Epidemiological Determinants of Successful Vaccine Development

Hiroshi Nishiura\textsuperscript{1,2\textcopyright} and Kenji Mizumoto\textsuperscript{1,3}

1. School of Public Health, The University of Hong Kong, Level 6, Core F, Cyberport 3, 100 Cyberport Road, Pokfulam, Hong Kong SAR, China.
2. PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama, 332-0012 Japan.
3. Department of International Health, Nagasaki University Institute of Tropical Medicine and GCOE, 1-12-4 Sakamoto, Nagasaki, 852-8523 Japan.

\textcopyright{} Corresponding author: E-mail: nishiurah@gmail.com Tel: +852 3906 2010, Fax: +852 3520 1945.

Abstract

Epidemiological determinants of successful vaccine development were explored using measurable biological variables including antigenic stability and requirement of T-cell immunity. Employing a logistic regression model, we demonstrate that a high affinity with blood and immune cells and pathogen interactions (e.g. interference) would be the risk factors of failure for vaccine development.

Key words: Vaccination; epidemiology; infectious disease; immunization; statistical model.

Introduction

The estimated low effectiveness of malaria vaccine and dengue vaccine in the field [1,2] have disappointed a certain fraction of vaccine enthusiasts. Both vaccines offered some hope in advance of the field studies, but clinical protection in infants and other groups appears to be lower than the earlier expectation based on laboratory experiments and smaller-scale clinical studies among older individuals. As part of lessons to be learnt from these unfavorable outcomes (and from success in earlier vaccines), there is a growing need to clarify what risk factors determine successful and unsuccessful vaccines [3].

Understanding common biological features among the existing (recommended) childhood vaccines, in contrast to recent unsuccessful vaccines, would be useful for designing optimal avenues for future vaccine development. Here we report the results from a simple epidemiological analysis which aimed to identify virological, pathophysiological and eco-evolutionary factors that may lead to a successful vaccine.

Methods

There are 14 infectious diseases against which vaccination among children aged from 0-6 years is recommended in the USA [4], including chickenpox, diphtheria, Haemophilus influenzae infection, hepatitis A virus infection, hepatitis B virus infection, influenza, measles, mumps, pertussis, polio, pneumococcal disease, rotavirus infection, rubella, and tetanus. On the contrary, four notable diseases without a promising vaccine include HIV/AIDS, tuberculosis, malaria and dengue virus infection [3,5]. In many instances, the difficulty in developing vaccines against these four infectious diseases has been implicitly attributed to two biological factors: (i) broad antigenic diversity (except for tuberculosis) and (ii) the limited role of humoral immunity and the requirement of T-cell immunity for protection [5]. However, these factors have not been quantitatively measured in common biological scales.

The existence of two distinct groups of diseases (i.e. with and without successful vaccine) provides us with an opportunity to implement a case-control
analysis that helps identify epidemiological determinants of success in vaccine development. For each pathogen, we scale the abovementioned two factors by considering the following five dichotomous variables that are measurable in practice. As an explanatory variable that would partially reflect antigenic stability, one can examine if a disease is caused by (i) an RNA virus or protozoa that is known to experience rapid evolution. As an indirect measure of evolutionary pressure, (ii) the existence of asymptomatic infection could act as a complementary factor that potentially mirrors the antigenic diversity. That is, if an infection involves a substantial fraction of asymptomatic infections, the infection is unlikely to impose strong clinical selection pressure, and thus, we regard it as reflecting low pathogenicity (or low virulence) or the infection caused by an evolved pathogen. Third, (iii) a high affinity to cells in blood and immune systems is considered as a variable that reflects the limited protection by humoral immunity (e.g. *Mycobacterium tuberculosis* in phagocytic immune cells can escape from humoral immune response). Moreover, (iv) the transmission via a common route (i.e. respiratory or through direct contact) is also considered as inversely associated with the requirement of cellular immunity, because the establishment of infection via the common route involves many opportunities of humoral immune response. Fifth, (v) the existence of interactions between different pathogens, including enhancement within an identical pathogen species and epidemiological interference between different pathogens, is considered, because the interaction can be consistent with both broad antigenic diversity (due to a rich multistain evolutionary dynamics) [6] and also with the requirement of cell-mediated immunity (e.g. non-specific temporal immunity shortly after influenza virus infection is believed not to be humoral [7]).

Both the outcome variable (i.e. with or without successful vaccine) and explanatory variables were dealt with as dichotomous. To examine statistical association between two dichotomous variables, we employed Fisher’s exact test. Subsequently, to address potential confounding effect, a multiple logistic regression model was used to identify factors that are associated with our outcome variable. A forward stepwise method was employed using the minimum corrected Akaike Information Criterion (AICc) to choose the best model. All statistical data were analyzed using a statistical software JMP version 9.0.0 (SAS Institute Inc., Cary, NC, USA).

**Results**

A total of 18 samples of infectious diseases, including 14 diseases with successful vaccines that are routinely adopted as recommended ones and 4 diseases without a successful vaccine, were examined. Table 1 shows the results of univariate and multivariate analyses. In both analyses, only two factors, i.e., a high affinity to blood and immune cells and the existence of interactions between different pathogens were positively and significantly associated with diseases without a successful vaccine. Multivariate logistic model with step-wise method also identified these two variables to be left in the final model (Table 1). Adjusted odds ratios of diseases without successful vaccine for high affinity to blood/immune cells and interactions between pathogens were as large as $1.2 \times 10^{15}$ (95% confidence interval (CI): 9.8, $\infty$) and $4.9 \times 10^{14}$ (95% CI: 2.2, $\infty$), respectively. Neither RNA-virus/protozoa nor the existence of asymptomatic transmission appeared to be a significant predictor of infectious diseases without a successful vaccine ($p=0.577$ and 0.082, respectively).

**Table 1. Epidemiological determinants of diseases with successful vaccine (n = 18).**

<table>
<thead>
<tr>
<th>Possible determinants</th>
<th>Diseases with successful vaccine (n=14)</th>
<th>Diseases with unsuccessful vaccine (n=4)</th>
<th>p-Value*</th>
<th>Odds ratio (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA virus or protozoa</td>
<td>6 (42.9%)</td>
<td>3 (75.0%)</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>Existence of asymptomatic infection</td>
<td>5 (35.7%)</td>
<td>4 (100.0%)</td>
<td>0.082</td>
<td></td>
</tr>
<tr>
<td>High affinity to cells in blood and immune systems</td>
<td>1 (7.1%)</td>
<td>4 (100.0%)</td>
<td>0.002*</td>
<td>$1.2 \times 10^{15}$ (9.8, $\infty$)</td>
</tr>
<tr>
<td>Transmission via respiratory route or non-sexual contact</td>
<td>10 (71.4%)</td>
<td>1 (25.0%)</td>
<td>0.245</td>
<td></td>
</tr>
<tr>
<td>Existence of interactions between pathogens including enhancement</td>
<td>3 (21.4%)</td>
<td>4 (100.0%)</td>
<td>0.011*</td>
<td>$4.9 \times 10^{14}$ (2.2, $\infty$)</td>
</tr>
</tbody>
</table>

*Two-tailed Fisher’s exact test. †Significant by univariate analysis. †Adjusted odds ratio of a specific factor that leads not to have successful vaccine. Parenthesis shows the lower and upper 95 percent confidence intervals. Forward stepwise logistic regression was employed using the minimum corrected Akaike Information Criterion (AICc) to choose the best model, and only two variables were left. Dependent nominal variable = disease without a successful vaccine, AICc=7.71, $\chi^2=5.60 \times 10^7$, $p=0.99$. 

http://www.medsci.org
Discussion

Although our epidemiological analysis has rested on crude analysis, and the finding is likely vulnerable to confounding effect, Table 1 critically shows that, rather than being an RNA virus or protozoa, other markers (i.e. high affinity with blood/immune cells and pathogen interactions) are more likely to predict diseases without a promising vaccine. Two practical implications should be noted. First, our epidemiological exercise has permitted us to identify a research gap in vaccine development. Namely, while studying antigenic evolution [8] and system biologic approach to vaccinology [3,5] is of utmost importance to address limited protective effect of recent vaccines, it is vital to understand the disease characteristics that prevent us to develop a promising vaccine [7,9-11]. For instance, to understand the mechanism of unsuccessful vaccine development in a rudimentary fashion, it is critical to clarify the reasons for an efficient escape of a pathogen from vaccine-induced immunity among those causing infection in blood and immune cells. Similarly, understanding eco-epidemiological mechanisms that underlie interactions between pathogens would play key roles in disentangling the complexity of the protective effect of vaccination in field studies (e.g. clarification on this point may help understand complex immunological response among vaccinated individuals in natural settings). Second, the abovementioned explanatory factors can identify diseases without a successful vaccine, and thus, allow us to objectively assess the feasibility of introducing a forthcoming future vaccine in advance of field studies. Such a predictive model can also be revised and updated by considering other (useful) explanatory factors and adding other diseases of public health interest to the samples.

Despite our identification of the potential epidemiological determinants of successful vaccine, the abovementioned model was only aimed to learn from the existing armament of childhood vaccines. As long as truly novel approaches (e.g. adjuvants, new regimes of immunization and novel antigens) are effectively used for developing a new vaccine, future vaccine development may be granted an opportunity to refute the simple expectation.

Acknowledgments

HN received funding support from the Japan Science and Technology Agency (JST) PRESTO program. This work also received financial support from the Harvard Center for Communicable Disease Dynamics from the National Institute of General Medical Sciences (grant no. U54 GM088558). The funding bodies were not involved in the collection, analysis and interpretation of data, the writing of the manuscript or the decision to submit for publication.

Competing Interests

The authors have declared that no competing interest exists.

References