Decreasing measles burden by optimizing campaign timing

Niket Thakkar\textsuperscript{a,1}, Syed Saqlain Ahmad Gilani\textsuperscript{b}, Quamrul Hasan\textsuperscript{c}, and Kevin A. McCarthy\textsuperscript{a}

\textsuperscript{a}Institute for Disease Modeling, Bellevue, WA 98005; \textsuperscript{b}Ministry of National Health Services, Regulations, and Coordination, Islamabad, Pakistan; and \textsuperscript{c}National Institute of Health, Islamabad, Pakistan

Edited by Burton H. Singer, University of Florida, Gainesville, FL, and approved April 19, 2019 (received for review October 25, 2018)

Measles remains a major contributor to preventable child mortality, and bridging gaps in measles immunity is a fundamental challenge to global health. In high-burden settings, mass vaccination campaigns are conducted to increase access to vaccine and address this issue. Ensuring that campaigns are optimally effective is a crucial step toward measles elimination; however, the relationship between campaign impact and disease dynamics is poorly understood. Here, we study measles in Pakistan, and we demonstrate that campaign timing can be tuned to optimally interact with local transmission seasonality and recent incidence history. We develop a mechanistic modeling approach to optimize timing in general high-burden settings, and we find that in Pakistan, hundreds of thousands of infections can be averted with no change in campaign cost.

Significance

Measles vaccine is a highly effective healthcare intervention, but getting vaccine to those in need remains a major problem. Complicating the issue, high-burden countries typically have low-quality infrastructure, severely limiting the number of infections detected and therefore limiting our understanding of local epidemiology. Here we show that statistical disease models can be fitted to sparse case data from Pakistan using a fast linear regression approach. This method yields estimates of the effects of past interventions, the seasonal likelihood of measles transmission, and the magnitude of future outbreaks under different intervention policies. We use these models to understand in general when and where vaccine should be distributed, and these results were used to inform Pakistan’s 2018 vaccination campaign planning.

Author contributions: N.T., S.S.A.G., Q.H., and K.A.M. designed research; N.T. performed research; N.T. analyzed data; N.T., Q.H., and K.A.M. wrote the paper; and S.S.A.G. and Q.H. collected data.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The code used in this paper has been deposited in the GitHub repository, https://github.com/NThakkar-IDM/campaign_timing.

\textsuperscript{1}To whom correspondence should be addressed. Email: nikhakkar@idmod.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1818433116/-/DCSupplemental.

Published online May 13, 2019.
infection prevalence at time $t$, as a discrete, stochastic dynamical system,

$$S_t = (1 - \mu_{t-1}) (B_t + S_{t-1} - I_t)$$  \hspace{1cm} [1]$$

$$I_t = \beta_t I_{t-1} S_{t-1} \varepsilon_t$$  \hspace{1cm} [2]$$

$$C_t \sim \text{Binom} \{I_t, p\}.$$  \hspace{1cm} [3]$$

Here, $B_t$ is an assumed known estimate of births missed by RI at time $t$, $\mu_t$ is the fraction of the susceptible population reached by any SIA at time $t$, $\alpha$ models inhomogeneous population mixing (5), and $\beta_t$ is the average number of infectious contacts per person at time $t$ which we assume has an annual periodicity. Transmission uncertainty is accounted for by $\varepsilon_t$, a zero-mean, log-normal random process, and laboratory-confirmed cases, $C_t$, are assumed to be drawn from a binomial distribution where $p$ is the laboratory-reporting rate, an unknown probability for cases to be selected for laboratory study. In Eq. 1 children missed by RI, $B_t$, contribute to $S_t$ while infections and SIAs serve to decrease $S_t$. Simultaneously, Eq. 2 models new infections occurring at rate $\beta_t$ as infectious and susceptible populations interact.

Since measles SIAs happen relatively infrequently, Pakistan’s campaign history can be used to reduce $\mu_t$ to the estimation of a single parameter. Subnational vaccination campaigns have been conducted six times in Pakistan since 2012 with wide variation in target population (17). Here we assume that nonzero $\mu_t = P_t \mu$ where $P_t$ is the known target population fraction and $\mu$ is an unknown SIA efficacy parameter common to all campaigns from 2012 to 2017.

Given the observed $C_t$ series and corresponding $B_t$ (via RI coverage estimates (6) and birth-rate estimates (18–20)) (Methods), the model can be fitted to data in a two-step linear-regression process described in Methods and SI Appendix, section 2. Model calibration yields estimates of $\alpha$, $\mu$, $p$, and $\beta$ with uncertainty due both to underreporting and to transmission stochasticity.

Fitting the model to national-level reports yields $p = 0.23 \pm 0.04\%$, indicating, in qualitative agreement with similar estimates from high-burden settings (21), that a single laboratory-confirmed case corresponds on average to ~400 infections in the population. Simultaneously, we find $\alpha = 0.93 \pm 0.03$, indicating that inhomogeneous population mixing is a small but statistically significant effect. Past SIA efficacy $\mu$ is estimated to be 40%, which shows that campaign efforts have had a significant effect on measles susceptibility in Pakistan.

In Fig. 1, national-level reports from 2012 to 2017 are aggregated by month (gray bars), showing that the majority of measles cases occur in the first half of the year. The inferred $\beta_t$ consistent with this case distribution is averaged by month and overlaid in red (SD cloud), showing that low transmission occurs between May and October (blue), Pakistan’s hot, summer rainy season. This correlation between measles transmission and rainfall or temperature agrees with findings from research in other settings (11, 22) and suggests that transmission fluctuations in part due to weather-related variation in contact rates.

Seasonal population migration and associated changes in population density have also been correlated with measles incidence in urban settings (23, 24). We test this hypothesis in SI Appendix, section 2 by computing annual variation in nighttime light satellite imagery brightness (25) near Pakistan’s largest cities, Karachi and Lahore. As shown in SI Appendix, Fig. S6, our inferred $\beta_t$ is highly correlated (Pearson correlation 0.725) with this measure of Pakistan’s urban population density, suggesting that annual rural-to-urban migration is also a driver of Pakistan’s transmission seasonality.

Interestingly, the increase in transmission precedes the rise in cases by 2–3 mo. This phase difference is in quantitative agreement with seasonality studies of measles in the preelimination United States (26), suggesting that although a measles infection’s duration is only 2–3 wk, high transmission is required for a considerable time before enough infections have occurred to spark an outbreak. Operationally speaking, this is a valuable insight since lows in the aggregated case count alone might incorrectly suggest that Pakistan’s low measles transmission season ranges from July to November.

Model seasonality and corresponding extrapolation ability are tested against laboratory reporting-rate scaled cases (black dots) in Fig. 2. In red, predicted $I_t$, given $C_{t-1}$ shows that the model is capable of reliable semimonthly prediction ($R^2 \approx 0.89$) with relatively low uncertainty (red cloud, 95% CI). A more substantial test of the model is shown in black, where $I_t$ is predicted for a full 6 y starting with $C_0$ in January 2012. This long-term model prediction has larger uncertainty (gray cloud, 95% CI) as expected and captures the major outbreaks in 2013, 2016, and 2017 ($R^2 \approx 0.35$), demonstrating that the inferred seasonality is consistent with the observed dynamics. The corresponding inferred $S_t$ is plotted in blue, showing stark decreases in susceptible population following SIAs (gray dashed lines) with heterogeneity between SIAs due largely to differences in target population. Finally, in SI Appendix, section 3, we demonstrate that when data past March 2017 are withheld from model fitting, out-of-sample model extrapolation successfully predicts the severity and timing of the 2017 outbreak.

**Optimal SIA Timing**

An effective vaccination campaign immunizes susceptible individuals to stifle measles transmission before it occurs. SIAs
Fig. 2. Testing model performance. (Lower) Semimonthly (red) and 6-yr (black) model extrapolations are compared with laboratory reporting-rate scaled cases demonstrating that the model predicts outbreak timing and magnitude. (Upper) The underlying susceptible population (blue) corresponding to the long-term projection highlights the potentially strong effect of SIAs (gray dashed lines). For all traces, shaded regions indicate 95% CIs.

achieve this in the model by decreasing both $S_t$ in Eq. 1 and the resulting $I_t$ in Eq. 2. Intuitively, based on the seasonality of Fig. 1, we expect that SIAs in Pakistan will have greatest impact in October or November since the susceptible population built up over the summer low-transmission season can be immunized before high transmission begins. Using the model, we demonstrate that this intuition is qualitatively correct, but a given population’s recent measles history also affects optimal SIA timing.

Hypothetical SIA policies can be quantitatively compared by calculating projected infections. Here, we focus on SIAs run in 2018 over the course of a full month with half the population targeted in each semimonthly model period, and we compute the sample distribution of total infections from 10,000 model runs starting with the data at the end of 2017 and forecasting for 3 yr. The 2018–2021 forecasting period was selected since, in practice, multiple SIAs will be run in >3-yr periods, and we are interested in comparing effects of single SIAs for simplicity. All hypothetical campaigns have efficacy equal to the inferred 2012–2017 efficacy, $\mu = 40\%$, to isolate the effects of SIA timing.

Expected infections under hypothetical 2018 SIA policies are plotted in black in Fig. 3A. As anticipated based on the seasonality, a November SIA has greatest impact, with ~440,000 fewer infections on average than an otherwise equivalent campaign run in January. Moreover, if the extra 10 mo to prepare leads to increases in SIA efficacy, we find that November rapidly becomes even more strongly favored (SI Appendix, Fig. S6). Throughout the low-transmission season (shaded blue region), campaigns become more and more effective. This is as we would expect since susceptible population buildup results in a wider-reaching campaign with greater herd-immunity effects (SI Appendix, section 5).

As a direct consequence of this, however, delays past November rapidly incur large costs since the 2018–2019 high-transmission season depletes the susceptible population and mitigates the effect of an SIA. This is demonstrated in Fig. 3A by extending the analysis to equivalent campaigns in 2019. Expected infections under these policies are plotted in red, and we find that a campaign delayed from November 2018 to May 2019 results in over 600,000 more measles infections on average over the 2018–2021 period.

Fig. 3B plots extrapolated model traces for SIAs before (in April, blue) and after (in November, green) the 2018 low-transmission season for more detailed comparison. While the April SIA mitigates infections in 2018, this comes at the expense of a large outbreak in 2020. On the other hand, the November SIA decreases the severity of the predicted 2020 outbreak at the expense of infections in 2018. This tradeoff indicates that transmission seasonality’s contribution to the optimal SIA timing acts in concert with the expected severity of upcoming outbreaks, an expectation which depends directly on measles’ recent history in a population. For Pakistan as a whole, 2017 was a relatively severe measles year, indicating that natural infection has decreased the susceptible population. Consistent with this intuition, model extrapolation predicts that 2020’s outbreak will be larger on average than 2018’s, and the November SIA is preferable as a result.

The interplay between seasonality and recent history is highlighted if we apply the model to Pakistan’s provinces individually. To do this, the model is fitted to province-level data assuming the national-level transmission seasonality of Fig. 1 with a contact rate scaled by the fraction of Pakistan’s population within the province. The assumption that measles transmission behaves qualitatively similarly across the country is necessary since individual provinces report too few laboratory-confirmed cases to reliably infer province-level transmission parameters. Province-level models are tested by the methods of Fig. 2 in SI Appendix, section 4. They show comparable predictive performance to the national-level model, indicating that the seasonality assumptions are valid.

Subnationally, Pakistan’s recent measles history has significant heterogeneity. For example, in Pakistan’s two most populated provinces, Punjab and Sindh, laboratory-confirmed measles cases per 100,000 in 2017 were at 0.9 and 6.8, respectively. While this is due in part to RI coverage differences between Punjab and Sindh (6), this also indicates that 2017 was an outbreak year in Sindh but not in Punjab. This heterogeneity is mirrored in province-level optimal SIA timing: Comparing April and November SIAs where data are available, we see in Fig. 3C that in provinces with high 2017 case counts the November campaign is more effective (purple) while in Punjab the April SIA performs better (red). Thus, in line with intuition
from the national level, optimizing SIA timing requires a balance between contributions due to seasonality and incidence history. The modeling approach presented here offers a robust means to solve this optimization problem in high-burden contexts.

Discussion

Measles vaccination campaign optimization is a complex general problem. Here, we have studied data from Pakistan to demonstrate that SIA timing is a critical factor and that two SIAs with equivalent efficacy and cost may have significantly different impacts solely as a result of their start date. With that in mind, transmission seasonality and recent measles burden, the drivers of optimal campaign timing, should be considered alongside operational constraints in future SIA planning.

From a methodological perspective, the TSIR model used in this work is a robust tool for evaluating competing SIA policies. While disease models with mass vaccination have been studied in the past (27–29), generalization of a least-squares–based model calibration method (5) to the high-burden context offers a simple, data-driven SIA optimization approach. Model extensions such as age structure (30), subnational spatial correlation (10), and disease importation (7) are active areas of research. These studies, in conjunction with the method presented here, may contribute to other aspects of SIA optimization, an important problem for measles eradication with widespread global health implications.

Methods

Pakistan Demographic and Surveillance Data. Population estimates for 2010 and 2015 and live-birth estimates for 2010, 2012, 2015, and 2020 were obtained from WorldPop (18–20). These were aggregated to the district level and linearly interpolated over time. Rates for the first dose of measles vaccine were estimated using the 2012–2013 DHS (6) and treated as constant over the 2012–2017 model period, while second-dose rates were estimated using caregiver-reported dose histories associated with laboratory-rejected cases.

Laboratory-confirmed and -rejected cases were obtained from Pakistan’s WHO-affiliated laboratory. The rejected cases and corresponding self-reported dose histories were used to estimate rates of second-dose measles vaccine coverage in all provinces. Combining these estimates of demographic quantities gives

$$B_t = \hat{B}_t [1 - 0.9V_{1,t}(1 - V_{2,t}) - 0.99V_{1,t}V_{2,t}],$$

where $V_{1,t}$ and $V_{2,t}$ are first- and second-dose measles vaccine coverage over time, and $\hat{B}_t$ is the estimated live births. The above model assumes the first vaccine dose has a 90% seroconversion rate and the second dose has a 99% seroconversion rate (16). For more details, see SI Appendix, section 1.

Fitting and Testing the Model. Model fitting to an observed $R_t$ series proceeds in two steps, accounting for uncertainty due to underreporting and transmission individually. In the first step, Eq. 1 is used to construct a weighted least-squares regression of $B_t$ against $R_t$ which yields, for a given $\mu_t$, estimates of $p$ and the relative fluctuations in the susceptible population. Further assuming that susceptible fluctuations are small and $\beta_t = \beta_{\text{mod24}}$, i.e., that seasonality varies only within a year, Eq. 2 defines a generalized linear autoregression of $R_t$. Solving this regression problem yields estimates of

Fig. 3. Optimizing SIA timing in Pakistan. (A) Comparing total expected infections in 2018–2021 (black, SE shading) under different SIA policies shows that November minimizes measles burden by taking advantage of susceptible buildup over the low-transmission season (blue region). As a result however, delays into the 2018–2019 high-transmission season (red, SE shading) are costly. (B) Model projections for pre–(April) and post–low-transmission season (November) SIAs (black dashed lines) demonstrate the tradeoff between 2018 and 2020 outbreak control. As a result, 2017 measles burden also plays a significant role in timing optimization. (C) Extending the model to the province level allows us to compare April and November SIA timing subnationally. Preference for April is mapped in red while preference for November is mapped in purple; gray provinces [Federally Administered Tribal Areas and Azad Kashmir, representing less than 5% of Pakistan’s total population (20)] are inaccessible to health workers while white areas indicate disputed territory. Heterogeneity in the 2017 laboratory-confirmed measles cases per 100,000 (indicated) is reflected in the timing optimization.
\( \beta \) and the remaining parameters including the variance due to transmission uncertainty.

As mentioned in the main text, we assume \( \mu = \gamma P_i \), where \( P_i \) is a known measure of target SIA population in the past and \( \mu \) is an efficacy parameter common to all SIAs from 2012 to 2017. Since the regression approach above can be carried out given a hypothetical \( \mu \), we take an approach similar to the profile-like optimization used by others (14, 15). In other words, a range of \( \mu \) are tested by repeated model fitting and subsequent goodness-of-fit optimization. For mathematical details of the full model calibration procedure and related sensitivity testing see SI Appendix, section 2.

Code and Data Availability. All analysis was done in Python 3.6.2, and the associated code can be found in the GitHub repository, https://github.com/NThakkar-IDM/campaign_timing (31). All data came from open-source providers noted in the references with the exception of the laboratory reports which can be obtained only with permission from the World Health Organization Country Office in Pakistan. To obtain permission, contact the corresponding author (N.T.) or submit a request to the WHO directly (details can be found in the WHO data policy, https://www.who.int/publishing/datapolicy/en/).

ACKNOWLEDGMENTS. We thank Pakistan’s Expanded Program on Immunization for providing the laboratory-reported data, Kurt Frey for useful conversations and significant help with compiling Pakistan’s past SIA calendar, Hil Lyons for assistance with computing birth rates from the DHS, and Caitlin Cornell for useful edits to the manuscript.