Clarifying Misconceptions Regarding Influenza Research,

GISAI D, and the Role of the US Centers for Disease Control

We read with interest the October 2, 2011 opinion piece by Steven Salzberg titled, *Is The Government Hiding Something About Swine Flu?*, which was published on various internet blogs by Dr. Salzberg. We fully support his right to express his opinions, and applaud his efforts to raise public awareness concerning influenza research and prevention. Unfortunately, however, Dr. Salzberg’s opinion piece contains certain fundamental misstatements and omissions that make his article dangerously misleading.

Dr. Salzberg’s contention appears to be that the CDC is intentionally withholding influenza genetic sequence data from others by submitting such information to a database (GISAID) that Dr. Salzberg alleges is non-public. Dr. Salzberg speculates that the CDC does this to prevent others from discovering and publishing important information contained in the data before the CDC can do so. Dr. Salzberg’s contentions are incorrect, and are based on fundamentally flawed premises.

As an initial matter, Dr. Salzberg does not accurately describe GISAID or its database. GISAID is an initiative comprised of scientists for scientists that is organized through a nonprofit foundation based in Washington, D.C., with the support of governments around the world. GISAID’s database was launched in 2008 to foster international sharing of avian and other influenza virus sequences and encourage collaborative research. Contrary to Dr. Salzberg’s assertion, GISAID’s database is freely and publicly accessible. It is overseen by a verifiable and unbiased mechanism of organizational bodies that operate independently of each other. Due to the very nature of the public-private partnership, the GISAID initiative enjoys wide ranging support, including from companies such as Oracle and Cisco Systems, as well as from governments such as the Federal Republic of Germany, which hosts GISAID’s EpiFlu™ database.

Many laboratories participating in WHO’s Global Influenza Surveillance and Response System—including CDC’s WHO Collaborating Center for Influenza Reference and Research—submit influenza data to GISAID because they, like many other influenza research laboratories, believe the GISAID database is the preeminent platform to make the data available to researchers and others around the world and to promote influenza research and prevention in a manner that is most likely to benefit the global community. Data submitted to GISAID undergoes a meticulous curation process to enhance data quality and consequently its scientific exploitation.

The GISAID database provides for a responsible sharing mechanism and adds to transparency by making its data accessible to anyone who positively identifies him or herself and accepts the GISAID EpiFlu™ Database Access Agreement (“DAA”). The DAA is intended to encourage the sharing of sequence and related data by requiring the user to acknowledge the data contributors (e.g., the originating laboratory or submitting laboratory), preventing the user from attaching restrictions to the data (e.g., by including it in a patent application), preventing the user from disclosing the data outside the GISAID community (e.g. to people who have not agreed to comply with the DAA) and requiring the user to make

---

1 http://www.cidrap.umn.edu/cidrap/content/influenza/panflu/news/jun2509gisaid-br.html
best efforts to collaborate with the originating laboratory. Prior to the emergence of GISAID, many countries were unwilling to contribute material to public archives such as GenBank because their contributions were not acknowledged by users of the data, even though the contributions of the material were indispensable to any use. The DAA also helps preserve access by the contributors of the data to the benefits derived from such data. Contributors (particularly those from developing countries whose populations may be at greatest risk from disease) require access to the benefits of technology derived from the data, such as vaccines and antiviral drugs. By requiring acknowledgment of the rights and interests of those providing data, as well as those analyzing and interpreting the data, GISAID has made an objective contribution to overcome the reluctance of some scientists, officials and governments to share information for the benefit of the global community. Thus, many scientists, laboratories, governmental entities and others have chosen to submit data to the GISAID database because, among other things, it embodies principles for responsible data sharing more recently developed for international genomics research and reaffirmed by funders of health research “to promote the efficient use of research data to accelerate improvement in public health.”

Dr. Salzberg’s assertion that GISAID engaged in a “bait and switch” and “sits on data” is patently incorrect. All data that is submitted to GISAID is immediately available, without exception, free of charge. For example, recent data contained in GISAID’s EpiFlu™ database is shared by labs of the GISRS (including the CDC) and is used to anticipate “what is coming in the next flu season”, and contribute to the basis for the semiannual WHO recommendations on the composition of human influenza vaccines (and candidate vaccine viruses for vaccine against potential zoonotic epidemics/pandemics). This mechanism (GISAID) contributes to optimizing the efficacy of the vaccines from which everyone will benefit, thereby assisting the CDC in fulfilling its public health mission.

Dr. Salzberg also appears to have a fundamental misunderstanding of the relative roles and responsibilities of the various human and animal health authorities. The World Health Organization and its network of WHO Collaborating Centers for Influenza Reference and Research (located in Atlanta, Beijing, London, Melbourne and Tokyo), along with its network of 136 National Influenza Centers in 106 countries around the globe, are responsible for monitoring circulation of influenza viruses in humans, while its sister United Nations agency, the Food and Agriculture Organization (FAO), monitors events affecting animal health globally. In addition, Ministries or Departments of Agriculture in each country monitor diseases affecting livestock, especially those for which there are significant economic and animal health implications. In recent years the WHO, FAO and OIE have worked closely together to encourage the sharing of data among the human and animal health sectors.

Dr. Salzberg fails to recognize that what is important is to insure that there is effective collaboration in influenza surveillance of both human and animal viruses, and at the human-animal interface, and that there are sustainable mechanisms for fostering and promoting interactive partnerships for the benefit of the global community. CDC has played a leading role in promoting this goal. The GISAID initiative has provided a mechanism and conditions

---

2 http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm

that have stimulated greater sharing of information within the influenza community, reinforcing an ethos that has sustained the WHO global surveillance network for more than 60 years.

Dr. Salzberg appears to believe that GenBank is preferable to GISAID.4 GenBank is an excellent data archive that plays an extremely important role in scientific research. We submit, however, that GISAID has certain clear advantages over GenBank with respect to influenza data. Among other things, GenBank is solely an archive of sequences. It is not a true relational database. This limits the research that may be done on any influenza virus due to the limited data fields provided. Typically, the information regarding any particular influenza virus isolate extends far beyond nucleotide sequences, and often includes substantial laboratory, epidemiological, clinical and other data that is not accommodated by GenBank. In contrast, GISAID was designed specifically for influenza research and therefore has many more data fields and offers many more ways to examine an isolate and associated metadata.5

We do, however, applaud Dr. Salzberg for raising the issue of influenza research and prevention. It is an important issue and reasoned and vigorous public discussion will benefit everyone. In order for any discussion to be productive, however, it must be based on accurate information. Unfortunately, the misinformation contained in Dr. Salzberg’s article detracts from the points he wishes to make, and is ultimately a disservice to the public good.

4 Although Dr. Salzberg promotes GenBank and denigrates GISAID, he inexplicably fails to disclose his close connection to NCBI (which operates GenBank) and his role in matters that led to the creation of the Influenza Virus Resource (also operated by NCBI), which (not coincidentally) is an influenza database.

5 Dr. Salzberg’s assertion that the WHO requires the CDC to submit genetic sequence data to both GISAID and GenBank is incorrect. The WHA Framework Dr. Salzberg references provides only that WHO GISRS laboratories should submit genetic sequences data to a publicly accessible database, and lists GISAID and GenBank as two examples. See http://whqlibdoc.who.int/publications/2011/9789241503082_eng.pdf. (Section 5.2.2 states, “Recognizing that greater transparency and access concerning influenza virus genetic sequence data is important to public health and there is a movement towards the use of public-domain or public-access databases such as Genbank and GISAID respectively.”) Only influenza viruses of human pandemic potential, not seasonal influenza viruses, are covered by the WHA Framework and there is no requirement that the laboratories submit the information to any particular database.