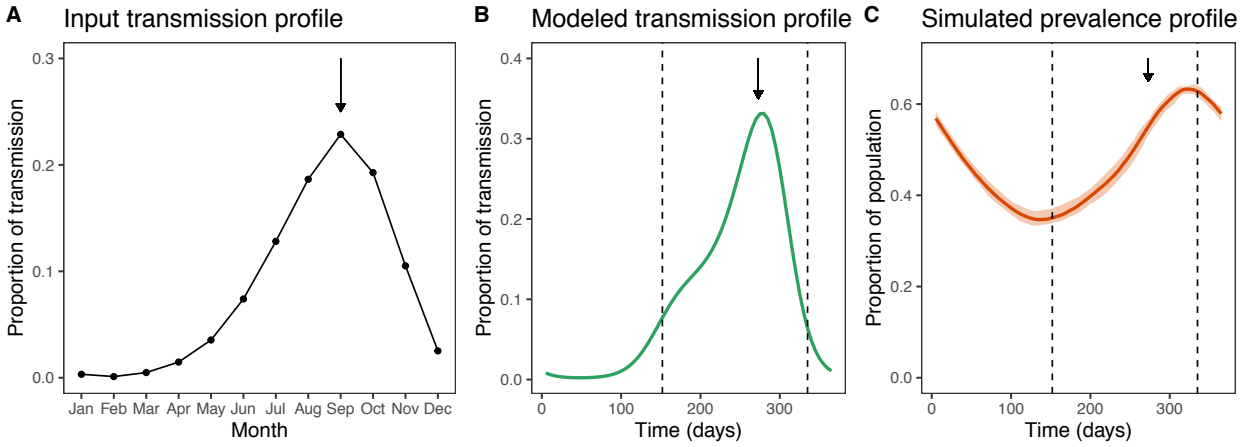
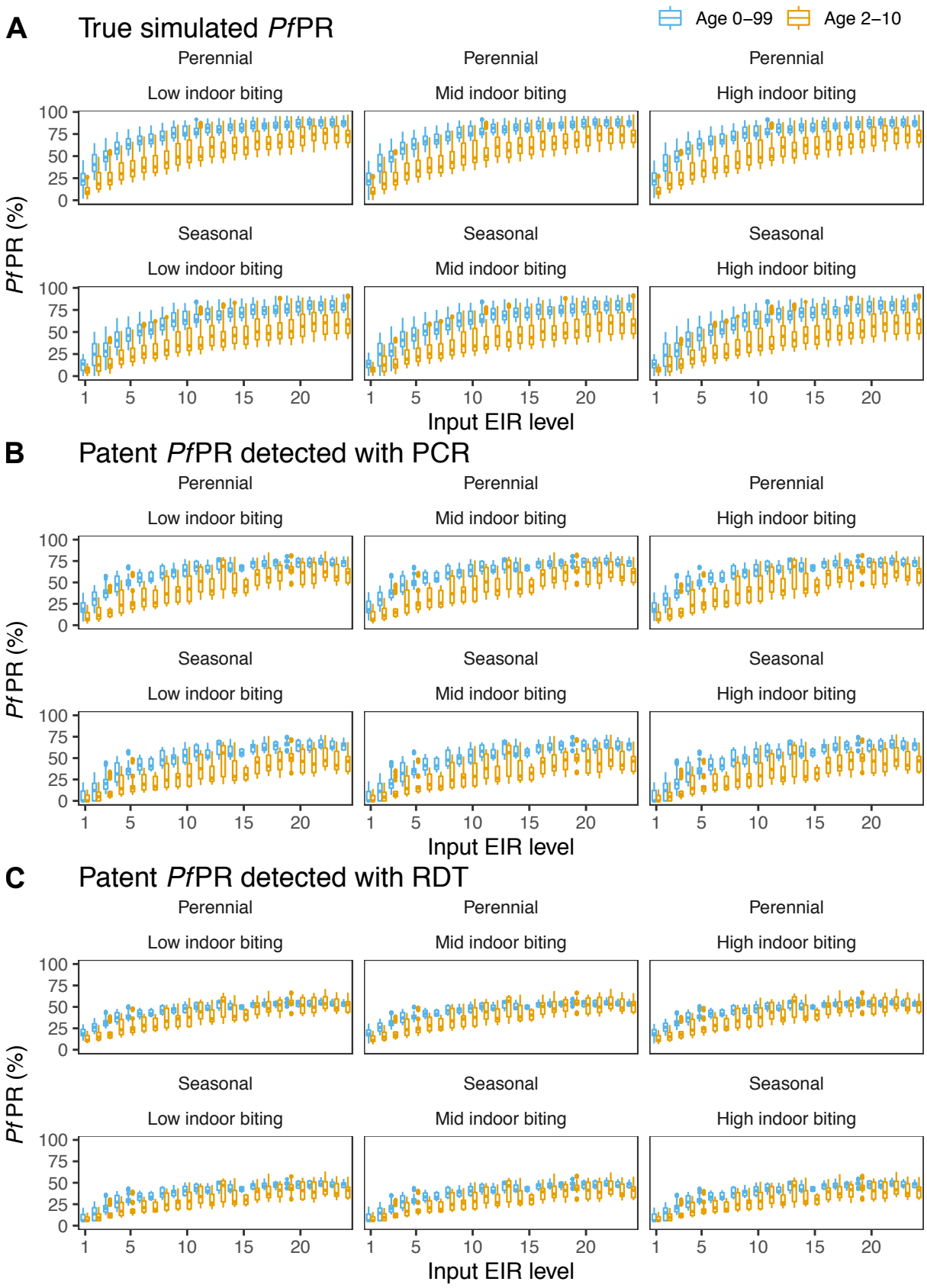


Figure S2.1

Illustration of the yearly malaria transmission and prevalence patterns in simulated seasonal settings.

(A) Observed, normalized, monthly seasonal pattern of malaria EIR in Namawala, Tanzania extracted from (*50*). **(B)** Corresponding input, 5-day seasonal EIR pattern used in OpenMalaria simulations, obtained by scaling and extrapolating the monthly seasonality profile from (*50*) to 5-day time steps. For this example, the simulated input EIR was 7.78 infectious bites per person per year. **(C)** Resulting simulated yearly *Pf*$PR_{0\text{-}99}$ profile. In all figures, the arrows indicate the month of September, the peak of transmission and show the delay between the peak of transmission and the resulting peak in malaria prevalence. The dotted vertical lines on figures **(B)** and **(C)** indicate the deployment times of first and second rounds of malaria interventions when applicable.


A
True simulated *Pf*PR
Age 0–99
Age 2–10
Perennial
Low indoor biting
Mid indoor biting
High indoor biting
Seasonal
*Pf*PR (%)
Input EIR level
B
Patent *Pf*PR detected with PCR
C
Patent *Pf*PR detected with RDT

1230

Figure S2.2

Simulated distributions of true and patent (detected with PCR or RDT) *Pf*PR$_{0\text{-}99}$ and *Pf*PR$_{2\text{-}10}$ for various input EIR levels in absence of interventions.

The input entomological inoculation rate (EIR) defines the simulated malaria transmission level. In every simulation experiment, EIR was uniformly sampled from the interval [1, 25]. In figures (A) – (C), each panel corresponds to a simulated setting and presents the distributions of true **(A)**, patent with PCR **(B)** and patent with RDT **(C)** *Plasmodium falciparum* prevalence (*Pf*PR, shown with boxplots, blue for 0-99 years old and orange for 2-10 years old) at varying EIR levels (x axis). The 6 represented settings are defined by the seasonality pattern (perennial shown in the first row, or seasonal shown in the second row of each figure) and mosquito indoor biting behavior (low- shown in the first column, mid- shown in the second column or high-indoor biting shown in the third column of each figure). Each EIR level on the x axis is defined as a set of continuous input EIR values which range between the current level and the current level - 1, e.g., an input EIR level of 1 contains EIR values in the interval (0, 1]. For each EIR level and setting, the case management levels, i.e., the probability of seeking care (access to treatment) within 5 days from the onset of fever (E_5), was varied within the interval [0.04 - 0.5]. PCR stands for polymerase chain reaction and RDT stands for rapid diagnostic test.

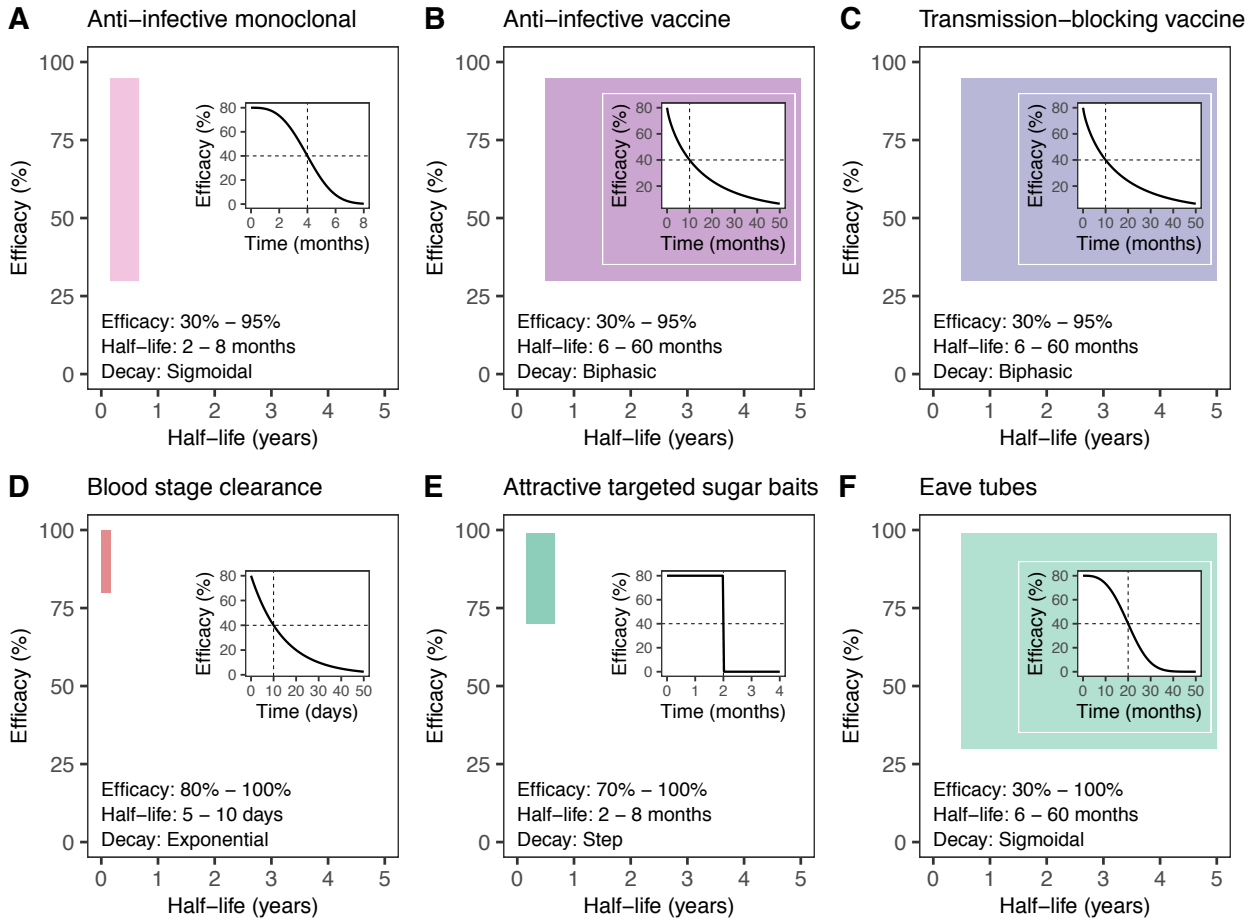


Figure S2.3

Representation of decay and the range of efficacy and half-life against different parasite or vector targets for intervention-agnostic malaria interventions.

The simulated malaria interventions (**A – F**) were modeled in terms of their targets in the malaria transmission cycle. The effect of each intervention is represented through the half-life of its decay (x axis) as well as the initial efficacy (y axis). The color blocks represent the range of parameter space of efficacy and half-life of decay considered in the current analysis for each intervention. The half-life and the color block does not represent the entire duration of effect, as that depends on the decay shape chosen for each intervention. The decay shape for each intervention is displayed in the right side insert of each plot where the dotted lines specify the half-life and corresponding half of the intervention efficacy. The definitions of all the parameter ranges for all interventions are provided in each figure on the lower left side and detailed in Table S2.1.

	Intervention	Coverage	Initial efficacy	Half-life or duration (years)	Decay type
Intervention profiles	Prevent infection				
	Anti-infective vaccine	0 - 1	0.3 - 0.95	0.5 - 5	Weibull (k = 0.8) (Sigmoidal)
	Anti-infective monoclonal antibody	0 - 1	0.3 - 0.95	0.167 - 0.667	Weibull (k = 3) (Biphasic)
	Blood stage clearance				
	Antimalarial drugs	0 - 1	0.8 - 1	0 - 0.1667	Exponential
	Transmission blocking				
	Vaccine	0 - 1	0.3 - 0.95	0.5 - 5	Weibull (k = 0.8) (Biphasic)
	Preprandial killing effect (affects only indoor mosquito biting)				
	Eave tubes	0 - 1	0.3 - 0.99	0.5 - 5	Weibull (k = 3) (Sigmoidal)
	Preprandial and postprandial killing effect (affects indoor and outdoor mosquito biting)				
	Attractive targeted sugar baits	0 -1	0.7 - 0.99	0.167 – 0.667	Step
Transmission settings	EIR range: 1 – 25, representing a *Pf*PR$_{0-99}$ of 13-88% and a *Pf*PR$_{0-2}$ of 7.2-74% Case management (baseline scenario) range: 0 – 0.8, corresponding to a probability of seeking care within 5 days from the onset of fever of 0-0.5 Seasonality levels: 1. high seasonal setting with one transmission peak over a year 2. perennial setting with constant yearly transmission Proportion of indoor-biting mosquitoes, out of total indoor and outdoor biting mosquitoes: 3. high (0.8) 4. medium (0.5) 5. low (0.2)				

Table S2.1 Description and ranges of simulation variables.

Within each OpenMalaria simulation, the varied parameters and ranges correspond to the profiles of applied interventions (see Fig S2.3 for visual ranges of vector and parasite targets), as well as the transmission setting characteristics. The profile of each modeled malaria intervention is defined by its target, the ranges of the deployment coverage, initial efficacy, half-life or duration of effect as well the type of decay. Where interventions are applied to individual humans, in the present demonstrative analysis this is equally applied across ages, and not targeted to certain population. The transmission setting is defined by the yearly EIR, seasonality level, as well as proportion of indoor mosquitoes.

Intervention	Minimized profile	Intervention properties constraints	Specifications of combination therapies
Anti-infective monoclonal antibody (Sigmoidal decay)	Coverage	Coverage ∈ [0 .. 80%] Efficacy = 85% Half-life = 4 months	Blood stage drug: Efficacy = 90% Half-life = 10 days
	Efficacy	Coverage = 60% Efficacy ∈ [30% .. 95%] Half-life = 4 months	
	Half-life	Coverage = 60% Efficacy = 85% Half-life ∈ [2 .. 8 months]	
Anti-infective vaccine (Biphasic decay)	Coverage	Coverage ∈ [0 .. 80%] Efficacy = 85% Half-life = 7 months	Blood stage drug: Efficacy = 90% Half-life = 10 days
	Efficacy	Coverage = 60% Efficacy ∈ [30% .. 95%] Half-life = 7 months	
	Half-life	Coverage = 60% Efficacy = 85% Half-life ∈ [2 months .. 5 years]	
Transmission-blocking vaccine (Biphasic decay)	Coverage	Coverage ∈ [0 .. 80%] Efficacy = 85% Half-life = 7 months	Blood stage drug: Efficacy = 90% Half-life = 10 days
	Efficacy	Coverage = 60% Efficacy ∈ [30% .. 95%] Half-life = 7 months	
	Half-life	Coverage = 60% Efficacy = 85% Half-life ∈ [2 months .. 5 years]	
Attractive targeted sugar baits (Step decay)	Coverage	Coverage ∈ [0 .. 80%] Efficacy = 85% Half-life = 4 months	Not applicable
	Efficacy	Coverage = 60% Efficacy ∈ [70% .. 99%] Half-life = 4 months	
	Half-life	Coverage = 60% Efficacy = 85% Half-life ∈ [2 .. 8 months]	
Eave tubes (Sigmoidal decay)	Coverage	Coverage ∈ [0 .. 80%] Efficacy = 85% Half-life = 3 years	Not applicable
	Efficacy	Coverage = 60% Efficacy ∈ [30% .. 99%] Half-life = 3 years	
	Half-life	Coverage = 60% Efficacy = 85% Half-life ∈ [6 months .. 5 years]	

Table S2.2

Specifications of the optimization procedure for TPP development.

For each intervention, we successively identified the minimum profiles of the intervention coverage, efficacy, and half-life. Precisely, we optimized each parameter separately (column "Minimized profile"), according to its feasibility constraints while setting the two other parameters

to the specified fixed values (column “Intervention properties constraints”). When deployed in combination with other drugs or vaccines, the additional interventions had fixed properties as well (column “Specifications of combination therapies”).

3 Results: Disease model simulations

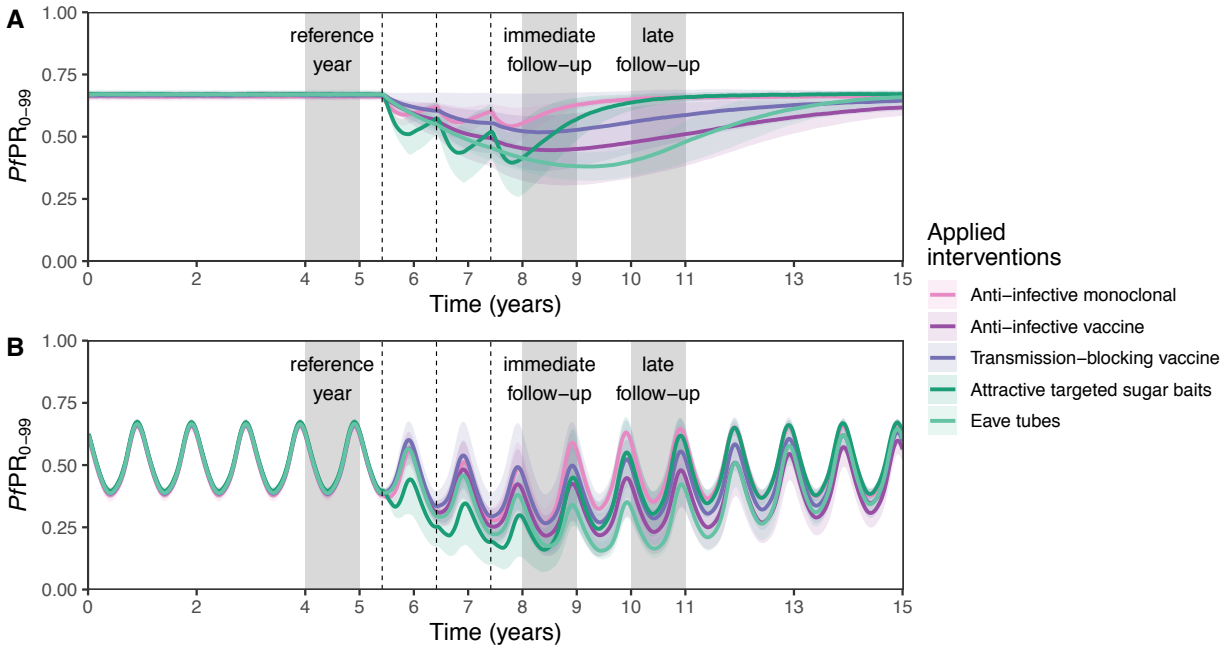


Figure S3.1.
Examples of OpenMalaria simulation outputs.

Time series of simulated malaria *Pf*PR$_{0\text{-}99}$ in a perennial **(A)** and seasonal **(B)** setting. Both figures display the prevalence of malaria cases, *Pf*PR$_{0\text{-}99}$, (y axis) across time (x axis). Interventions targeting different stages in the malaria transmission cycle (different colors) are applied once per year at the beginning of June (vertical dotted lines, in this example for three years of deployment). The effect of each intervention is assessed by evaluating the *Pf*PR$_{0\text{-}99}$ reduction in all ages relative to the year prior deployment (first grey block). Two outcomes are assessed, following an immediate and late follow-up (second and third grey blocks), depending on whether the average prevalence is calculated across the next year after deployment, or across the third year after deployment, respectively.

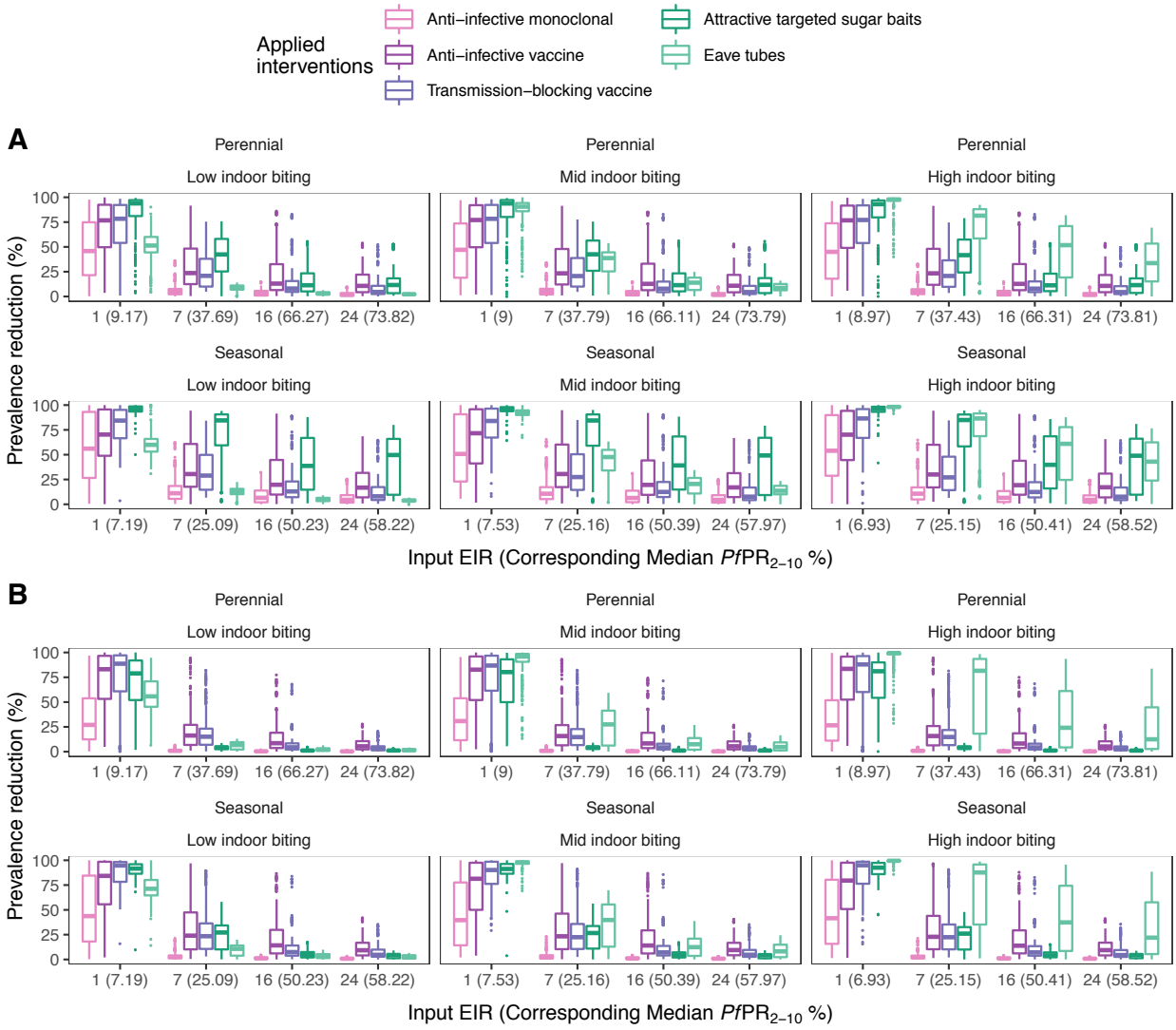


Figure S3.2.

Distributions of prevalence reduction following yearly deployment of single interventions.

Prevalence reduction was calculated by comparing the initial prevalence in the year before any interventions were deployed to the yearly prevalence obtained in the following year (short follow-up, panel **A**) and in the third year (long follow-up, panel **B**) after deployment of interventions. Each individual figure corresponds to a simulated setting and presents the distributions of $Pf\mathrm{PR}_{0\text{-}99}$ reduction (shown with boxplots) at varying EIR as well as corresponding simulated $Pf\mathrm{PR}_{2\text{-}10}$ levels (x axis). Each boxplot displays the interquartile range (box), the median value (horizontal line), the largest and smallest values within 1.5 times the interquartile range (whiskers), and the remaining outside values (points). The 6 represented settings in each panel are defined by the seasonality pattern (perennial or seasonal) and mosquito indoor biting behavior (low, mid or high indoor biting). Each EIR level on the x axis is defined as a set of continuous input EIR values which range between the current level and the current level - 1, e.g., an input EIR level of 1 contains all EIR values in the interval (0, 1]. The definitions and ranges of all the EIR levels is included in Table S2.1.

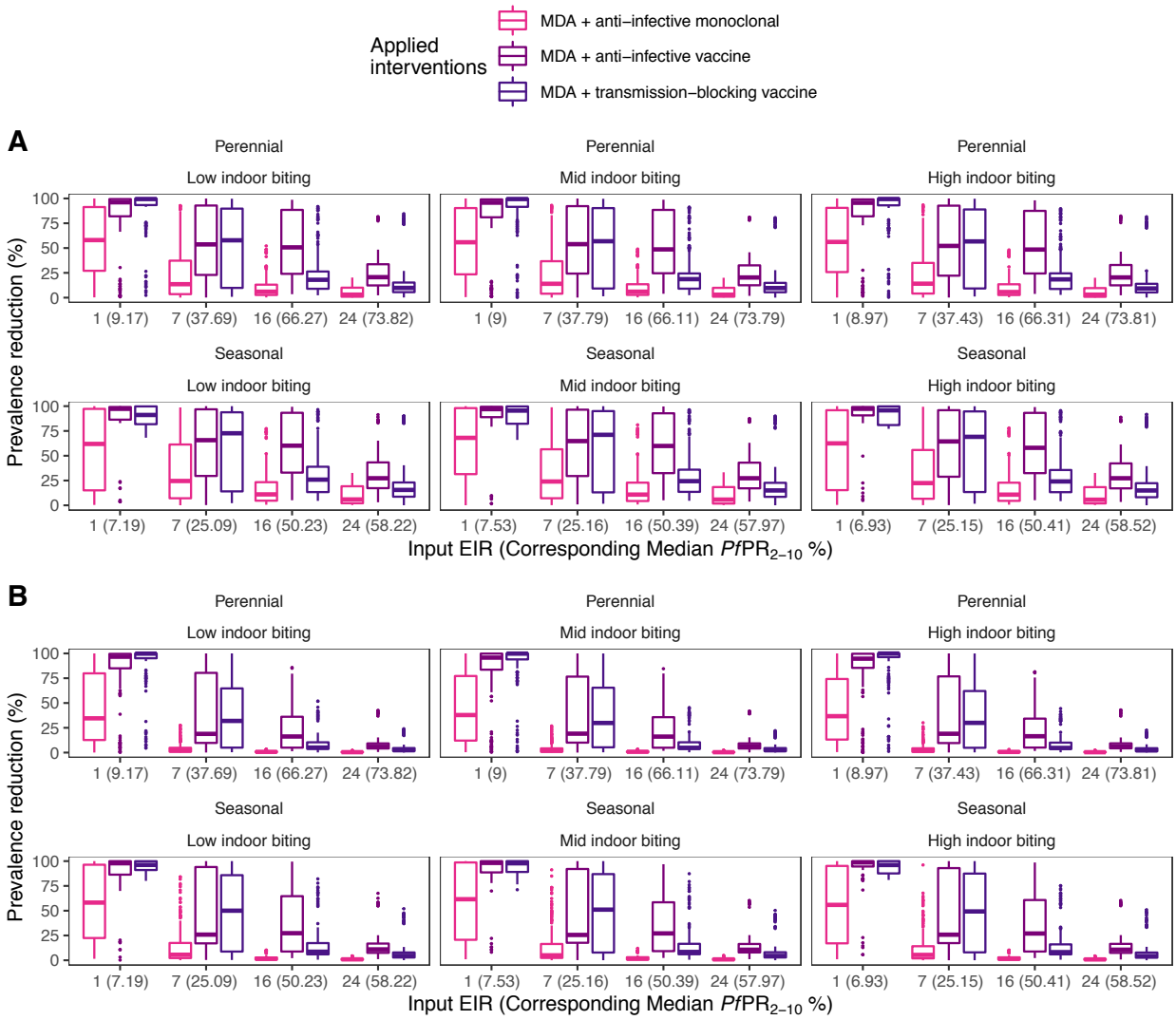


Figure S3.3.

Distributions of prevalence reduction following yearly deployment of combinations of interventions.

Prevalence reduction was calculated by comparing the initial prevalence in the year before any interventions were deployed to the yearly prevalence obtained in the following year (short follow-up, panel **A**) and in the third year (long follow-up, panel **B**) after deployment of interventions. Each individual figure corresponds to a simulated setting and presents the distributions of $Pf\mathrm{PR}_{0\text{-}99}$ reduction (shown with boxplots) at varying EIR as well as the corresponding simulated $Pf\mathrm{PR}_{2\text{-}10}$ levels (x axis). Each boxplot displays the interquartile range (box), the median value (horizontal line), the largest and smallest values within 1.5 times the interquartile range (whiskers), and the remaining outside values (points). The 6 represented settings in each panel are defined by the seasonality pattern (perennial or seasonal) and mosquito indoor biting behavior (low, mid or high indoor biting). Each EIR level on the x axis is defined as a set of continuous input EIR values which range between the current level and the current level - 1, e.g., an input EIR level of 1 contains EIR values in the interval (0, 1]. The definitions and ranges of all the EIR levels are included in Table S2.1. MDA stands for mass drug administration.

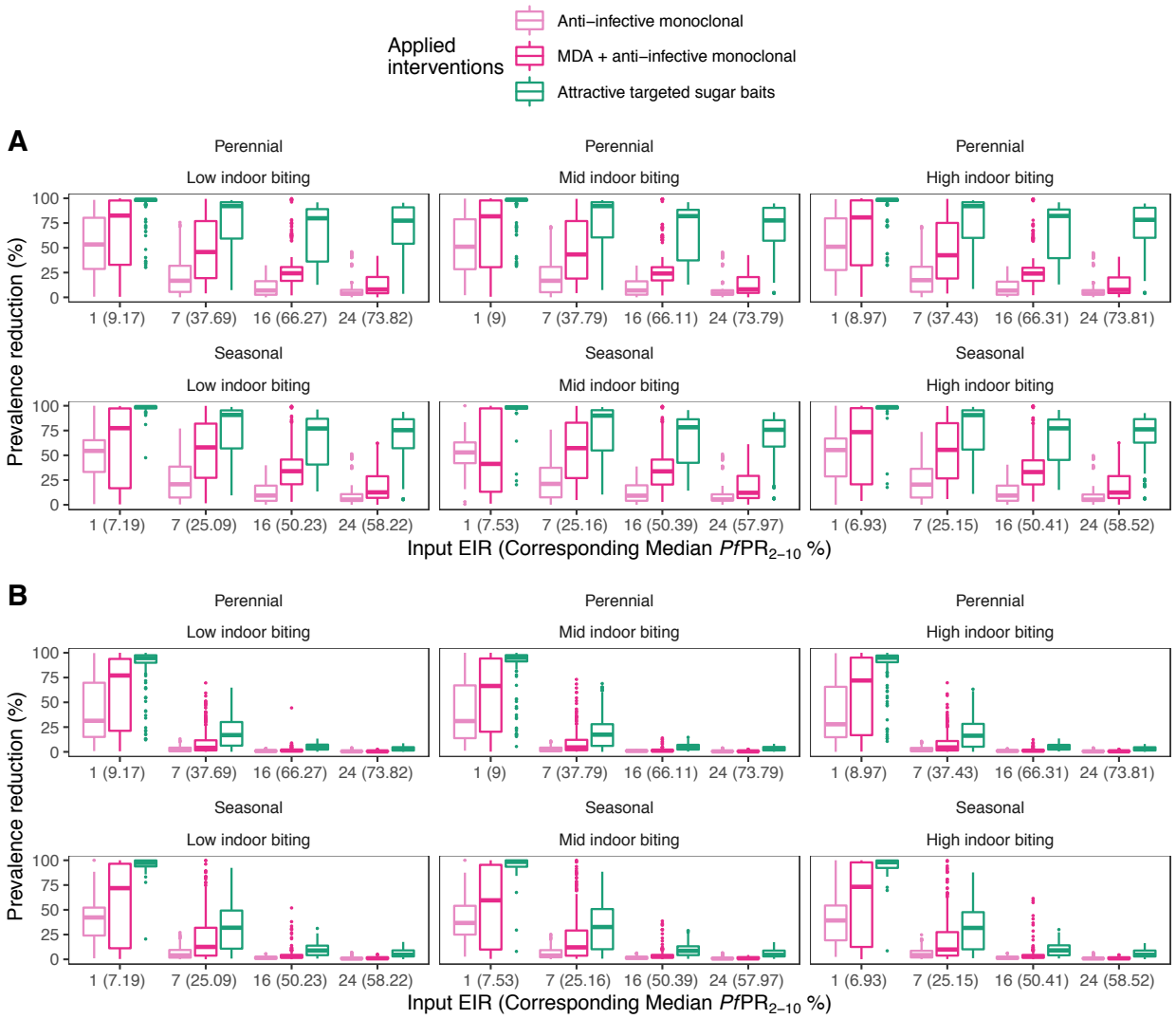


Figure S3.4.

Distributions of prevalence reduction following deployment of single and combinations of interventions twice per year.

Prevalence reduction was calculated by comparing the initial prevalence in the year before any interventions were deployed to the yearly prevalence obtained in the following year (short follow-up, panel **A**) and in the third year (long follow-up, panel **B**) after deployment of interventions. Each individual figure corresponds to a simulated setting and presents the distributions of $Pf\mathrm{PR}_{0\text{-}99}$ reduction (shown with boxplots) at varying EIR as well as corresponding simulated $Pf\mathrm{PR}_{2\text{-}10}$ levels (x axis). Each boxplot displays the interquartile range (box), the median value (horizontal line), the largest and smallest values within 1.5 times the interquartile range (whiskers), and the remaining outside values (points). The 6 represented settings in each panel are defined by the seasonality pattern (perennial or seasonal) and mosquito indoor biting behavior (low, mid or high indoor biting). Each EIR level on the x axis is defined as a set of continuous input EIR values which range between the current level and the current level - 1, e.g., an input EIR level of 1 contains EIR values in the interval (0, 1]. The definitions and ranges of all the EIR levels for all simulated settings is included in Table S2.1. MDA stands for mass drug administration.

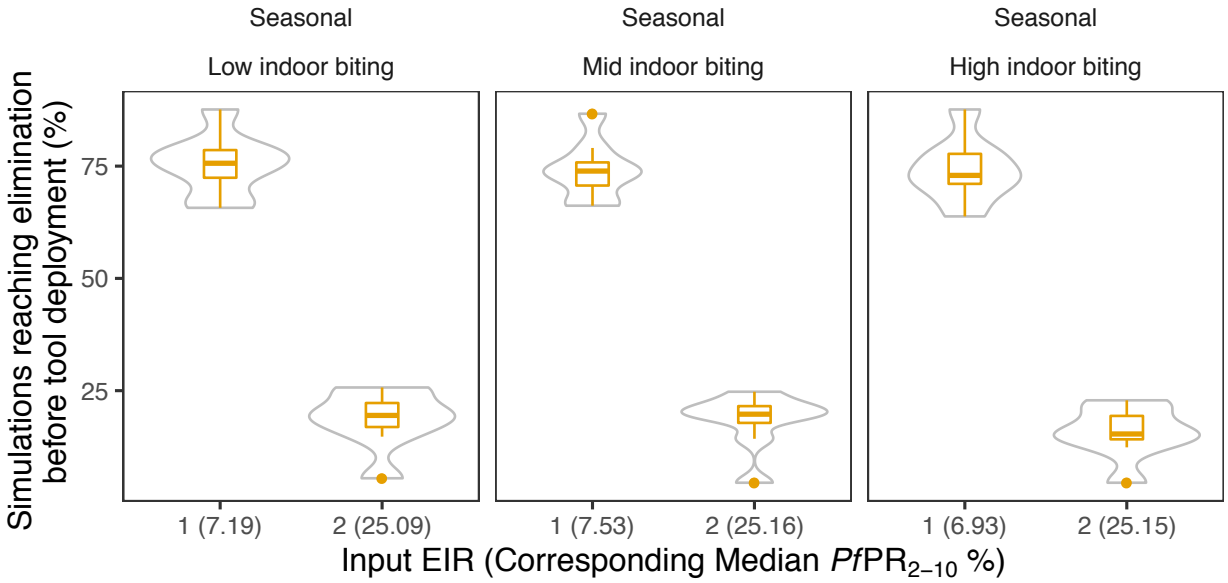


Figure S3.5.

Simulations reaching malaria elimination before intervention deployment.

The violin plots and boxplots in each panel present the distributions of the percentage of simulations reaching malaria elimination ($Pf\mathrm{PR}_{0\text{-}99} = 0$) before intervention deployment (this can arrive due to case management and only occurs in seasonal settings), across all simulated interventions and intervention combinations.

4 Results: Emulator performance

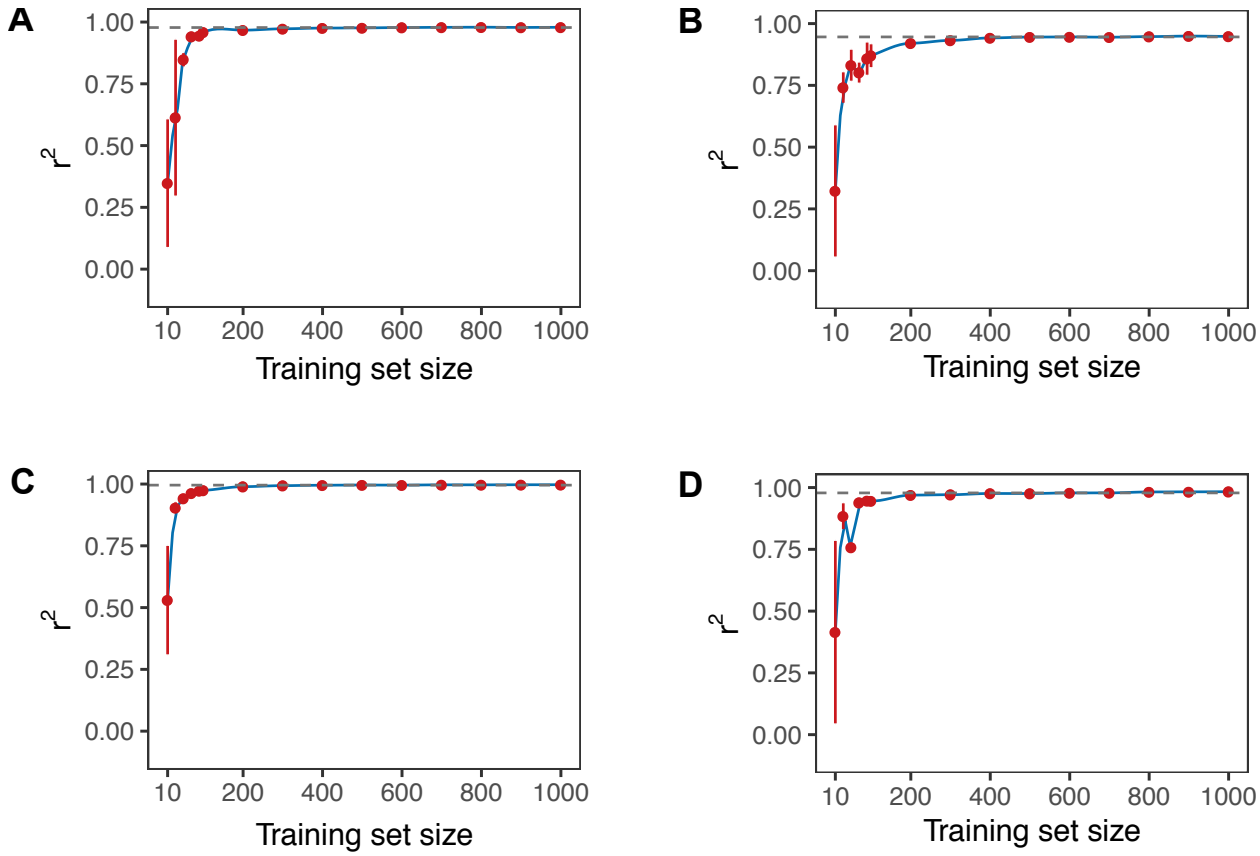


Figure S4.1.
Assessment of the performance of the trained GP depending on the training set size.
Each figure presents the Pearson correlation coefficient r^2 between true and predicted values on a broad range of out-of-sample test sets of varying length, when simulating deployment of an anti-infective monoclonal antibody deployed once per year **(A)** or twice per year **(B)** as well as in combination with a blood-stage drug once **(C)** or twice per year **(D)**.

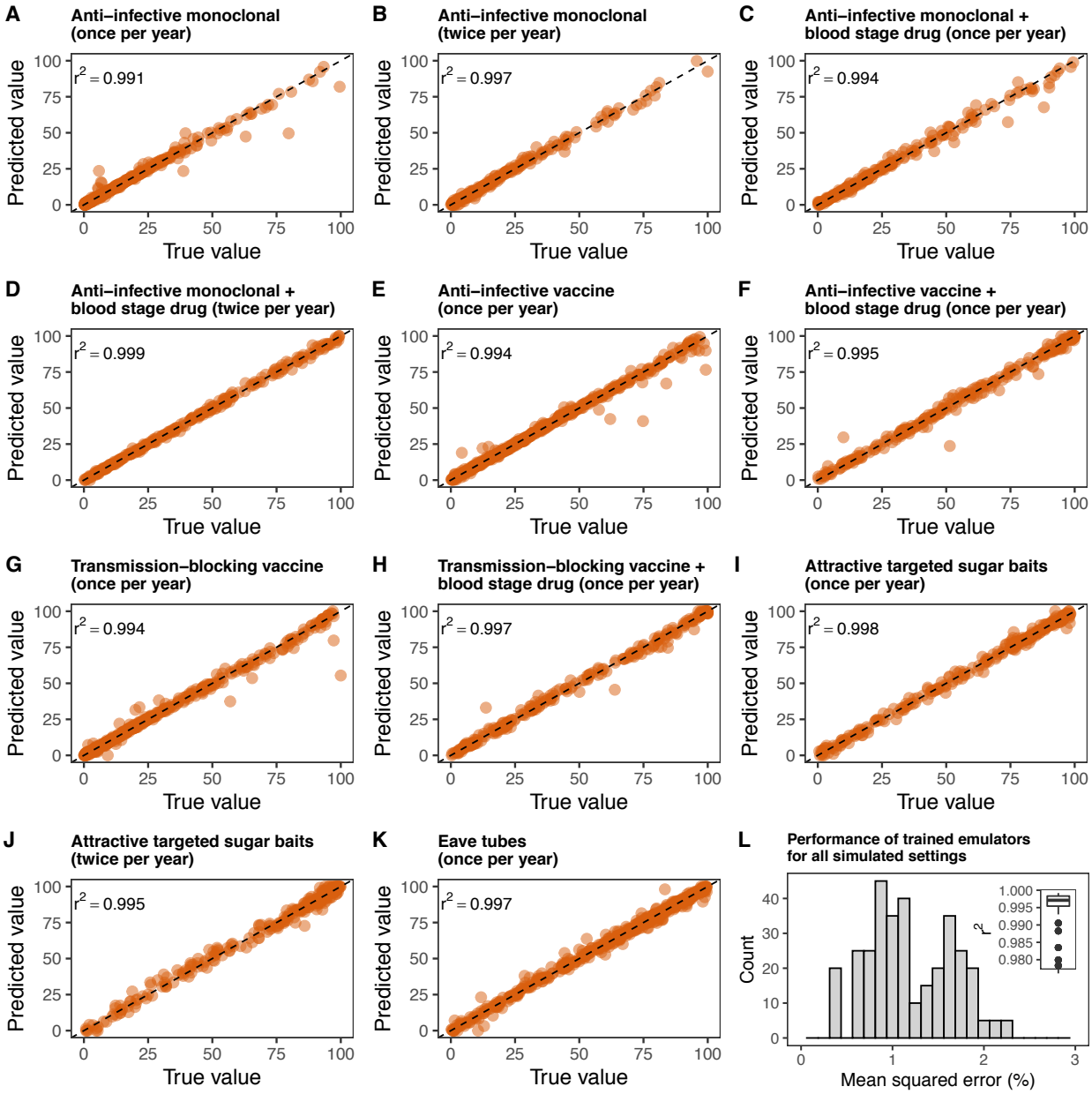


Figure S4.2.

Performance of the trained GP emulators predicting immediate intervention impact.

For a wide range of deployed interventions and transmission settings (see Materials and Methods), GP emulators were trained to predict the immediate impact of each intervention, i.e., the resulting average *Pf*PR$_{0\text{-}99}$ reduction in the year following deployment of the intervention. The performance of the trained emulators was assessed by inspecting the Pearson correlation coefficient (r^2) and the mean squared error between true and predicted values on an out-of-sample test set. Figures **(A) – (K)** display the true and predicted values of each trained emulator across all deployed interventions in a seasonal transmission setting with high indoor biting. Figure **(L)** summarizes r^2 and the mean squared error of all the trained emulators for all simulated transmission settings and interventions (the simulated settings were defined by seasonality and mosquito biting patterns, see Table S2.1 for detailed values per setting).

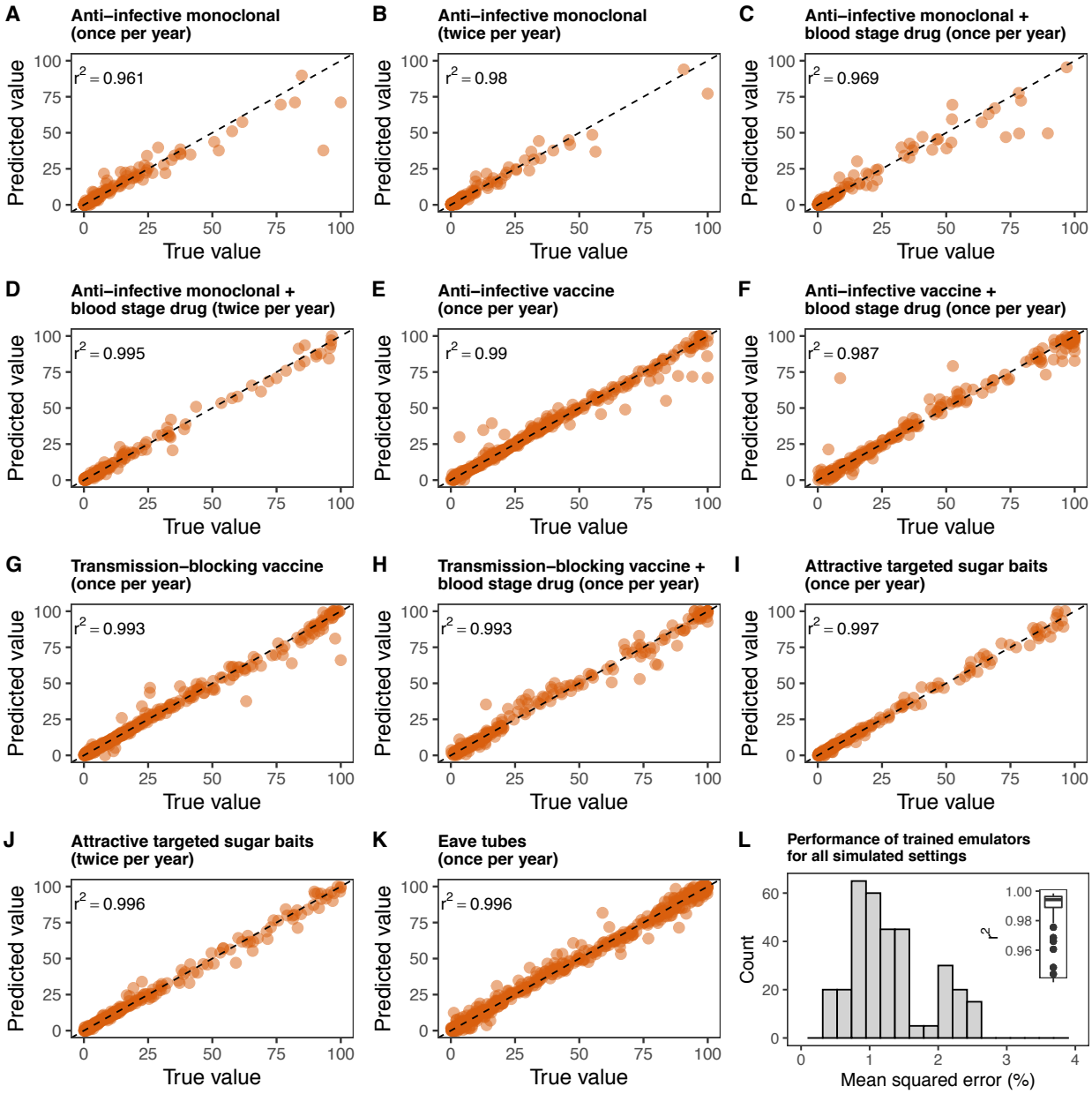


Figure S4.3.

Performance of the trained GP emulators predicting long-term intervention impact.

For a wide range of deployed interventions and transmission settings (see Methods), GP emulators were trained to predict the immediate impact of each intervention, i.e., the resulting average $Pf\mathrm{PR}_{0\text{-}99}$ reduction in the third year following deployment of the intervention. The performance of the trained emulators was assessed by inspecting the Pearson correlation coefficient (r^2) and the mean squared error between true and predicted values on an out-of-sample test set. Figures **(A) – (K)** display the true and predicted values of each trained emulator across all deployed interventions in a seasonal transmission setting with high indoor biting. Figure **(L)** summarizes r^2 and the mean squared error of all the trained emulators for all simulated transmission settings and interventions (the simulated settings were defined by seasonality and mosquito biting patterns, see Table S2.1 for detailed values per setting).

Intervention(s) (deployment)	Training set size	Test set size	Cross-validation r^2 and (mean error)	Test set r^2 and (mean error)
Anti-infective monoclonal antibody (once/year)	10000	5000	**Immediate:** 0.99 (1.02%) **Long:** 0.96 (1.15%)	**Immediate:** 0.99 (0.63%) **Long:** 0.97 (0.68%)
Anti-infective monoclonal antibody (twice/year)	5000	2500	**Immediate:** 0.99 (1.11%) **Long:** 0.97 (1.32%)	**Immediate:** 0.99 (0.91%) **Long:** 0.99 (0.83%)
Anti-infective monoclonal antibody + Blood stage drug (once/year)	10000	5000	**Immediate:** 0.99 (1.34%) **Long:** 0.96 (1.74%)	**Immediate:** 0.99 (1.18%) **Long:** 0.98 (1.05%)
Anti-infective monoclonal antibody + Blood stage drug (twice/year)	5000	2500	**Immediate:** 0.99 (1.26%) **Long:** 0.97 (1.98%)	**Immediate:** 0.99 (0.98%) **Long:** 0.99 (1.12%)
Anti-infective vaccine (once/year)	10000	5000	**Immediate:** 0.99 (1.08%) **Long:** 0.99 (1.3%)	**Immediate:** 0.99 (0.99%) **Long:** 0.99 (1.16%)
Anti-infective vaccine + Blood stage drug (once/year)	5000	2500	**Immediate:** 0.99 (1.18%) **Long:** 0.99 (1.63%)	**Immediate:** 0.99 (1.57%) **Long:** 0.99 (2.25%)
Transmission-blocking vaccine (once/year)	10000	5000	**Immediate:** 0.99 (1.13) **Long:** 0.99 (1.25%)	**Immediate:** 0.99 (0.89%) **Long:** 0.99 (1.07%)
Transmission-blocking vaccine + Blood stage drug (once/year)	5000	2500	**Immediate:** 0.99 (1.25%) **Long:** 0.99 (1.53%)	**Immediate:** 0.99 (1.68%) **Long:** 0.99 (2.23 %)
Attractive targeted sugar baits (once/year)	5000	2500	**Immediate:** 0.99 (1.26%) **Long:** 0.98 (1.71%)	**Immediate:** 0.99 (1.98%) **Long:** 0.99 (1.19%)
Attractive targeted sugar baits (twice/year)	5000	2500	**Immediate:** 0.99 (1.09%) **Long:** 0.99 (1.98%)	**Immediate:** 0.99 (1.03%) **Long:** 0.99 (1.29%)
Eave tubes (once/year)	10000	5000	**Immediate:** 0.99 (1.11%) **Long:** 0.99 (1.3%)	**Immediate:** 0.99 (0.89%) **Long:** 0.99 (1.26%)

Table S4.1.
Performance of the trained GP emulators predicting immediate and long-term intervention impact.

For each modelled transmission setting defined by case management level and mosquito biting patterns and for each intervention (Table S2.1), a comprehensive set of simulation scenarios was built by sampling uniformly the parameter space (defined in Table S2.1) and simulation with OpenMalaria. In this manner, a training and a test set were constructed. The training set was used to train, for each setting and intervention, a Heteroskedastic GP model in a 5-fold cross-validation procedure. The performance of the trained GP was assessed by computing the Pearson correlation coefficient r^2 as well as the mean error between the true and predicted outcomes on both out-of-sample cross-validation and test sets. For each intervention and follow-up (immediate or long-term), the average r^2 and mean error for all the GP models trained across 6 settings (seasonal or perennial, high, medium or low mosquito indoor biting) are reported.

1405 **5 Results: Summary of key intervention impact determinants, optimal intervention**
1406 **profiles, and vaccine results**
1407

Intervention	Summary of analysis results	Relevant figures
Therapeutic interventions - Anti-infective monoclonal antibodies - Anti-infective vaccines - Transmission-blocking vaccines	**Key determinants of impact** - The main driver of intervention impact was coverage - The second determinant of intervention impact depended on intervention half-life. For interventions with short half-lives such as monoclonal antibodies, the half-life was the second driver, while for long-term interventions such as vaccines, efficacy played a key role. - As opposed to long-term vaccines whose impact is mainly driven by coverage and efficacy, interventions with short half-life (e.g., anti-infective monoclonal antibodies) rely on the case management to prevent resurgence - The various biting patterns of mosquitoes did not influence the intervention determinants of impact	S6.1 and Figure 2
	Optimal intervention profiles - As opposed to vaccines, anti-infective monoclonal antibodies require high efficacy and deployment coverage while achieving limited reduction in *Pf*PR$_{0\text{-}99}$ with very little impact in perennial settings - Increasing the deployment frequency for anti-infective monoclonal antibodies from once to twice per year, extended the landscape of feasible health targets but mainly in seasonal settings - Combination with a blood-stage drug proved more impactful as compared to increasing the deployment frequency for anti-infective monoclonal antibodies, extending the achievable health goals in perennial settings as well	S7.1-S7.4 S8.1-S8.3 and Figure 3
Vector control interventions - Attractive targeted sugar baits - Eave tubes	**Key determinants of impact** - As with short-term therapeutic interventions such as anti-infective monoclonal antibodies, attractive targeted sugar baits rely on case management for preventing resurgence - We see limited difference between key drivers for attractive targeted sugar baits in different biting settings because mosquitoes sugar feed before indoor or outdoor biting. In contrast, we observe that intervention properties of eave tubes rather than health system access to treatment are larger drivers of impact in indoor biting	S6.2 and Figure 2

	settings, as mosquitoes in those settings will be more likely to contact the eave tube.	
	Optimal intervention profiles - Increasing deployment frequency from once to twice per year for attractive targeted sugar baits, resulted in a significant increase in intervention impact and less requirements in terms of coverage and half-life - Increasing efficacy of attractive targeted sugar baits did not have a significant impact	S7.5-S7.6 S8.4-S8.55 and Figure 3

Table S5.1.

Key findings guiding target product profiles of new malaria interventions.

A summary of key results concerning impact determinants and minimal intervention profiles as well as references to the corresponding illustrative figures is provided.

6 Results: Key determinants of impact

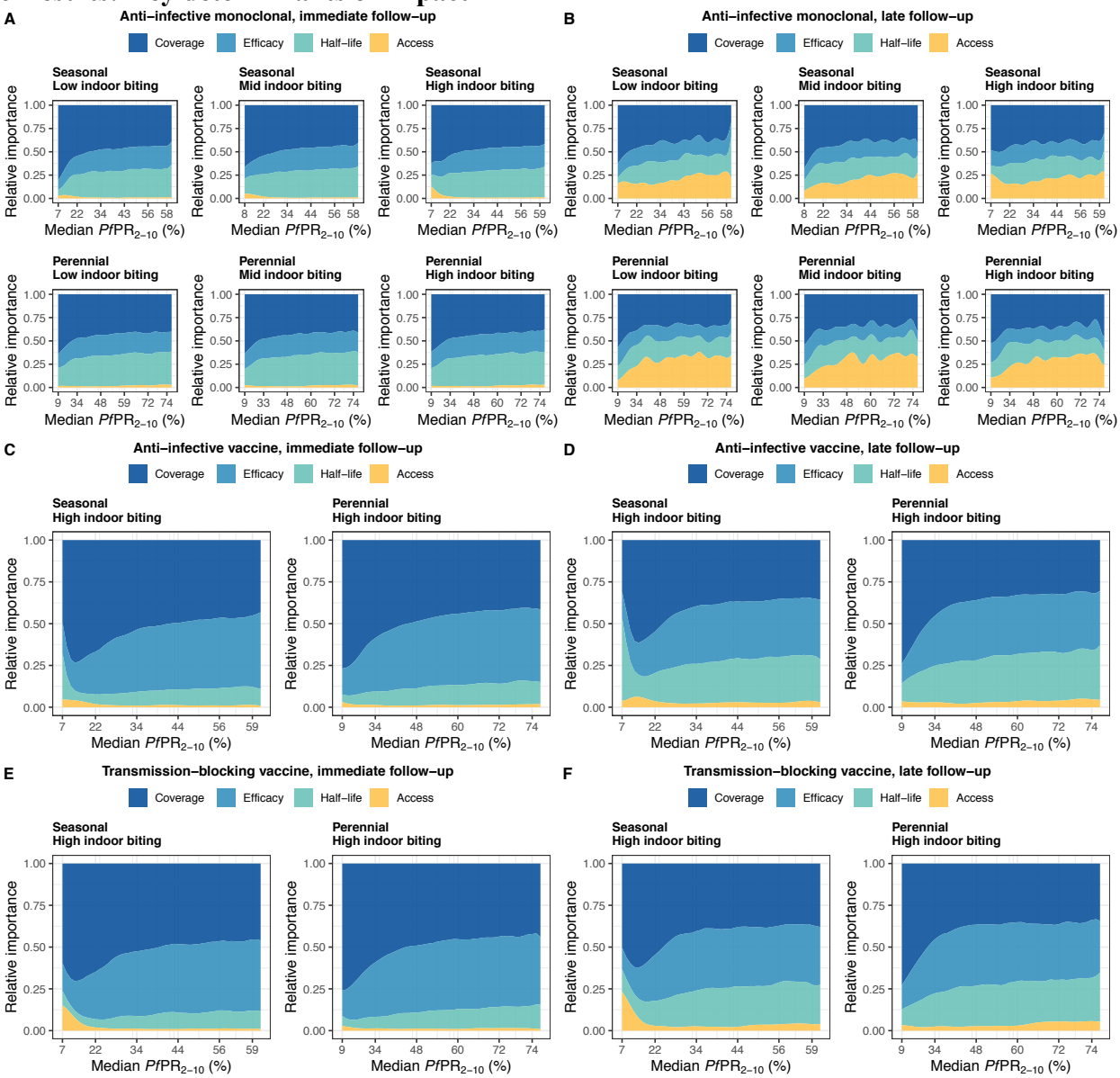


Figure S6.1.

Key drivers of impact for therapeutic malaria interventions across different transmission settings.

Results of sensitivity analysis identifying the determinants of intervention impact on *Pf*PR$_{0\text{-}99}$ reduction for anti-infective monoclonal antibodies **(A, B)**, anti-infective vaccines **(C, D)** and transmission-blocking vaccines **(E, F)**. The distinct colors represent proportions of the GP emulator output variance (relative importance) attributable to intervention efficacy, half-life, deployment coverage, as well as health system access. Determinants of impact are shown for both immediate and late follow-up, when interventions are applied once per year for three years in different transmission settings (see full intervention specifications in the Methods section). The transmission settings are defined by two seasonal settings (seasonal and perennial) and three types of mosquito biting patterns (low, medium and high indoor biting). The mosquito biting patterns had little to no effect on the results of the sensitivity analysis for these therapeutic interventions

(see results for all settings for monoclonal antibodies in figures A and B). Therefore, only the results for seasonal and perennial settings with high indoor mosquito biting are displayed for the vaccine interventions.

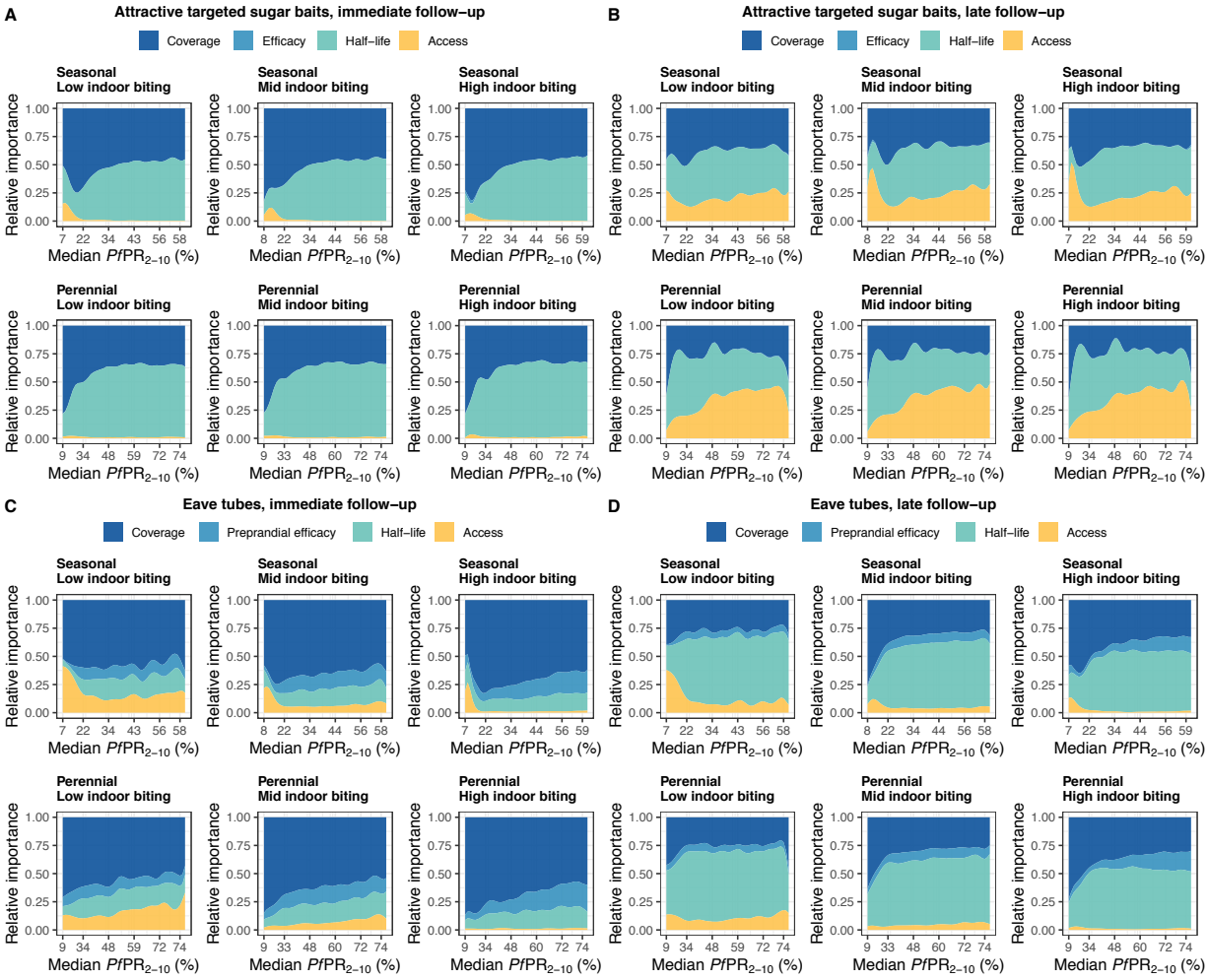


Figure S6.2.

Key drivers of impact for vector control malaria interventions across different transmission settings.

Results of sensitivity analysis identifying the determinants of intervention impact on *Pf*PR$_{0\text{-}99}$ reduction for attractive targeted sugar baits **(A, B)** and eave tubes **(C, D)**. The distinct colors represent proportions of the GP emulator output variance (relative importance) attributable to intervention efficacy, half-life, deployment coverage, as well as health system access. Determinants of impact are shown for both immediate and late follow-up, when interventions are applied once per year for three years in different transmission settings (see full intervention specifications in the Methods section). The transmission settings are defined by two seasonal settings (seasonal and perennial) and three types of mosquito biting patterns (low, medium and high indoor biting). Like for the therapeutic interventions in the previous figure, we see limited difference between key drivers for attractive targeted sugar baits in different biting settings as mosquitoes sugar feed before indoor or outdoor biting. In contrast, we observe that intervention properties of eave tubes rather than health system access to treatment are larger drivers of impact in indoor biting settings, as mosquitoes in those settings will be more likely to contact the eave tube.

7 Results: Feasible landscapes of optimal, constrained intervention profiles

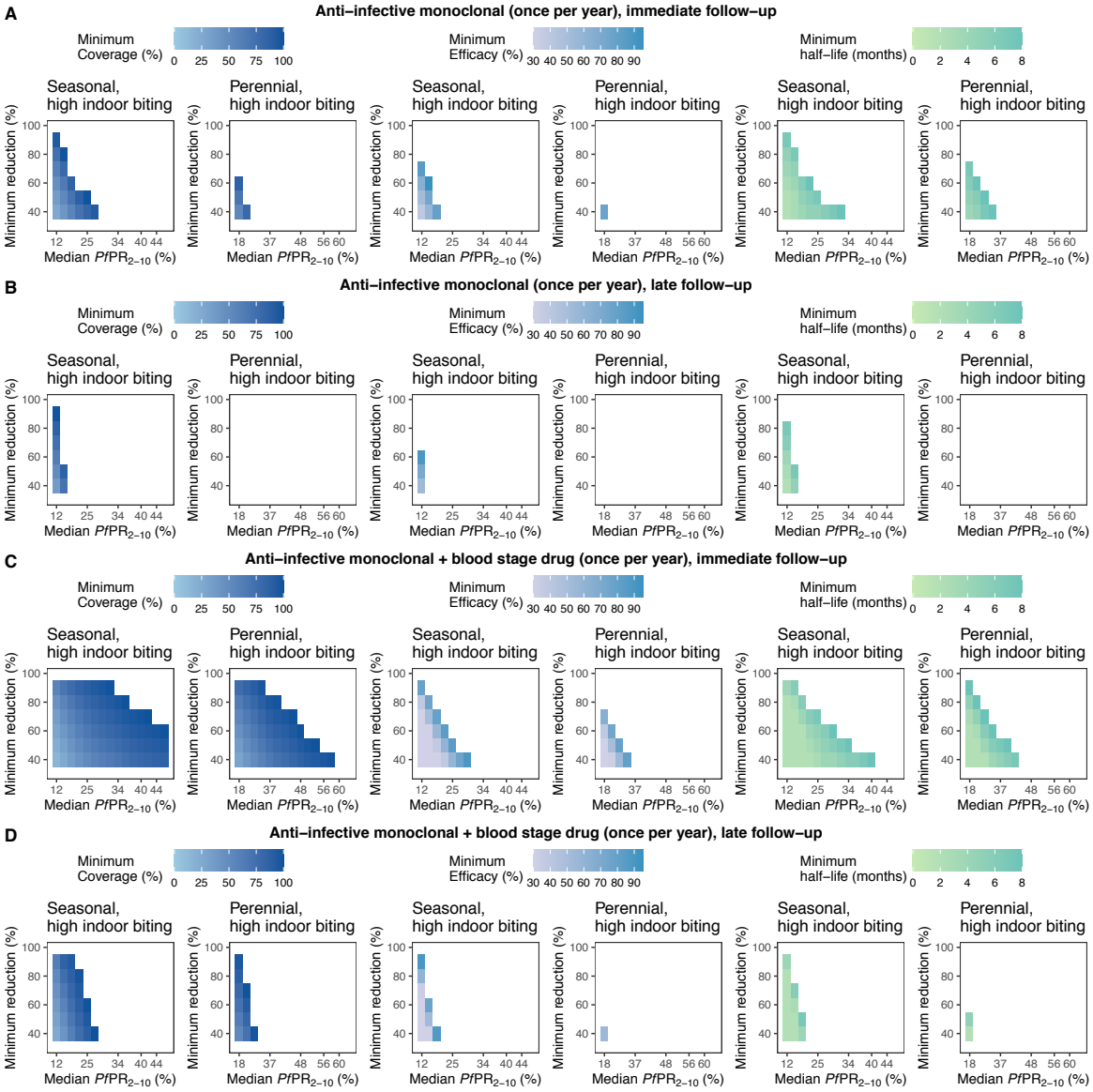


Figure S7.1.

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed once per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in $Pf\mathrm{PR}_{0\text{-}99}$, y axis) across different simulated true $Pf\mathrm{PR}_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for an anti-infective monoclonal antibody delivered alone and assessing immediate **(A)** and late **(B)** follow up, as well as when delivered in combination with a blood stage drug assessing immediate **(C)** and late **(D)** follow-up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

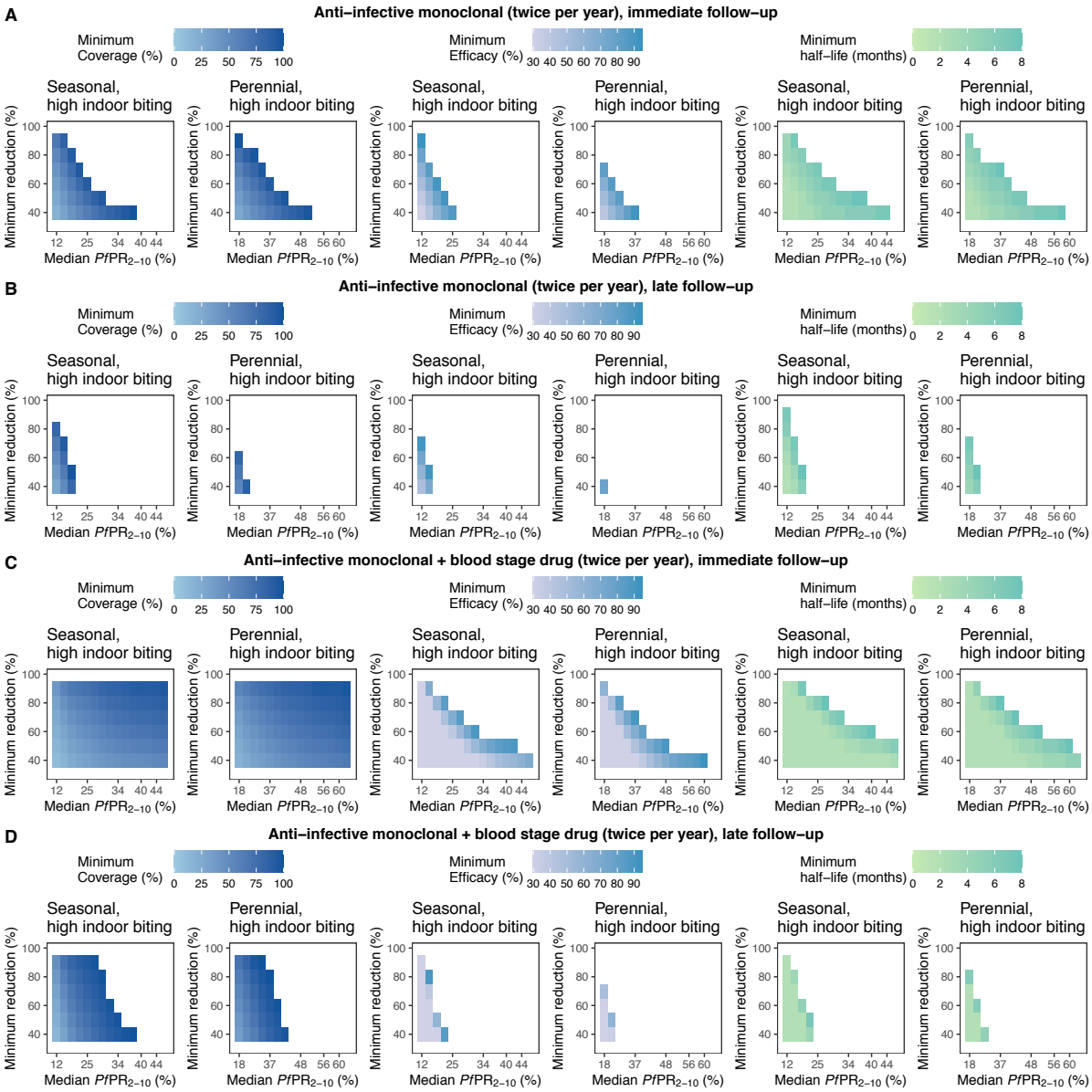


Figure S7.2.

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective monoclonal antibody deployed twice per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in *Pf*PR$_{0\text{-}99}$, y axis) across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for an anti-infective monoclonal antibody delivered alone and assessing immediate **(A)** and late **(B)** follow up, as well as when delivered in combination with a blood stage drug assessing immediate **(C)** and late **(D)** follow-up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

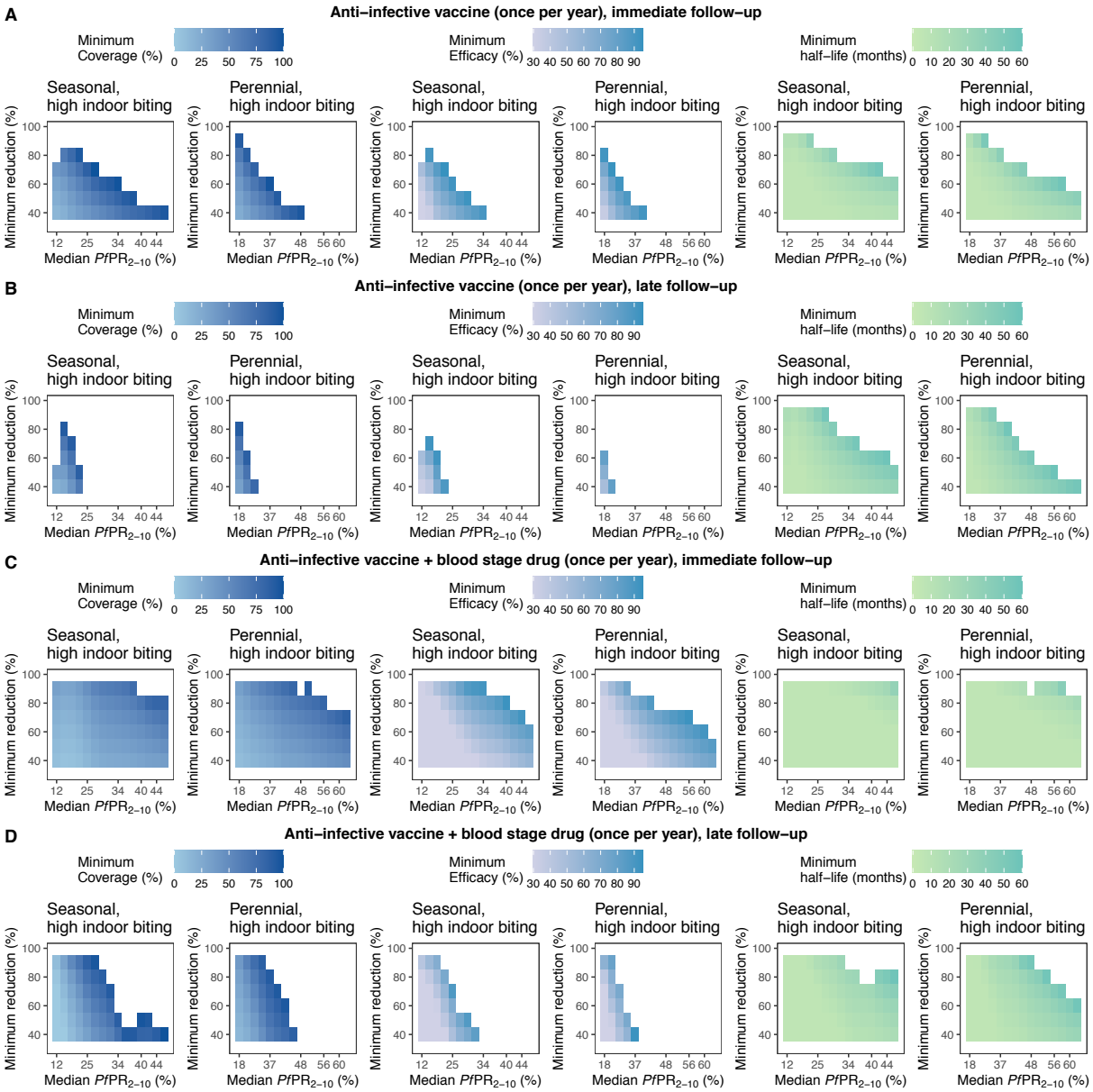


Figure S7.3

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for an anti-infective vaccine deployed once per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in *Pf*PR$_{0\text{-}99}$, y axis) across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for an anti-infective vaccine delivered alone and assessing immediate **(A)** and late **(B)** follow up, as well as when delivered in combination with a blood stage drug assessing immediate **(C)** and late **(D)** follow-up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

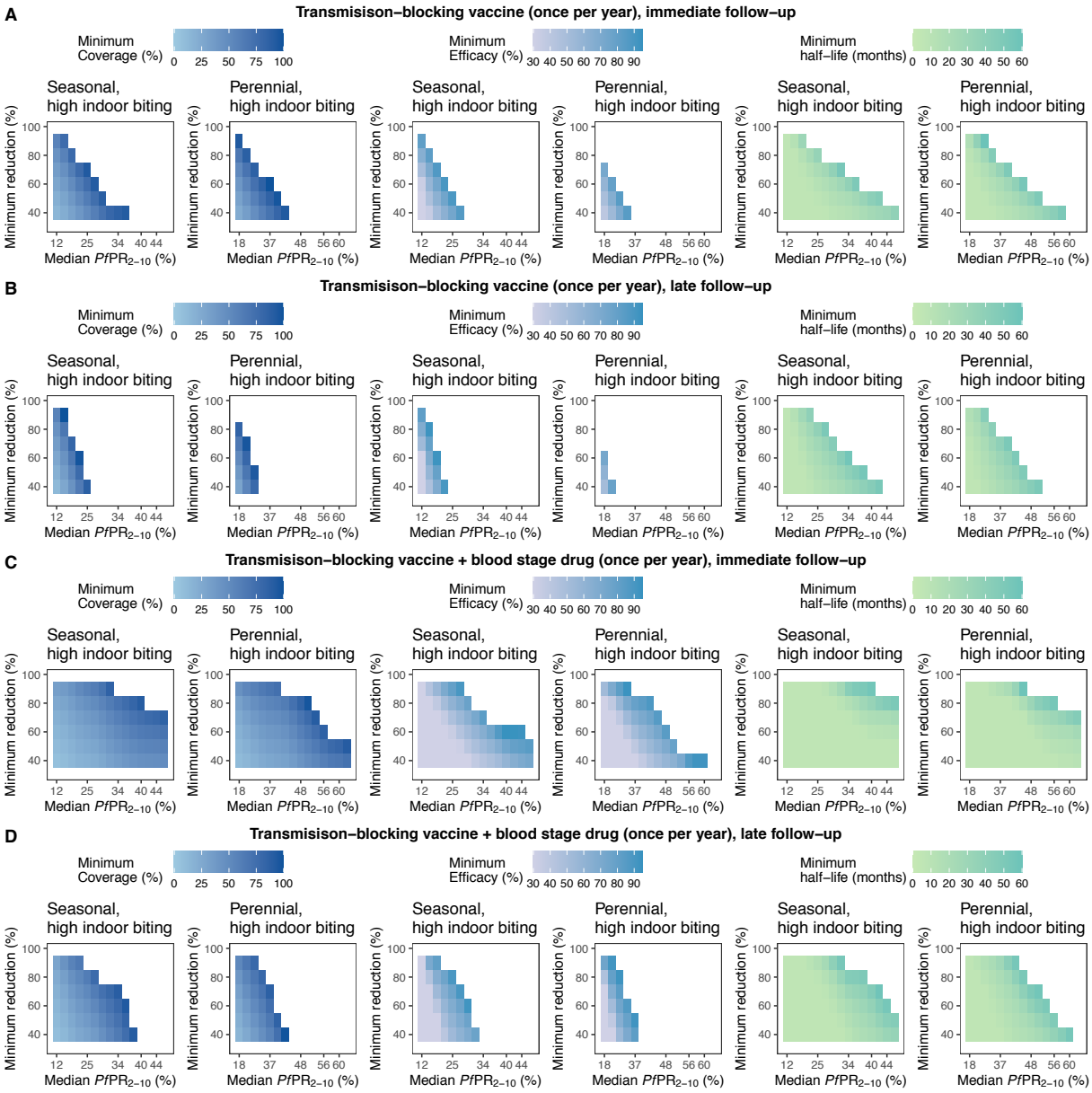


Figure S7.4.

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for a transmission-blocking vaccine deployed once per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in $Pf\mathrm{PR}_{0\text{-}99}$, y axis) across different simulated true $Pf\mathrm{PR}_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for a transmission-blocking vaccine delivered alone and assessing immediate **(A)** and late **(B)** follow up, as well as when delivered in combination with a blood stage drug assessing immediate **(C)** and late **(D)** follow-up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

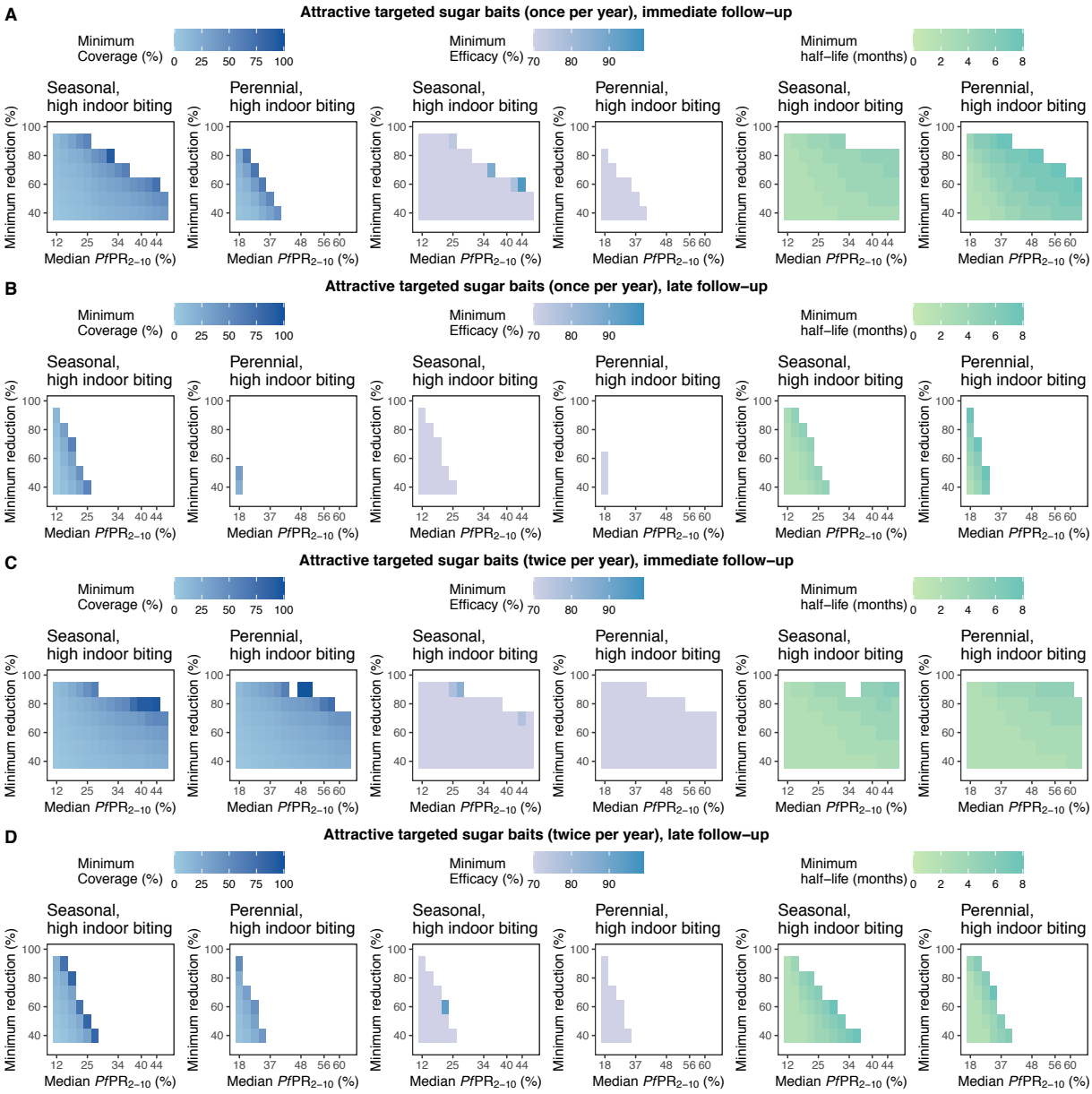


Figure S7.5.

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for attractive targeted sugar baits deployed once or twice per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in *Pf*PR$_{0\text{-}99}$, y axis) across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for attractive targeted sugar baits delivered alone once per year and assessing immediate **(A)** and late **(B)** follow up, as well as when delivered twice per year assessing immediate **(C)** and late **(D)** follow-up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

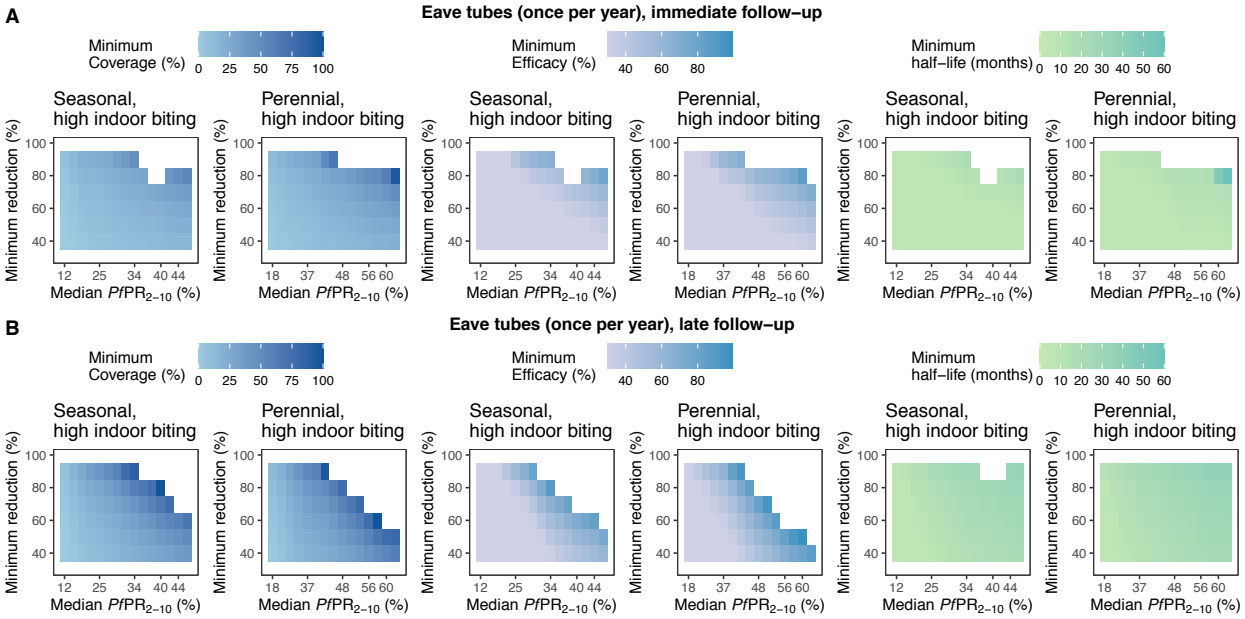


Figure S7.6.

Feasible landscapes of optimal, constrained intervention profiles (TPPs) for eave tubes deployed once per year.

The heatmaps represent landscapes of optimal, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life) required to achieve various health goals (quantified by minimal reduction in $Pf\mathrm{PR}_{0\text{-}99}$, y axis) across different simulated true $Pf\mathrm{PR}_{2\text{-}10}$ settings (rounded values, x axis) with seasonal or perennial transmission and high indoor mosquito biting (results for other biting patterns not shown as they are similar). Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown for eave tubes delivered alone and assessing immediate **(A)** and late **(B)** follow up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

8 Results: Optimal intervention profiles

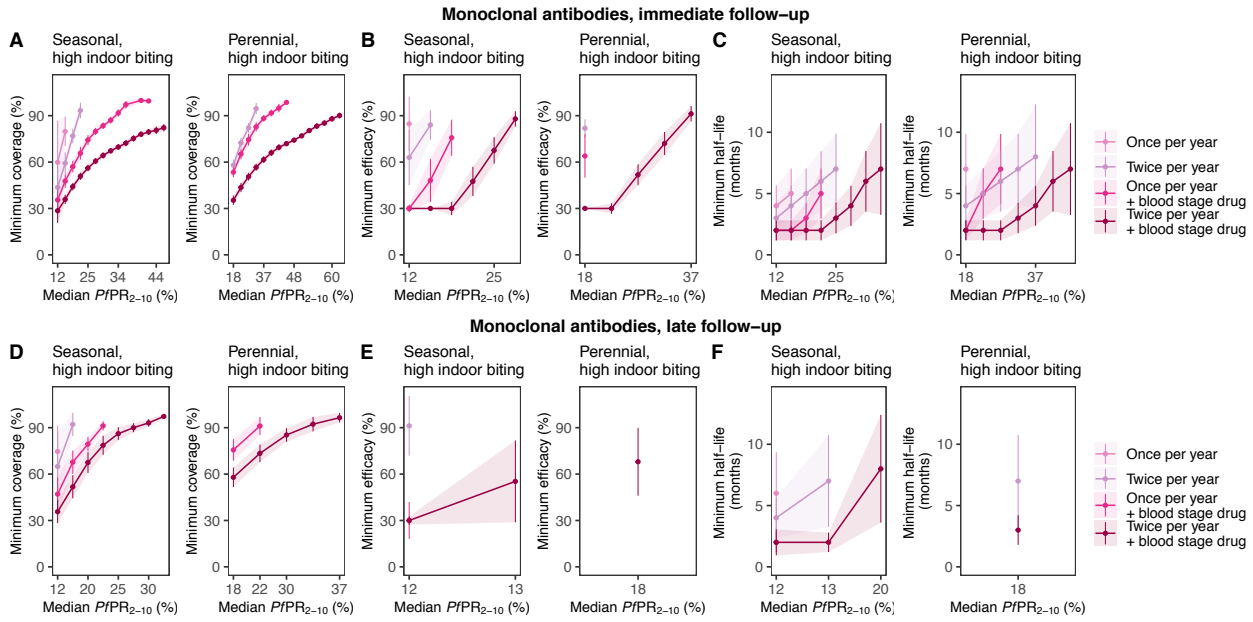


Figure S8.1

Optimal intervention profiles (TPPs) for anti-infective monoclonal antibodies under various deployment regimes to achieve a *Pf*PR$_{0\text{-}99}$ reduction of at least 70%.

Each figure displays minimum, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life, y axis) required to achieve a minimal reduction in *Pf*PR$_{0\text{-}99}$ of 70% across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown when assessing *Pf*PR$_{0\text{-}99}$ reduction at immediate **(A - C)** and late **(D - F)** follow up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

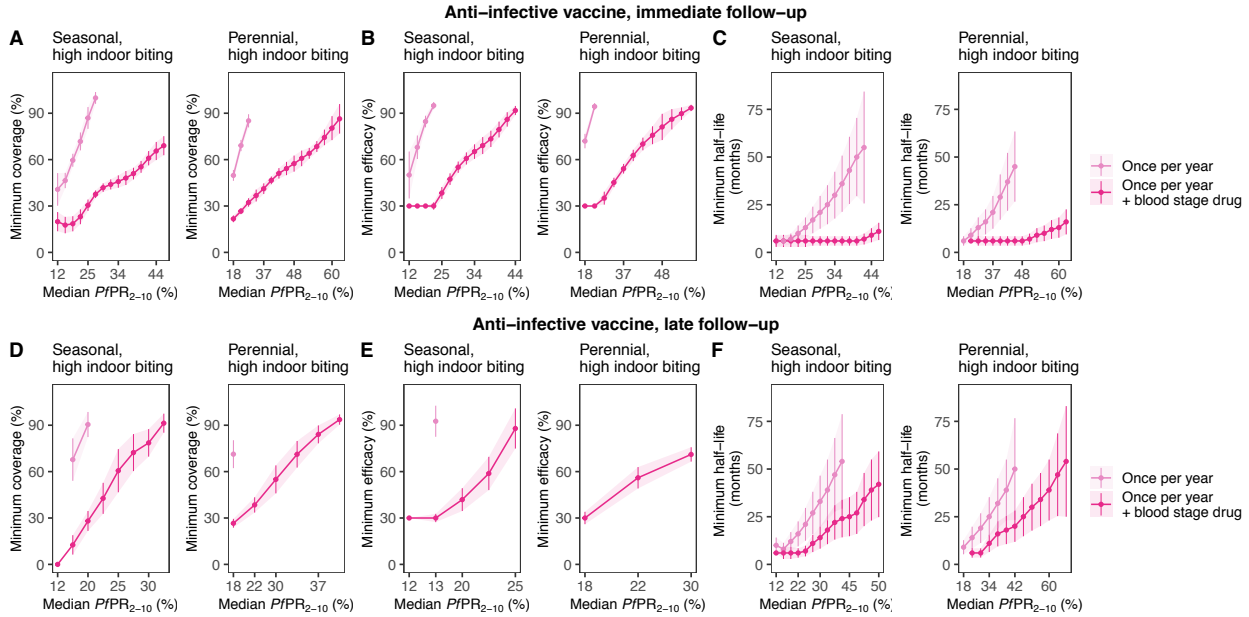


Figure S8.2

Optimal intervention profiles (TPPs) for anti-infective vaccines under various deployment regimes to achieve a *Pf*PR$_{0\text{-}99}$ reduction of at least 70%.

Each figure displays minimum, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life, y axis) required to achieve a minimal reduction in *Pf*PR$_{0\text{-}99}$ of 70% across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown when assessing *Pf*PR$_{0\text{-}99}$ reduction at immediate **(A - C)** and late **(D - F)** follow up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

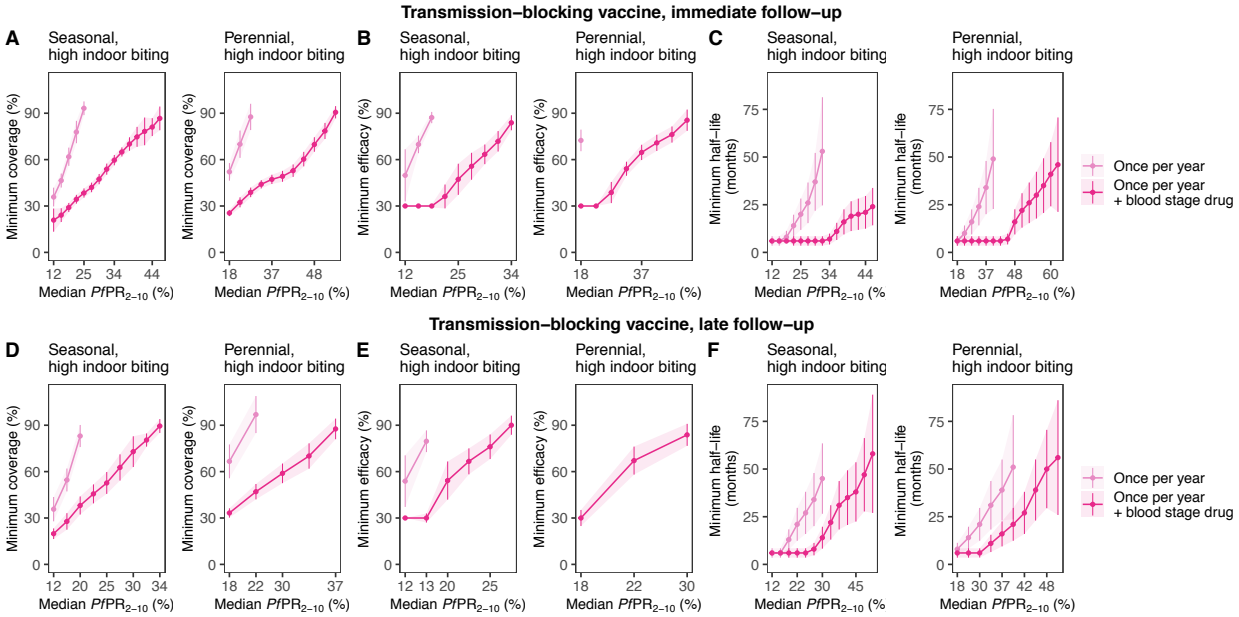


Figure S8.3

Optimal intervention profiles (TPPs) for transmission-blocking vaccines under various deployment regimes to achieve a *Pf*PR$_{0\text{-}99}$ reduction of at least 70%.

Each figure displays minimum, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life, y axis) required to achieve a minimal reduction in *Pf*PR$_{0\text{-}99}$ of 70% across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown when assessing *Pf*PR$_{0\text{-}99}$ reduction at immediate **(A - C)** and late **(D - F)** follow up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.

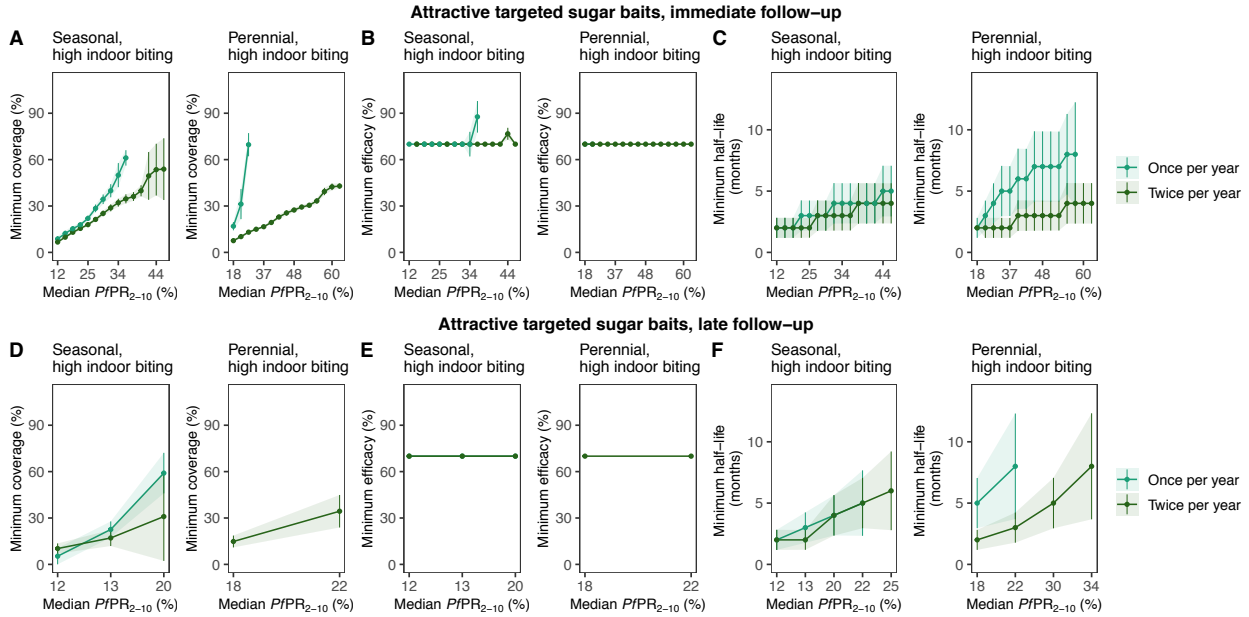
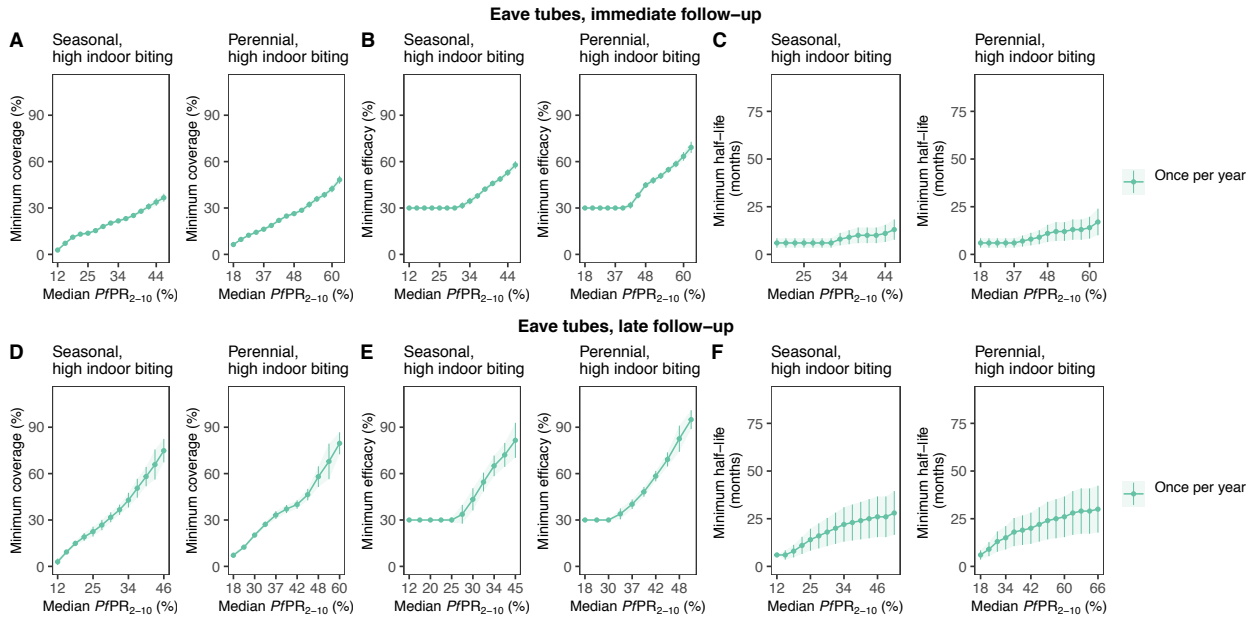


Figure S8.4

Optimal intervention profiles (TPPs) for attractive targeted sugar baits under various deployment regimes to achieve a *Pf*PR$_{0\text{-}99}$ reduction of at least 70%.

Each figure displays minimum, constrained intervention characteristic profiles (minimum coverage, efficacy, and half-life, y axis) required to achieve a minimal reduction in *Pf*PR$_{0\text{-}99}$ of 70% across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal transmission and high indoor mosquito biting. Each intervention characteristic was minimized in turn, while keeping the other characteristics fixed (fixed parameter values for each optimization are specified in Table S2.2). Results are shown when assessing *Pf*PR$_{0\text{-}99}$ reduction at immediate **(A - C)** and late **(D - F)** follow up. The simulated case management level (E_5) for all the displayed optimization analyses was assumed 25%.



1595

1596 **Figure S8.5**

1597 **Optimal intervention profiles (TPPs) for eave tubes to achieve a *Pf*PR$_{0\text{-}99}$ reduction of at**
1598 **least 70%.**

1599 Each figure displays minimum, constrained intervention characteristic profiles (minimum
1600 coverage, efficacy, and half-life, y axis) required to achieve a minimal reduction in *Pf*PR$_{0\text{-}99}$ of
1601 70% across different simulated true *Pf*PR$_{2\text{-}10}$ settings (rounded values, x axis) with seasonal
1602 transmission and high indoor mosquito biting. Each intervention characteristic was minimized in
1603 turn, while keeping the other characteristics fixed (fixed parameter values for each optimization
1604 are specified in Table S2.2). Results are shown when assessing *Pf*PR$_{0\text{-}99}$ reduction at immediate
1605 **(A - C)** and late **(D - F)** follow up. The simulated case management level (E_5) for all the displayed
1606 optimization analyses was assumed 25%.

1607

1608