Winston Churchill famously said: “I cannot forecast to you the action of Russia. It is a riddle, wrapped in a mystery, inside an enigma; but perhaps there is a key.” He could have been referring to forecasting influenza epidemics and the effectiveness of influenza vaccines. Currently available influenza vaccines have moderate efficacy, established by placebo-controlled clinical trials, and moderate effectiveness, demonstrated in observational studies [1, 2]. However, there is variation in vaccine effectiveness by age and vaccine type, as well as substantial year-to-year variation. Some of this variation is well understood. With current technology, vaccine strains are selected many months in advance, leading to periodic major antigenic mismatches between the strains used for vaccine and the predominant circulating virus. Such a mismatch occurred in the 2014–2015 influenza season between the influenza A(H3N2) virus strain in the vaccine (A/Texas/50/2012) and the strain that circulated in North America (A/Switzerland/9715293/2013-like), resulting in minimal vaccine effectiveness [3–5]. The need to express the antigens of the selected virus in egg-adapted seed strains can lead to unexpected low yields and delays in vaccine availability. Strain optimization and low yield contributed to the delay in production of vaccine to respond to the 2009 pandemic caused by influenza A(H1N1) virus (A[H1N1]pdm09) [6].

To monitor vaccine effectiveness, annual observational studies are conducted in the United States [7], Canada [8], and Europe [9]. These studies use the test-negative case-control study design. This method involves enrollment of patients with medically attended respiratory illness and testing them with sensitive and specific assays for influenza virus. The odds ratio is calculated from the rate of influenza vaccination among those with influenza, compared with those who have a negative test result. Multivariate logistic regression is used to control for potential confounders such as age, presence of high-risk conditions, and calendar time. Effectiveness is calculated as $[1 - \text{odds ratio}] \times 100$. When the sample size is large enough, effectiveness can be calculated for specific age groups, virus type, and vaccine type. Preliminary results can be available early in the influenza season, to help guide policy [4]. The design has been validated using data from randomized trials [2], but it is still susceptible to potential bias, particularly from differential testing of vaccinated and unvaccinated patients [10, 11].

Annual effectiveness studies have provided some insights. Correlating strain-specific data with effectiveness has helped explore how effectiveness is related to age and the degree of antigenic drift [7, 12]. Such studies have also provided surprises, including a suggestion that, in some settings, receipt of vaccine in the previous years was associated with lower effectiveness [13].

The article by Gaglani et al in this issue of *The Journal of Infectious Diseases* provides another result that could not have been forecasted [14]. The authors used test-negative case-control data from the US Flu Vaccine Effectiveness Network for the 2013–2014 season, when (A[H1N1]pdm09—essentially the same virus that caused the 2009 pandemic—was the predominant strain. They compared the effectiveness of inactivated influenza vaccine (IIV) and live attenuated influenza vaccine (LAIV). This analysis was timely because, in June 2013, the Advisory Committee on Immunization Practices (ACIP) recommended that LAIV be the preferred vaccine for healthy children aged 2–8 years. The ACIP based this preference on data from randomized trials [15] and observational studies [16] that showed higher relative efficacy and effectiveness of LAIV, compared with IIV, in children aged 2–8 years. Surprisingly, Gaglani et al found that, during the 2013–2014 season, quadrivalent LAIV was not effective in children (adjusted vaccine effectiveness, 17%; 95% confidence interval, 39% to −51%). In contrast, IIV was effective with an adjusted vaccine effectiveness of 60% (95% confidence interval, 36%–64%). A similar lack of effectiveness of LAIV was observed in a separate test-negative study among children, funded by the manufacturer [17], and in data from the Armed Forces Surveillance Network [18]. In contrast, LAIV was effective in children against influenza B virus infection, during 2013–2014, and against influenza A(H3N2) virus infection, during 2011–2012 and 2012–2013.

How can we explain the surprising failure of LAIV against influenza A(H1N1)
virus during 2013–2014? Gaglani et al suggest that the problem may lie in the thermostability of the A(H1N1)pdm09 vaccine construct. The hemagglutinin and neuraminidase genes for each year’s vaccine are expressed by reverse genetics in the cold-adapted vaccine strain [19], so the resultant virus is a 6:2 reassortant. The hemagglutinin protein from the original A(H1N1)pdm09 strains early in the pandemic contained glutamic acid at position 47 (E47) in the stalk, a critical locus in the fusion domain of hemagglutinin. Later circulating viruses contained lysine at this position (K47). Cotter et al demonstrated that the K47 virus was more infectious than the E47 virus in ferrets, probably reflecting further adaptation to humans. In addition, vaccine constructs with E47 were more heat labile, which would be predicted to affect stability and potency [20]. Indeed, data presented to the ACIP by the manufacturer suggested that lots shipped earlier in the year at higher ambient temperatures were associated with lower effectiveness than those shipped later in the season.

There are other possible explanations that cannot be excluded with the available data. Because the influenza A(H1N1) virus in 2013–2014 had circulated with little antigenic change and the H1N1 component of vaccines was unchanged since 2009, it is possible that immune mechanisms limited replication of LAIV in recipients, attenuating the immune response. Immune-mediated limitation of replication has been hypothesized to explain why LAIV is less effective than IIV in older adults [21]. Undetected problems during manufacturing, shipping, or storage could also have contributed to lower effectiveness.

The larger problem stems from our incomplete knowledge of correlates of natural and vaccine-induced immunity to influenza. Our understanding of natural immunity to influenza remains incomplete >80 years after the Robert Shope’s isolation of influenza virus. Antibody to hemagglutinin is used for regulatory approval but is, at best, an imperfect correlate of effectiveness for IIV. LAIV generally elicits less robust humoral responses but induces long-lived CD4+ T-cell responses, even in some subjects who do not have a humoral response [22, 23]. It is becoming clear that T-cell–mediated responses are critical in long-term and cross-strain protection against influenza [24].

Many have called for new influenza vaccines [1], but, in the meantime, careful studies will help us understand and better use existing vaccines. The characteristics of an optimal vaccine might include rapid and inexpensive production, broad protection against drifted and pandemic strains, and production of T-cell and B-cell memory. This goal is unlikely to be achieved by novel vaccine constructs alone. It will require a deeper understanding of the enigmas and mysteries of influenza pathogenesis and immunity.

Notes

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (grant 1R01AI04593-01) and BioFire Diagnostics.

Potential conflict of interest. A. T. P. has received grants from BioFire Diagnostics, personal fees and other compensation from Antimicrobial Therapy, and personal fees from WebMD, all outside the submitted work. The author has submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References