SPECIAL REPORT

Global Public Health Database Support to Population-Based Management of Pandemics and Global Public Health Crises, Part II: The Database

Frederick M. Burkle, Jr., MD, MPH, DTM, FAAP, FACEP;[1,2] David A. Bradt, MD, MPH, FACEP, FACEM, FAFPHM, DTM&H;[3] Joseph Green, PhD;[4] Benjamin J. Ryan, PhD, MPH[5]

1. Professor (Ret.) Senior Fellow and Scientist, Harvard Humanitarian Initiative, Harvard University, T.H. Chan School of Public Health, Cambridge, Massachusetts, USA
2. Global Scholar, Woodrow Wilson International Center for Scholars, Washington, DC, USA
3. Dept of International Health, Johns Hopkins School of Public Health, Baltimore, Maryland, USA
4. Acting Director of Applied Science, Pacific Disaster Center, Hawaii, USA
5. Clinical Associate Professor, Department of Environmental Science, Baylor University, Waco, Texas, USA

Correspondence:
Frederick M. Burkle, Jr., MD
Senior Fellow & Scientist
Harvard Humanitarian Initiative
c/o 452 Iana Street
Kailua, Hawaii 96743, USA
Email: Skipmd77@aol.com;
fburkle@hsph.harvard.edu

Conflicts of interest: none

Keywords: global public health; pandemics; population-based management; public health; triage

Abbreviations:
AI: artificial intelligence
COVID-19: coronavirus disease 2019
GPH: global public health
PH: public health
PBM: population-based management
PBMT: population-based management team
WHO: World Health Organization

Received: September 20, 2020
Revised: October 9, 2020
Accepted: October 12, 2020

doi:10.1017/S1049023X20001363

Abstract

This two-part article examines the global public health (GPH) information system deficits emerging in the coronavirus disease 2019 (COVID-19) pandemic. It surveys past, missed opportunities for public health (PH) information system and operational improvements, examines current megatrend changes to information management, and describes a new multi-disciplinary model for population-based management (PBM) supported by a GPH Database applicable to pandemics and GPH crises.

Burkle FM Jr., Bradt DA, Green J, Ryan BJ. Global Public Health Database support to population-based management of pandemics and global public health crises, Part II: the database. *Prehosp Disaster Med.* 2021;36(1):105–110.

Introduction

The 2019 novel coronavirus (SARS-CoV-2, 2019-nCoV) disease (COVID-19) has become the first pandemic of the 21st century. Viral contributory causes include high virus shedding from upper respiratory tract secretions, symptomatic and pre-symptomatic virus transmission, reproduction ratio estimates up to 3.58,[1] little effective antiviral therapy, and no effective vaccine. Health system and community contributory causes include wide-spread community seeding through unrestricted air travel, inadequate implementation of non-pharmacologic control measures, limited contact tracing and viral diagnostic capabilities, overwhelmed clinical care facilities, and threat minimization by elected officials.

These contributory causes have drawn attention to potential remedies for pandemic management by public health (PH) authorities. Part I reviewed past missed opportunities for PH information systems and operational improvements, then examined important PH megatrends heralding future advances in pandemic management—evidence-based decision making, data literacy, ascent of PH information services, and ascent of remote management. Part I also reviewed the emergence of population-based management (PBM) of health crises. A multi-disciplinary model led by PH professionals is required to ensure data driven decisions can be made that reflect the PH needs and risks of local communities. This approach builds on the evolution of disaster management systems, which are used in most countries across the world, to regularly deal with local, state, national, and international crises. This article examines database requirements and modalities of functioning for PBM teams (PBMTs).

Digital Technology

Digital technologies may help perform a wide range of functions in pandemic management, as summarized in Table 1.[2] Health authorities in some countries have attributed their disease control successes to early adoption of such technology.[3-5]

Nonetheless, major issues have arisen on data oversight, privacy, protection of personally identifiable health information, and redress for errors that implementing jurisdictions must confront. The scientific community has seen retractions of papers published in *The Lancet* and *The New England Journal of Medicine* due to post-publication concerns over provenance

Function	Technology	Early Adopters
Infection Screening	Digital thermometers, infrared thermal cameras, web-based tools	China, Iceland, Singapore, Taiwan
Contact Tracing	GPS, mobile phone apps, facial recognition technology	China, Germany, Singapore, South Korea
Quarantine and Isolation	GPS, mobile phone apps, surveillance of mobile devices for voluntary or involuntary tracking	China, Iceland, South Korea, Taiwan, Australia
Clinical Management	Lab test pooling, telemedicine, video-conferencing, AI algorithms for clinical outcome prediction	Australia, Canada, China, USA
Medical Logistics	Mobile apps for transactions, barcode scanners for commodities, robotics, and drones for delivery	US, China
Epidemic/Pandemic Tracking	Data dashboards, real-time data transmitted by smartphones and PDAs	China, Singapore, Sweden, USA

Table 1. Applications of Digital Technology in Pandemics
Source: Adapted from Whitelaw.[2]
Abbreviations: AI, artificial intelligence; PDA, personal digital assistant.

of proprietary data and inconsistencies from electronic health records.[6] This experience is cautionary about the importance of data source/origin, data audit trails, and oversight in managing such datasets.

Core Databases

Core Databases for PBM are characterized below in three overriding domains:

1. Background Information (Table 2) describes the affected population pre-pandemic—demography, access to essential services, health status, and socioeconomic status.
2. Clinical Case and Virological Data (Table 3) encompasses case definitions, case identification and outcomes, and clinical dynamics (case-fatality rates, death rates, and time course to death).
3. PH Control Measures and Consequences (Table 4) encompasses details of contact tracing, quarantine measures, travel restrictions, and all epidemic modelling.

The Database Challenge

The technical standards for disease surveillance and surveillance systems comprise an extensive literature. The Centers for Disease Control and Prevention (CDC; Atlanta, Georgia USA) published its first guidelines in 1988, which were re-issued in 2001.[7,8] The World Health Organization (WHO; Geneva, Switzerland) Program on Disease Control in Humanitarian Emergencies published numerous adaptations of communicable disease and surveillance guidelines for epidemic-prone diseases in natural disasters and complex emergencies through the 2000s. Recent emerging infectious diseases such as Ebola have prompted additional reviews of data needed for outbreak evaluation culminating in descriptions of an emerging science of outbreak analytics.[9,10]

Taking into account the rate of global population growth, it was predicted in 2019 there would be 1.7MB data created every second for every person on Earth, bringing forth new data problems, "most notably where and how to store it."[9] With unprecedented movement of migrants and refugees from the Middle East, compounded by massive relocations of urban populations because of COVID-19, are only two entities that will impact data collection critical to future infectious diseases. The era of data-centric computing is here, rate of product change accelerating, storage options expanding, and already digital transformation has taken center stage as a key driver. Outbreaks, epidemics, and pandemics are driving streaming data, larger memory servers, video, artificial intelligence (AI), and machine learning (crucial to guarantee data security), and need for unprecedented "big data and analysis."[11] Multi-disciplinary AI questions to the proposed Global Public Health (GPH) Database will be increasingly common, as it should be in pandemics. All these demands require new training and professions. Cost effective and resilient Cloud storage infrastructure can adapt to rapid workload requirements such as transient or spiking demands of world-wide PBM requirements. With the advent of "fast and dense memory," the intelligent storage can be trained to be largely self-managing, responding to detect anything from race conditions or access to care changes. Harris concludes that the era of data-centric computing is here with the potential to bring top-priority to the demands of the proposed PBM system.[11]

The Economist is more detailed in their assessment in managing future pandemics accelerated by population growth and globalization that have markedly increased the spread of infectious diseases:[12]

- Data analysis can pinpoint outbreaks and reveal where they may unfold next.
- Can show where people were moving within 24 hours, right from the initial outbreak, and could distribute that information widely, especially PBMTs.
- Can take advantage of diverse skill sets in tracking new, less-understood pandemics, favoring the proposed multi-disciplinary focus of their teams to include: "physicians, veterinarians, ecologists, data scientists, epidemiologists, environmental health specialists, geographers, designers, and software developers," to help make sense of raw data and AI requests and critical mobility-related information.[12]
- The WHO concludes that "working together with data firms and health agencies (ie, PBMTs) before a disaster strikes mitigates searching for data at the start of an outbreak, which is already too late, reiterating a multi-stakeholder collaboration archives more than working in isolation."[12]

https://doi.org/10.1017/S1049023X20001363 Published online by Cambridge University Press

A. **Population at Risk**

1. Census — M — F — Total
    - a. < 1
    - b. 1-4
    - c. 5-14
    - d. 15-44
    - e. ≥ 45
    - f. total
2. Demography
    - a. race, ethnicity, religion
    - b. languages
    - c. socioeconomic status
    - d. urban vs rural
    - e. average family/household size

B. **Settlements**

1. Locations
    - a. community size, geographic separation
    - b. access to transportation hubs and communication links
2. Population Mobility
    - a. diurnal and seasonal occupational travel flows
    - b. recreational and tourist travel flows
    - c. migration flows
3. Community Density
    - a. geographic separation of families within the settlement
4. Hazard Exposure (eg upcoming seasonal natural hazards)
5. Discrimination and Conflict within the Community
6. Emergency Preparedness
7. Known Coping Mechanisms

C. **Access to Essential Services**

1. Public Utilities (infrastructure)
    - a. % population with access to improved water
    - b. % population with access to improved, unshared sanitation
    - c. % population with electricity
    - d. % population with cell phones
2. Health Services
    - a. skilled health care workers (all)/10,000 p (22)
    - b. physicians/10,000 p (0.2)
    - c. nurses/10,000 p (1)
    - d. midwives/10,000 p (1)
    - e. community health workers/10,000 p (10)
    - f. % population with provider coverage (100%)
    - g. consultations/p/yr (1)
    - h. immunization rates—measles vac in 1 y/o (90%)
    - i. hospital beds/10,000 p (10 - 120)
    - j. admissions/10,000 p
3. Underserved Areas (by location and type of service)

D. **Health Status**

1. Mortality and Oher Rates
    - a. CBR (10 - 50)
    - b. CMR (5 - 25)
    - c. AGR (%)
    - d. IMR (3 - 200)
    - e. U5MR (3 - 300)
    - f. MMR (5 - 2,100)

Table 2. Datasets and Visualizations Pre-Pandemic *(continued)*

https://doi.org/10.1017/S1049023X20001363 Published online by Cambridge University Press

2. Nutritional Status
    a. % infants with low birth weight (LBW) (7% with < 2.5 kg)
    b. % U5 with growth faltering (30%)
    c. % with micronutrient deficiencies (vitamins A-D; Fe, I)
3. Endemic Diseases and Comorbidities
    a. Chronic Illnesses and Conditions
        (1) hypertension
        (2) diabetes
        (3) obesity
        (4) disability
    b. Infectious Diseases
        (1) HIV prevalence
        (2) TB prevalence
    c. Other Major Diseases
4. Routine/Baseline Disease Surveillance Data
    a. Baseline infectious incidence rates
    b. Baseline hospital admission rates
    c. Baseline death rates ratios
5. Health Beliefs and Traditions
6. Sources of Care
    a. Health infrastructure
7. Referral System

E. **Macroeconomic Status**

1. Labor force occupations
2. Unemployment rate
3. GNI per capita ($US) ($700 - $70,000)
4. Inflation rate
5. % population in extreme poverty (< $1.90/d income)
6. Current health expenditures per capita ($30 - 10,000)

F. **Educational Status**

1. Adult literacy rate
2. Female literacy rate
3. Adolescent literacy rate

Table 2. *(continued)* Datasets and Visualizations Pre-Pandemic

Artificial Intelligence in support of the GPH Database

Expectations were high for AI to help fight COVID-19. However, before AI tools can make an impact, global collaboration and high-quality data and model sharing are needed. *Nature Machine Medicine* emphasized that early warning and alerts, prediction and detection of outbreak of diseases, real-time disease monitoring world-wide, analysis and visualization of spreading trends, prediction of infection rate and infection trend, and rapid decision making to identify the effectiveness of management.[13] Other authors emphasize challenges that need to be addressed before AI can have a beneficial global impact.[14-16] They point out that a first challenge is knowing where to start with developing AI tools that can be most effective, which requires close cooperation with practitioners at the health care frontline, those that manage the PDMTs. The best solutions may involve adapting already validated systems rather than building new tools from scratch. Furthermore, Hu, et al described how clinical needs are evolving as the pandemic is moving through different stages, from early detection and anticipation, to containment and mitigation, and finally eradication. During these transitions, the specific types of AI models may need to change too. Ideally, the AI system itself becomes a smart, intelligent, and adaptive system.[13]

Education and Training for PBM and the GPH Database

The origins of this pandemic can be attributed to climate change and the emergence of complex GPH crises such as climate extremes, biodiversity loss, emergencies of scarcity, rapid unsustainable urbanization, migrant and refugee surges, domestic and international terrorism, the civilianization of war and conflict, and the global rise of resistant antibiotics that have resulted in an unprecedented rise in direct and indirect mortality and morbidity, among others.[17] In 2019, *The Tohuku Journal of Experimental Medicine* recognized the emergence of complex GPH crises that are beyond the current decision-making and operational capabilities of traditional disaster management and its providers, most of who are community-level practitioners representing every discipline.

The 1930s "disaster cycle" concept described a phase-related approach to meeting the strategic, operational, research, educational, and training components required of disasters; this presents an opportunity for the structured development of a Health Crisis Management Framework to oversee the phase-related strategic and operational requirements for prevention, preparedness, response, recovery, and rehabilitation challenges of major GPH crises.[18] No longer will physicians, for example, Clinical and Virological Data:

A. **Clinical Data**

1. Case Definitions (suspected, probable, confirmed)
2. Severe Cases (admissions, ICU cases, ventilated cases)
   a. Age/gender/race data
   b. Comorbidities
   c. Clinical course, complications, and outcomes
   d. Data on use of antivirals, antibodies, and vaccines
   e. Discharge diagnoses
   f. Causes of death
3. Clinical Dynamics—population morbidity and mortality rates, age & gender-specific CFRs, time course to death
4. Hotspot Definitions—cases linked temporally or geographically triggering further outbreak investigation and/or control measures

B. **Virological Data**

1. Sentinel Site Testing
2. Confirmatory testing for new cases in the community meeting the surveillance definition (before sustained community transmission is identified)
   a. daily new positive cases
   b. 14 day running average
3. Confirmatory testing for hospital admissions and unexplained deaths that are compatible with the disease
4. Community seroprevalence testing once community transmission is established
5. On-demand individual testing (depending on resource availability, may be targeted at vulnerable groups)
6. Antiviral Sensitivity Testing

Table 3. Datasets and Visualizations Intra-Pandemic

Modelling, Control Measures, and Impact:

A. **Epidemic Modelling Data**

1. Population sizes and subgroups at risk
2. Compartment models
3. Spot maps (geolocated plots of disease outbreaks)
4. Kinetics models of transmission dynamics—incubation period, reproduction ratio, contacts per case
5. Epidemic curves
6. Epidemic projections showing impact of non-pharmaceutical mitigation measures

B. **Public Health Control Measures**

1. Travel restrictions
2. Arrival testing
3. Contact tracing effectiveness
4. Quarantine requirements (site specification, duration, monitoring), social acceptance of quarantine measures, overall adequacy of quarantine
5. Social distancing recommendations
6. Mask recommendations
7. Restrictions on social gatherings
8. Lockdown onset, duration
9. Population sizes & community densities of populations affected

C. **Socio-Economic Impact Data**

1. Displacements
2. Unemployment rates
3. Evictions
4. Business closures
5. Direct costs in health care
6. Indirect costs of business losses
7. Government expenditures for unemployed persons and distressed businesses

Table 4. Datasets and Visualizations Intra-Pandemic

focus on the response phase alone. Future PBMTs will be trained across the entire disaster cycle. Operational knowledge on prevention and preparedness will lessen the number of victims more than any other action. The Japanese Editorial launched academic interest in training multi-disciplinary Health Care Crisis Managers and Scientists to fill the full- and part-time roles required of PBM, PBMTs, and those managing the GPH Database. For example, the university-centered CRIMEDIM (Research Center in Emergency and Disaster Medicine) in Navaro, Italy has worked diligently to offer both Masters and PhD degrees leading to training Health Crisis Managers and Scientists from many contributing disciplines to fill these proposed PBMT positions. A global effort is recommended post-COVID of multi-disciplinary educational programs to help prepare the potential PBMT and GPH Database experts that this study recommends.

https://doi.org/10.1017/S1049023X20001363 Published online by Cambridge University Press

References

1. Lai CC, Shih TP, Ko WC, et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. *Int J Antimicrob Agents.* 2020;55(3):105924.
2. Whitelaw S, Mamas M, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digital Health.* 2020;2(8);e435-e440.
3. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *JAMA.* 2020;323(14):1341-1342.
4. University of Oxford. Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown. https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown. Accessed September 6, 2020.
5. World Health Organization Western Pacific Region. New Zealand takes early and hard action to tackle COVID-19. https://www.who.int/westernpacific/news/feature-stories/detail/new-zealand-takes-early-and-hard-action-to-tackle-covid-19. Accessed September 6, 2020.
6. The Lancet Digital Health. Transparency during global heath emergencies. *Lancet Digital Health.* 2020;2(9):E441.
7. Klaucke DN, Buehler JW, Thacker SB, et al. Guidelines for evaluating surveillance systems. *MMWR.* 1988;37(S-5):1-18.
8. German RR, Lee LM, Horan JM, et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR.* 2001;50(RR-13):1-35.
9. Cori A, Donnelly CA, Dorigatti I, et al. Key data for outbreak evaluation: building on the Ebola experience. *Phil Trans R Soc.* 2017;372(1721):20160371.
10. Polonsky JA, Baidjoe A, Kamvar AN, et al. Outbreak analytics: a developing data science for informing the response to emerging pathogens. *Phil Trans R Soc.* 2019;374(1776):20180276.
11. Harris R. Data storage: everything you need to know about emerging technologies. https://www.zdnet.com/article/innovations-in-data-storage-an-executive-guide-to-emerging-technologies-and-trends/. Accessed September 6, 2020.
12. The Economist/Intelligence Unite. From chaos to coherence: managing pandemics with date. https://expectexceptional.economist.com/managing-pandemics-with-data.html. Accessed September 2, 2020.
13. Finding a role for AI in the pandemic. *Nat Mach Intell.* 2020;291(2).
14. Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. *AI Soc.* 2020. Epub ahead of print.
15. Bragazzi NL, Dai H, Damiani G, Behzadifar M, Martini M, Wu J. How big data and artificial intelligence can help better manage the COVID-19 pandemic. *Int J Environ Res Public Health.* 2020;17(9):3176.
16. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr.* 2020;14(4):337-339.
17. Benis A, Notea A, Barkan R. Risk and disaster management: from planning and expertise to smart, intelligent, and adaptive systems. *Stud Health Technol Inform.* 2018;247:286-290.
18. Burkle FM. Challenges of global public health emergencies: development of a health-crisis management framework. *The Tohoku Journal of Experimental Medicine.* 2019;249(1).