The world’s microbiology laboratories can be a global microbial sensor network

Thomas F. O’Brien, John Stelling
Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

The microbes that infect us spread in global and local epidemics, and the resistance genes that block their treatment spread within and between them. All we can know about where they are to track and contain them comes from the only places that can see them, the world’s microbiology laboratories, but most report each patient’s microbe only to that patient’s caregiver.

Sensors, ranging from instruments to birdwatchers, are now being linked in electronic networks to monitor and interpret algorithmically in real-time ocean currents, atmospheric carbon, supply-chain inventory, bird migration, etc. To so link the world’s microbiology laboratories as exquisite sensors in a truly lifesaving real-time network their data must be accessed and fully subtyped.

Microbiology laboratories put individual reports into inaccessible paper or mutually incompatible electronic reporting systems, but those from more than 2,200 laboratories in more than 108 countries worldwide are now accessed and translated into compatible WHONET files. These increasingly web-based files could initiate a global microbial sensor network.

Unused microbiology laboratory byproduct data, now from drug susceptibility and biochemical testing but increasingly from new technologies (genotyping, MALDI-TOF, etc.), can be reused to subtype microbes of each genus/species into sub-groupings that are discriminated and traced with greater sensitivity. Ongoing statistical delineation of subtypes from global sensor network data will improve detection of movement into any patient of a microbe or resistance gene from another patient, medical center or country. Growing data on clinical manifestations and global distributions of subtypes can automate comments for patient’s reports, select microbes to genotype and alert responders.

Key words: Drug resistance, microbial; science and technology information networks.

Los laboratorios de microbiología del mundo pueden convertirse en una red de detección microbiana

Los microbios que nos afectan se diseminan por enfermedades locales y globales, y los genes resistentes que bloquean los tratamientos disponibles para combatirlos se reproducen dentro de ellos y se transmiten de unos a otros. Todo lo que sabemos sobre dónde rastrearlos y cómo contenerlos proviene de los únicos lugares en donde es posible examinarlos: los laboratorios de microbiología del mundo.

Sin embargo, la mayoría de estos laboratorios reportan el microorganismo que afecta a cada paciente específico solamente a los responsables de la atención de ese paciente en particular.

Los sensores, que van desde instrumentos hasta observadores de aves, se encuentran hoy conectados por redes electrónicas destinadas a monitorizar e interpretar por medio de algoritmos y en tiempo real las corrientes oceánicas, el carbono de la atmósfera, los inventarios de las cadenas de suministro, la migración de las aves, etc. La vinculación de los laboratorios de microbiología del mundo para que actúen como refinados detectores en una red dedicada a salvar vidas, requiere, no obstante, que sus hallazgos sean sometidos a subtipificación y que sus datos se puedan consultar sin restricción.

Los reportes de los laboratorios de microbiología generalmente son documentos en papel que son inaccesibles o están en sistemas electrónicos de notificación mutuamente incompatibles. No obstante, actualmente los resultados de más de 2.200 laboratorios en más de 108 países han sido traducidos a los archivos compatibles de WHONET y están disponibles para consulta. Con estos archivos en la internet se podría iniciar una red global de detección microbiana.

Los subproductos de información provenientes de los laboratorios de microbiología que hoy no se utilizan, como los datos sobre sensibilidad a medicamentos y los de pruebas bioquímicas, y aquellos que próximamente comenzarán a generar las nuevas tecnologías (genotipificación, técnicas de ionización suave [Matrix-Assisted Laser Desorption/Ionization Time-of-Flight, MALDI-TOF]), etc., pueden reutilizarse en la subtipificación de microorganismos de cada género y especie clasificados en subgrupos susceptibles de ser discriminados y rastreados con mayor precisión.

La delineación estadística de los subtipos que actualmente se lleva a cabo en base en los datos de la red global de sensores mejorará la detección de la transmisión de cualquier microbio o gen resistente de un paciente a otro paciente, centro médico o país. La creciente cantidad de datos relativos a
Microbiology laboratory data uniquely needs to be accessed and subtyped.

When a caregiver sends a specimen from a patient to a laboratory it sends a report back to that caregiver. Reports from the microbiology laboratory differ from those of the other laboratories, such as hematology or biochemistry. The others report measurements of analytes, e.g., serum sodium or hemoglobin that are entirely contained within each patient. A microbiology laboratory reports that in or on that patient is a living microbe, which came from and may go to some place or someone else, and may also have information on where or who else.

That microbiology laboratory or another may have information on whether any of the patients it tested had a microbe similar to one it is now reporting, and noticing who, when and where could reveal spread of a strain or incursion of a new one. Such noticing is limited by microbiology’s using only the one-patient-to-one-caregiver reporting that is sufficient for the other laboratories lacking this need, but also because informatics for that noticing is only now becoming available (1).

The world’s microbiology laboratories will have more information to track such spread when their data are all analyzed together, since microbial threats are increasingly seen as global epidemics (2-19). The strain of microbe infecting a patient anywhere now and the resistance gene blocking its treatment were rarely there a few years or decades earlier (20-26). Each had emerged somewhere, with its early global spread perhaps recorded unnoticed in some laboratory files, but then spread widely largely untracked and uncontained for lack of fully integrated real-time surveillance with alerting of responders (3, 27-30). Prompt recognition of the first incursion into a hospital or country of a new strain or resistance gene gives the best chance to contain its further spread (31).

Microbiology data thus differ not only from data of the other laboratories but also from most other categories of patient data – blood pressure, body mass index, etc. Data of the other categories are so often interdependent that many kinds of healthcare analyses need to be across multiple such data categories. Microbiology reports an encounter of two independent living organisms, patient and microbe. The microbe has meaningful past lineage on other people and places, and more ahead (32).

Microbiology data can thus often be accessed and analyzed in useful ways independently of other healthcare data (33).

The world’s microbiology laboratories can form a global microbial sensor network.

Each report of the world’s microbiology laboratories goes mostly to one caregiver to guide care of one patient. Advances in informatics, however, now open the possibility of recycling these millions of already-paid-for reports into an integrated network database to track spreading microbes and antimicrobial resistance genes everywhere, detect their outbreaks early and coordinate and focus their containment.

Systems of this general kind are being developed and broadly termed electronic sensor networks. Sensors, ranging from weather-monitoring instruments, through inventory-tracking workstations to birdwatchers are deployed widely and their sensed observations interconnected electronically, interpreted algorithmically and variously displayed (34,35). The e-bird network, for example, produces maps of the migration of any species of bird which locate its observed sightings day-by-day, displaying those on any day or advancing rapidly through a season.

It would be helpful to have similar displays for a country, region or the whole world of the daily sightings (reports) of different kinds of emerging specific infecting microbes, with separate mapping of individual sightings of each antibiotype-defined or other subtype. It would be more useful to have, as do some sensor networks, continuous multi-parameter screening of data intake with automated alerting of pre-selected responders for various specific findings. Each of the world’s thousands of microbiology laboratories can be
seen as a real-time sensor and an increasingly discriminating reporter of the microbes infecting patients in its area.

Accessing the information of the world’s microbiology laboratories

The first need for building a global microbial sensor network from the information of the world’s microbiology laboratories is to access that information from each laboratory.

Their reports are now largely inaccessible either because they go only into paper reports and logbooks or into diverse electronic laboratory information systems (LISs) that are incompatible with one another. A resource for accessing the reports of a growing number of those laboratories from all parts of the world, however, has emerged from the global WHONET initiative.

WHONET is a free software program developed and distributed by the World Health Organization Collaborating Centre for Surveillance of Antimicrobial Resistance, based at the Brigham and Women’s Hospital and Harvard Medical School, Boston, USA. Microbiology laboratories put their data into WHONET either by direct data entry or by an automatic translation from a laboratory information system facilitated by a data conversion utility (BackLink) which is included in WHONET (36,37).

WHONET empowers each laboratory to analyze its data in multiple ways, e.g., percentage of all isolates of any kind or of any requested subgrouping that tested susceptible, intermediate or resistant to any or all tested antimicrobials for any time period, percentages and/or line-listings of all isolates that tested resistant to each combination of antimicrobials, scatter-plots of measurements of levels of susceptibility of isolates of any type to any pair of antimicrobials, etc. Additional functions, e.g. an outbreak-detection algorithm (SaTScan), continue to be added (38).

As laboratories entered or translated their reports into WHONET their resulting WHONET files shared the same codes and structure and were thus inter-compatible and able to be merged to form both national and international multicenter surveillance networks (39). As a result, in a world where any two medical centers or medical center systems rarely have inter-compatible data they could share, there are now more than 2200 microbiology laboratories in more than 108 countries around the world that have inter-compatible WHONET files. As WHONET now transitions to a web basis a growing subset of these files can be accessed to pilot a global microbial sensor network.

Improving microbiology laboratory information by subtyping

Microbiology laboratories have been paid for a century to report the genus/species identity of a microbe in a patient’s specimen—initially to anticipate the clinical syndromes but later also the probabilities of resistance to various antimicrobials associated with each identity. But species of microbes don’t spread and cause outbreaks. Strains of species spread, and so distinguishing strains from one another would optimize detection of spread. An outbreak of five cases in a month might be noticed if it were seen as the only isolates of a particular strain in that hospital that year, but not if seen only as five of several hundred isolates of that species that year.

Subtyping isolates of one genus/species into smaller groupings will thus improve detection of outbreaks and tracking of spreading strains. Each of the subtypes of a larger grouping of microbes has had a more recent common ancestor and less time to diverge than the larger grouping. Ultimate subtyping to, or nearly to, the strain level would optimize detection and tracking, as shown by the many outbreaks tracked and contained by the long-practiced serological subtyping of non-typhoidal salmonellae (40). But any subtyping that subdivided a genus/species into any number of sub-groups could enhance such detection and tracking in rough proportion to the number of such sub-groups.

Most microbiology laboratories could subtype partially, and many extensively, with data they produce now, and most could soon subtype more. Most could subtype now by the antimicrobial susceptibility measurements they make for the up-to-half of their isolates that are insusceptible to at least one of the tested antimicrobials (41). Many could subtype further by the panels of up to 48 biochemical tests they use now only to identify to the genus/species level. A laboratory may now labor to identify a rare species, but then issue a quarter of its reports only as *Escherichia coli*, which has many subgroups with differing epidemiology and clinical manifestations (42,43).

Technology for further subtyping is growing. Microbiology laboratories are beginning to use matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) instruments (44). They may
use them to identify only the genus/species of each microbe, as done now, and vendors may set them to do only that. MALDI-TOF can distinguish subtypes within each genus/species, however, but making that routine may require both wider recognition of the value of subtyping for strain-tracking and outbreak detection, as sketched here, and informatics to manage it (45).

Other subtyping technologies are coming into increasing use, such as multi-locus sequence typing (MLST), polymerase chain reaction (PCR), including that for 18s ribosomal RNA, and ultimately full genome nucleotide sequencing (46-49). These may be done selectively for problems of a patient or a hospital’s infection control, but informatics to capture, integrate and compare their results across hospitals will amplify their value.

Reports from many academic centers throughout the world are being published that commonly have used multiple tests of these special kinds to describe in detail locally sampled epidemic or endemic microbes or their genetic elements. The publications may appear years after sampling and would be difficult to interrelate or overview. Many tests of these kinds, for example, are being done in many laboratories on strains carrying the KPC or NDM carbapenemases, but most of their results are now sequestered in the files of scattered laboratories or publications. Integrating and analyzing together all those results in near-real time would greatly improve understanding of the spread of what are dreaded now as among the most menacing kinds of antimicrobial resistance (27,50).

Example of the need for an integrated global subtyping process

We have previously reported an example of the need for developing a globally integrated subtyping process (33). The minimal inhibitory concentrations (MICs) of two antibiotics for 5 blood and urine isolates of *E. coli* over a month from one patient distinguished them from any of the 3615 other isolates of *E. coli* tested by that laboratory that year. The patient had received a kidney transplant several months earlier at a hospital in a distant continent. The observation made in figure 2 of that report was noticed later by chance in a WHONET scatter-plot, which happened to be made for a different reason.

The strain of *E. coli* isolated repeatedly from urine and blood of this returning foreign transplant recipient but from no other patient in the hospital that year was presumably acquired at the foreign transplant hospital, where it may have been widespread, and imported in the infected kidney. Many or most of the strains of resistant bacteria or of the resistance genetic elements circulating in any country, community, hospital or hospital ward may have first appeared in this way and in an index case such as this before spreading. Many may not have infected another cultured patient, as this one was not found to have done, but enough did to create the problems.

This presumptive import happened to be noticed as a distinctive subtype by the very unlikely chance observation of only two of the measurements made routinely by the laboratory, which also routinely records on such isolates the results of 47 biochemical tests and 4-6 MIC values for each of 17 tested antimicrobials. Statistical comparisons of the results of all of these tests for this patient’s isolates with all of those for this hospital’s other *E. coli* isolates might have added further evidence for its being a distinctive subtype.

The important lesson from this and other experiences is that optimal detection of such microbe movements needs integrated observation of so many variables on a global scale as to usually escape human noticing and require an informatics-supported system with automated alerting.

Informatics for a global microbial sensor network tracking microbial subtypes

The data management and processing needed to develop the optimal delineation of such subtypes and to automate the noticing and reporting of their problem interrelationships on a global level is now becoming available. Recent advances in informatic technology have made possible “Big Data” projects for management of massive databases, now often much larger than would be needed for a global microbial sensor network tracking microbial subtypes.

An approach could be to begin with a subtyping system based initially only on qualitative antibiotypes, e.g., the combinations of antibiotics to which a microbe’s level of resistance exceeds the susceptible-intermediate breakpoint. These have generally been sufficiently stable on repeated isolates of the same genus/species from the same patient to infer sufficient identity of that strain in another patient to detect most such transfers (30,48,51,52). They have also been shown to greatly increase sensitivity of detection of clusters of cases in trials of the SaTScan program (53,54).
Artificial intelligence routines can be developed to help delineate the initial qualitative antibiotype-based subtypes and to upgrade and reconfigure them and update their findings as global data grows. The same process could be extended to explore the advantages and problems of further subtyping by the use of quantitative antibiotypes, the combinations of antimicrobials that result from categorizing a microbe by the combinations of the measured levels of resistance to each of the antimicrobials to which it was tested. The resulting exponential expansion of the number of different antibiotypes would reduce their reproducibility but would enhance their sensitivity, as they did for the presumptive imported *E. coli* described above, which was detected by its quantitative antibiotype.

The same process could be applied progressively to adding available data from results of other kinds of testing to further discriminate subtyping. An early example would be the results of the many biochemical tests generated routinely by microbe-identifying instruments now in wide use, which we have shown to support biotyping. Increasingly used MALDI-TOF instruments appear to have, as mentioned above, potential for highly discriminating routine subtyping of all reported microbes if the value of subtyping is recognized and its vendors adjust it to report subtypes.

Additional subtyping power will come from selectively but increasingly used genotyping tests. These include MLST, which pioneered the integrated global filing and interpretation of test results that we advocate here, PCR and the ultimate, full genome sequencing (46). These offer the highest discrimination of subtyping, but their selective use limits their availability. Integrating their results with the commonly available subtyping methods results, however, will amplify the value of both. Many of the isolates of the now-threatening KPC and NDM-expressing strains of *Enterobacteriaceae* as mentioned above have, for example, been tested somewhere with one or more of these methods and retrieving and integrating all of that data could better delineate the subtyping and so also the epidemiology of that menace.

**A global electronic network interprets each patient’s microbe and also alerts responders.**

As a subtyping “engine” continues to distinguish and subdistinguish subtypes from the multiple kinds of data accumulating in its growing global database it can also record for each subtype its geographic distributions over time and its clinical characteristics, e.g., predominantly hospital or community-isolated, preferred anatomical sites of isolation, etc. It can also summarize these into commentary for each subtype and maintain an accessible updated dictionary of such comments. The electronic network would then locate the subtype of each patient’s microbe as a laboratory is about to report it and automatically present that subtype’s current comment for optional inclusion in the laboratory’s report to that patient’s caregiver.

Another set of algorithmic analyses running on the growing global database would screen for unusual time-space distribution of any subtypes, as we have done successfully with SaTScan in hospitals and for certain kinds of microbes in regions of countries (53,54). Whenever a statistical threshold is exceeded for an area, an alert is sent automatically to responders pre-selected for that area. The areas for many such alerts would span multiple hospitals or even countries, and many of those responders would thus be in the only organizations with such broad jurisdiction, area public health agencies, which have often been less involved with antimicrobial resistance.

This could fill what appears to be now a huge gap in the detection of spread particularly of subtypes of multidrug-resistant bacteria. Most hospitals have infection preventionists, each of whom works intensively to control such spread within their own hospital but often with little knowledge of what is in the next hospital, or in the chronic care facilities that send them patients or just beginning to come into that part of the country. Comprehensive, globally interpreted data shared between an area’s preventionists and its public health agencies with automated alerting of appropriate responders in both will help to close this gap.

**Conflicts of interest**

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Funding**

Research reported in this publication was supported by the National Institute of General Medical Science of the National Institutes of Health under award number R01GM103525.

**References**


Microbiology laboratories as a sensor network


