Published in final edited form as:
Curr Fungal Infect Rep. 2024 March ; 18(2): 88–94. doi:10.1007/s12281-024-00494-4.

Fungal Infections and Social Determinants of Health: Using Data to Identify Disparities

Samantha L. Williams[1], **Kaitlin Benedict**[1], **Mitsuru Toda**[1]

[1]Mycotic Diseases Branch, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Mailstop H24-11, Atlanta, GA, USA

Abstract

Purpose of Review—Fungal diseases disproportionately affect certain demographic populations, but few studies have thoroughly investigated the drivers of those disparities. We summarize data sources that can be considered to explore potential associations between fungal diseases and social determinants of health in the United States.

Recent Findings—Sociodemographic disparities are apparent in fungal diseases, and social determinants of health (e.g., income, living conditions, and healthcare access) may be associated with increased risk of infection, severe disease, and poor health outcomes.

Summary—Numerous data sources are available in the United States to analyze the potential association between fungal diseases and underlying social determinants of health. Each source has benefits and limitations that should be considered in the development of analysis plans. Inherent challenges to all fungal disease data (e.g., underdiagnosis, underreporting, and inability to detect people who do not seek medical care) should be noted and accounted for in interpretation of results.

Keywords

Fungal disease; Mycotic disease; Health disparities; Health equity research; Social determinants of health

Introduction

Fungal diseases represent a diverse group of infections, ranging from superficial infections of the skin, nails, and mucous membranes (e.g., dermatophytosis and vulvovaginal candidiasis) to invasive disease that can lead to serious illness and death (e.g., aspergillosis, blastomycosis, and candidemia). Approximately 6.5 million incident invasive fungal

---

[*]Samantha L. Williams, pog3@cdc.gov.

Author Contribution S.W. and M.T. conceptualized the manuscript. S.W. wrote the main manuscript text and prepared the table. All authors reviewed the manuscript.

Ethics Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Competing Interests The authors declare no competing interests.

Disclaimer The findings and conclusions of this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control (CDC).

infections are estimated occur worldwide each year, and the true incidence of cases is likely higher than documented because these diseases often are underdiagnosed and underreported [1]. These diseases also inflict considerable financial burden on patients and healthcare systems; estimated direct and indirect costs total more than $11.5 billion annually in the United States [2]. Certain populations are disproportionately affected by some fungal infections, but health disparities related to fungal diseases and the underlying mechanisms that drive them are not well understood [3, 4].

Health disparities reflect differences in health outcomes that are often linked to systematic social, economic, or environmental disadvantage. They are often connected to social determinants of health, conditions in which people live (e.g., education, income, occupation and working conditions, housing, and access to healthcare) that are influenced by wider forces and systems (e.g., economic policies, social norms, and political systems) [5]. Similarly, disproportionate exposure to certain adverse environmental features (e.g., air pollution and hazardous waste sites) may affect health outcomes. Differences in social determinants of health or environmental burden may be exacerbated by residence in rural or urban settings [6]. Identification of health disparities is an essential step to advance health equity research, the examination of patterns and factors that contribute to health inequities, and ultimately achieve health equity.

Literature on health disparities in fungal diseases is limited. Many studies have reported long-standing differences in demographic characteristics for fungal diseases. For example, racial and ethnic disparities are readily apparent for certain fungal diseases (e.g., coccidioidomycosis, cryptococcosis, and *Candida* bloodstream infections) [3, 7]. Additionally, males are generally more likely to be diagnosed with invasive fungal infections (IFI) compared with females, and people aged 65 years or older are often at greater risk for IFIs compared with younger people [3]. However, few studies have comprehensively explored the reasons for these demographic differences, particularly across fungal pathogens.

Some of the documented disparities may be attributable to genetic or biologic factors, but risk of fungal infection is likely also associated with differences in social determinants of health. For example, studies have shown that people with lower median household incomes are more likely to be hospitalized for most IFIs in the United States, and risk of superficial fungal infections is greater among lower socioeconomic status groups internationally [3, 4]. Certain occupations, particularly low-wage or outdoor occupations, have been associated with fungal infections and outbreaks, and exposure risk may be exacerbated by poor working conditions [4, 8]. Overcrowding and poor living conditions have demonstrated increased risk of superficial fungal infection [4, 9], while low-quality housing may produce conditions conducive to indoor mold and allergens [10]. Rates of some IFIs, like histoplasmosis, are higher in residents of rural areas, while rates of coccidioidomycosis, pneumocystosis, and aspergillosis are greater among urban residents [3]. Lack of access to affordable healthcare services (e.g., long distance to a healthcare center, lack of health insurance, and high medical costs) is also associated with greater risk of predisposing conditions for IFIs, and low health literacy may hinder interactions with the healthcare system [4, 11]. Many of these socioeconomic factors are interrelated and therefore may have a cumulative impact on increasing risk of fungal infections among affected populations.

Data Challenges

There is growing interest in understanding health disparities in fungal infections and associations with underlying social determinants of health, but limitations of existing data sources present several challenges for researchers. An overarching limitation of most sources of fungal disease data is the inability to capture cases among persons who do not seek healthcare. In addition, the nonspecific symptoms of many fungal diseases, along with low awareness, can lead to delayed or missed diagnosis. Although associations between social determinants of health and healthcare-seeking behavior, medical care, and diagnosis have been reported [12-16], it is not well understood how these factors may differ by sociodemographic subgroups for fungal diseases.

Datasets are often subject to small sample sizes for certain uncommon IFIs (e.g., mucormycosis and blastomycosis). Even with more common IFIs, stratified analyses by racial, ethnic, or socioeconomic subgroups may be difficult; small samples may also complicate the identification of confounding variables. Potential solutions may involve alternate sampling strategies if appropriate, such as oversampling populations of interest, or pooling data across years, which may be complicated if data collection methods or instruments are inconsistent over time [17]. Importantly, data for variables relevant to health disparities, such as race, ethnicity, and income, are frequently either not collected in or are missing from existing sources of IFI data. Although imputation techniques or proxy variables may be used to address the issue of missing data, they are inherently imperfect and require additional assumptions [17]. Finally, race and ethnicity categorizations may contribute to measurement error if patients or survey respondents do not identify with any of the listed groups or do not report race or ethnicity, which disproportionately affects certain groups [18, 19]. Collection of more granular race and ethnicity data may be preferable but could result in small samples.

Data Sources

With the aforementioned challenges in mind, numerous data sources are available to explore potential health disparities in fungal diseases. Here, we describe different types of data sources available in the United States that might be useful in such analyses, including their benefits and limitations. Detailed advantages and considerations for each data source are summarized in Table 1.

Public Health Surveillance Data

Public health surveillance data are systematically collected on an ongoing basis and are essential to help characterize disease and measure trends over time. Surveillance data for all diseases and conditions are especially useful to monitor case counts, emerging populations at risk of disease, and geographic distribution and expansion. In the United States, reportable status is designated at the state level and mandates healthcare professionals and laboratories to notify public health departments of diagnosed cases or positive laboratory tests for given diseases. However, not all fungal diseases are reportable, and of those that are, they are not reportable in all states, which may limit geographic representation [20]. While standardized surveillance case definitions are intended to facilitate comparison across jurisdictions, some jurisdictions do not follow these standardized definitions.

Surveillance data are not always available in a timely manner; although provisional data are typically presented at regular intervals (e.g., weekly and monthly), reconciled data may lag by months to years. The breadth, quality, and completeness of the data also depend on the information reported by the healthcare provider or laboratory. Key data elements for evaluating health disparities, such as race, ethnicity, disease severity, health outcomes, and factors related to social determinants of health, are frequently incomplete or not collected.

Surveillance systems are meant for continuous data collection, and therefore, opportunities to add or adjust data elements are often limited. However, surveillance practices and case definitions may change over time based on improved disease knowledge or changes in laboratory diagnostic methods (e.g., advances in new diagnostic tools and availability of laboratory tests). These artifacts of reporting should be considered, if applicable, in analyses of surveillance data.

Hospital Discharge and Hospital Survey Data

Hospital discharge and survey data can be used to explore health disparities in fungal diseases through healthcare encounters. Hospital discharge datasets enable researchers to assess patient interactions with the healthcare system, including diagnosis, treatment, disease management, and health outcomes. These data are especially well-suited for studying severe fungal diseases that require hospitalization and may be particularly influenced by the presence of underlying medical conditions and social determinants of health. Hospital discharge data generally comprise large sample sizes that can be nationally representative, and data are collected longitudinally, though availability may lag by several years, limiting timeliness.

Demographic information, including race and ethnicity, is often available and fairly complete, though accuracy often differs by racial and ethnic groups [21]. Payer type (e.g., private insurance, Medicare, and Medicaid) provides insight into insurance status, which may be associated with IFI, likely due to its connection with socioeconomic status and access to quality healthcare [3, 4]. Healthcare charges and associated costs are also included and can be used to analyze financial burden in connection with health disparities. Hospital survey datasets may also contain facility-level data related to social determinants of health and health equity, particularly around measures related to healthcare quality.

Hospital discharge and survey datasets almost certainly underestimate the true number of healthcare visits associated with fungal diseases. Disease classification based on medical coding (e.g., *International Classification of Diseases, Tenth Revision* [ICD-10] codes) can be inaccurate or nonspecific. For example, a recent study found that only 44% of culture-confirmed candidemia cases received an ICD-10 code for candidiasis [22]. Hospital discharge data may also disproportionately represent more severe cases of IFI, which can potentially misrepresent health disparities and limit generalizability to all cases.

Insurance Claims Data

Insurance claims databases offer similar benefits and are subject to similar limitations as hospital discharge datasets. They are made up of large samples and may contain longitudinal data related to healthcare encounters, including diagnosis, treatment, outcomes, and cost. Insurance information can help identify disparities in fungal disease rates and outcomes through comparison across insurance types. However, information about social determinants of health may be limited.

Data may also be skewed toward specific types of insurance, and therefore, researchers should be cautious about generalizing results to the total population.

Laboratory Testing Data

Laboratory testing data provide insights into diagnostic testing practices and results for fungal diseases and associated trends over time, though the utility for measuring health disparities may be restricted. These datasets often contain limited demographic data, and do not include information about socioeconomic status, treatment, or health outcomes, therefore identification of associated disparities is not feasible. However, geographic and facility data points may be beneficial, particularly when used in conjunction with other data sources, such as the census or location-based population datasets, to further assess differences. For researchers who use laboratory testing data, information on diagnostic testing is detailed and often up-to-date in near real-time, though referral bias and limited insight into contributing laboratories may hinder generalizability.

Electronic Medical Record Data

Electronic medical records (EMR) are the digital versions of medical charts, including patient information such as demographics and medical history, as well as clinical information related to diagnoses, medications, and immunizations. EMR data are generated for most people (88%) who interact with the healthcare system [23], resulting in large sample sizes, though population representation and inclusion bias is dependent on the healthcare centers included in the analysis. Information is detailed, high-quality, and often available in close to real-time.

Data elements in EMR may overlap with other sources, including insurance claims, hospital discharge data, and laboratory testing. The inclusion of these data points highlights the breadth of data available in EMR, though the limitations of those data sources (e.g., poor documentation, miscoding, and changes to diagnostic testing practices) should also be taken into consideration for EMR analyses.

Accessibility to EMR data can range based on the scope of the analysis. Data may be more easily obtained for an individual institution or a narrow geographic area, but access is often more limited at a national scale or across numerous medical centers. Additionally, analysis may require rigorous and laborious data cleaning as there is no standardized data structure across EMR platforms; different systems may use different structures and terminologies for certain data points. For fungal diseases, important information is often included in free text responses (e.g., imaging report notes and physician comments), which would similarly entail substantial data cleaning prior to analysis, though advancements in machine learning and artificial intelligence may offer techniques to more efficiently mine text data.

Survey Data

Surveys offer the opportunity to ask specific questions relevant to the topic of interest and direct those questions to applicable populations (e.g., healthcare providers, patients, and certain racial and ethnic groups). They can therefore help to fill gaps in literature because survey questions can be customized to the needs of a particular analysis; other data sources often necessitate more indirect connections or assumptions to answer research questions. National surveys generally boast large samples, though surveys can also be conducted to a targeted group of participants. Some people may be more likely to respond to a given survey based on the subject matter, and this response bias can influence survey results and may impact whether or not a survey is representative of the population of interest; data are also self-reported and therefore subject to recall bias.

Patient and provider survey datasets may be more likely to contain high-quality demographic and socioeconomic data than the data sources previously mentioned, which is useful for evaluation of health disparities. Some national surveys collect data relevant to social determinants of health (e.g., household size, household income, and education) that can be leveraged as denominators for various groups. In addition, survey data may be available soon after survey completion and therefore timelier than other data sources.

Vital Statistics Data

Vital statistics capture data related to births and deaths. For IFI, death data is likely the most useful information within vital statistics records to explore associations between patient characteristics and fungal disease mortality. Vital statistics mortality data comprise the entire deceased population, but reported fungal disease-associated deaths likely underestimate the true number of fungal disease-associated deaths to varying degrees depending on the pathogen. For example, an analysis of deaths related to coccidioidomycosis, an environmental fungal disease endemic to the southwestern United States, found that coccidioidomycosis-associated deaths were underreported in death registry data by two- to sevenfold in Arizona [24]. Underreporting may be a factor of misclassification or nonspecific attribution; approximately 20–30% of death certificates are thought to have issues with completeness, meaning they lack description of a clear chain of events from the immediate to the underlying cause of death, fail to report other conditions that contributed to death, or provide nonspecific information [25]. A previous study estimated that only 50% of IFIs are diagnosed prior to death, and autopsies reached their lowest documented rate in in 2020 (7.4%) [26, 27]. Vital statistics contain fairly complete demographic data, such as sex, age, race, and ethnicity, but data elements regarding socioeconomic status or social determinants of health are limited. Notably, fungal diseases are rarely listed as a primary cause of death; researchers using vital statistics data would benefit from pulling underlying and multiple cause of death data.

Health Equity Indices

Health equity indices are geographic place-based tools that measure a combination of social, and in some cases environmental conditions that affect health at a community level. These indices comprise a variety of indicators (e.g., unemployment, vehicle access, and walkability), which are generally grouped into larger themes or domains (e.g., socioeconomic status, housing type and transportation, and built environment) and then ranked in comparison with other neighborhoods. Indices provide an overall score for each community, usually by census tract or by county, and there are often additional scores for each theme or domain for more detailed analyses.

Two such indices include the Centers for Disease Control and Prevention (CDC)/Agency for Toxic Substances and Disease Registry (ATSDR) Social Vulnerability Index (SVI) and the CDC/ATSDR Environmental Justice Index (EJI). These tools aim to help public health officials identify communities at greater risk of poor health outcomes stemming from social determinants of health and environmental burden [28, 29].

Health equity indices can be used to explore potential associations between fungal disease and health equity or environmental justice based on index scores. Their finer geographic scale compared to other datasets also allows researchers to analyze the potential effect of rural–urban status on index scores and fungal disease incidence. The data necessitate analysis at the community level (e.g., county and census tract), and must be interpreted in relative terms since the index scores are generated through comparative rankings across the unit of analysis (county or census tract). This kind of analysis also requires a separate dataset for the disease or condition of interest at the same spatial scale as the index data; notably, the EJI is only available at the census tract level, and corresponding fungal disease data may be difficult to obtain at this level of specificity. Overall, these indices can allow for mapping and detection of communities at greater risk of fungal disease based on social determinants of health and environmental burden.

Conclusions

Evidence has shown disparities in fungal infections, but the relationship between these disparities and underlying social determinants of health is unclear. Fungal disease data are limited overall, and analyses are often complicated by small sample sizes and incomplete data elements. This review aims to facilitate further exploration of fungal disease disparities in the United States by summarizing key data sources and associated considerations for analysis. A similar review of international data sources would be valuable for global research on associations between fungal infections and social determinants of health. Emerging data hubs that link data sources by patient (e.g., National Institutes of Health All of Us Research Hub) [30] or geography (e.g., Agency for Healthcare Research and Quality Social Determinants of Health Database) [31] may also represent a future direction of population health analytics that could enhance health equity research for fungal diseases. Improved understanding of the mechanisms driving fungal disease disparities is essential to identify populations at risk of poor outcomes and inform effective public health interventions.

Data Availability

No datasets were generated or analyzed during the current study.

References

1. Denning DW. Global incidence and mortality of severe fungal disease. The Lancet Infectious Diseases [Internet]. 2024 Jan 12 [cited 2024 Mar 4];0(0). Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(23)00692-8/fulltext.
2. Benedict K, Whitham HK, Jackson BR. Economic burden of fungal diseases in the United States. Open Forum Infect Dis. 2022;9(4):ofac097. [PubMed: 35350173]
3. Rayens E, Rayens MK, Norris KA. Demographic and socioeconomic factors associated with fungal infection risk, United States, 2019. Emerg Infect Dis. 2022;28(10):1955–69. [PubMed: 36149028]
4. Jenks JD, Prattes J, Wurster S, Sprute R, Seidel D, Oliverio M, et al. Social determinants of health as drivers of fungal disease. eClinicalMedicine. 2023;66:102325. [PubMed: 38053535]
5. Social determinants of health [Internet]. [cited 2024 Jan 8]. Available from: https://www.who.int/health-topics/social-determinants-of-health. Accessed 28 Nov 2023
6. Weeks WB, Chang JE, Pagán JA, Lumpkin J, Michael D, Salcido S, et al. Rural-urban disparities in health outcomes, clinical care, health behaviors, and social determinants of health and an action-oriented, dynamic tool for visualizing them. PLOS Global Public Health. 2023;3(10):e0002420. [PubMed: 37788228]
7. Tsay SV, Mu Y, Williams S, Epson E, Nadle J, Bamberg WM, et al. Burden of Candidemia in the United States, 2017. Clin Infect Dis. 2020;71(9):e449–53. [PubMed: 32107534]
8. Green BJ. Occupational Fungal Exposure in the United States. In: Viegas C, Viegas S, Gomes A, Täubel M, Sabino R, editors. Exposure to microbiological agents in indoor and occupational environments [Internet]. Cham: Springer International Publishing; 2017 [cited 2024 Jan 12]. p. 3–33. Available from: 10.1007/978-3-319-61688-9_1.
9. Seth D, Cheldize K, Brown D, Freeman EE. Global Burden of Skin Disease: Inequities and Innovations. Curr Derm Rep. 2017;6(3):204–10. [PubMed: 29226027]
10. Rauh VA, Landrigan PJ, Claudio L. Housing and health. Ann N Y Acad Sci. 2008;1136(1):276–88. [PubMed: 18579887]
11. Goldman AL, McCormick D, Haas JS, Sommers BD. Effects of the ACA's Health Insurance Marketplaces on the previously uninsured: a quasi-experimental analysis. Health Affairs. 2018;37(4):591–9. [PubMed: 29608372]
12. Moser JW, Applegate KE. Imaging and insurance: do the uninsured get less imaging in emergency departments? J Am Coll Radiol. 2012;9(1):50–7. [PubMed: 22221636]
13. Richardson A, Allen JA, Xiao H, Vallone D. Effects of race/ethnicity and socioeconomic status on health information-seeking, confidence, and trust. J Health Care Poor Underserved. 2012;23(4):1477–93. [PubMed: 23698662]
14. Selove R, Kilbourne B, Fadden MK, Sanderson M, Foster M, Offodile R, et al. Time from screening mammography to biopsy and from biopsy to breast cancer treatment among Black and White, women medicare beneficiaries not participating in a health maintenance organization. Women's Health Issues. 2016;26(6):642–7. [PubMed: 27773529]
15. Lukachko A, Olfson M. Race and the clinical diagnosis of depression in new primary care patients. Gen Hosp Psychiatry. 2012;34(1):98–100. [PubMed: 22019462]
16. Howell EA, Brown H, Brumley J, Bryant AS, Caughey AB, Cornell AM, et al. Reduction of peripartum racial and ethnic disparities: a conceptual framework and maternal safety consensus bundle. Obstet Gynecol. 2018;131(5):770. [PubMed: 29683895]
17. Bilheimer LT, Klein RJ. Data and measurement issues in the analysis of health disparities. Health Serv Res. 2010;45(5 Pt 2):1489–507. [PubMed: 21054368]
18. Magaña López M, Bevans M, Wehrlen L, Yang L, Wallen GR. Discrepancies in race and ethnicity documentation: a potential barrier in identifying racial and ethnic disparities. J Racial and Ethnic Health Disparities. 2017;4(5):812–8.

19. Dembosky JW, Haviland AM, Haas A, Hambarsoomian K, Weech-Maldonado R, Wilson-Frederick SM, et al. Indirect estimation of race/ethnicity for survey respondents who do not report race/ethnicity. Med Care. 2019;57(5):e28–33. [PubMed: 30520838]
20. Reportable Fungal Diseases by State | Fungal Diseases | CDC [Internet]. 2023 [cited 2023 Nov 26]. Available from: https://www.cdc.gov/fungal/fungal-disease-reporting-table.html. Accessed 26 Nov 2023
21. Johnson JA, Moore B, Hwang EK, Hickner A, Yeo H. The accuracy of race & ethnicity data in US based healthcare databases: a systematic review. Am J Surg. 2023;226(4):463–70. [PubMed: 37230870]
22. Benedict K, Gold JAW, Jenkins EN, Roland J, Barter D, Czaja CA, et al. Low sensitivity of international classification of diseases, tenth revision coding for culture-confirmed candidemia cases in an Active Surveillance System: United States, 2019–2020. Open Forum Infect Dis. 2022;9(9):ofac461. [PubMed: 36196298]
23. National Center for Health Statistics. FastStats. 2023 [cited 2024 Jan 17]. Electronic Medical records/electronic health records (EMRs/EHRs). Available from: https://www.cdc.gov/nchs/fastats/electronic-medical-records.htm. Accessed 26 Nov 2023
24. Jones JM, Koski L, Khan M, Brady S, Sunenshine R, Komatsu KK. Coccidioidomycosis: an underreported cause of death—Arizona, 2008–2013. Med Mycol. 2018;56(2):172–9. [PubMed: 28595294]
25. Understanding Death data quality: cause of death from death certificates. https://www.cdc.gov/nchs/nvss/covid-19.htm#understanding-death-data-quality. Accessed 18 Dec 2023
26. Dignani MC. Epidemiology of invasive fungal diseases on the basis of autopsy reports. F1000Prime Rep. 2014;6:81. [PubMed: 25343038]
27. Hoyert D. Autopsies in the United States in 2020 [Internet]. National Center for Health Statistics (U.S.); 2023 May [cited 2024 Jan 15]. Available from: https://stacks.cdc.gov/view/cdc/126588. Accessed 8 Jan 2024
28. CDC/ATSDR Social Vulnerability Index (SVI) [Internet]. 2023 [cited 2024 Jan 8]. Available from: https://www.atsdr.cdc.gov/placeandhealth/svi/index.html. Accessed 8 Jan 2024
29. CDC. Centers for Disease Control and Prevention. 2023 [cited 2024 Jan 8]. Environmental Justice Index (EJI). Available from: https://www.atsdr.cdc.gov/placeandhealth/eji/index.html. Accessed 8 Jan 2024
30. About the Research Hub – All of Us Research Hub [Internet]. [cited 2024 Jan 15]. Available from: https://www.researchallofus.org/about-the-research-hub/. Accessed 8 Jan 2024
31. Social Determinants of Health Database (Beta Version) [Internet]. [cited 2024 Jan 15]. Available from: https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html. Accessed 8 Jan 2024

Table 1.

Advantages and considerations of potential data sources for analysis of health disparities related to fungal diseases

Data Source	**Advantages**	**Considerations**
Public health surveillance data	- Systematically collected data using standardized case definitions - Useful to identify and monitor big-picture trends over time	- Data related to race/ethnicity, social determinants of health, and health outcomes can be limited - Dependent on data reported by the laboratory/healthcare provider - Only certain fungal diseases are reportable - Often geographically limited - Limited flexibility to add or modify data elements - Can lag by months to years - Surveillance practices and case definitions may change over time - Subject to diagnostic testing practices - Availability and performance characteristics of laboratory diagnostic tests may change over time
Hospital discharge data and hospital survey data	- Large sample size - Can be nationally representative - Often include race/ethnicity data, payer status (i.e., private insurance, Medicare, Medicaid), charges/costs, outcomes - May include questions related to social determinants of health	- Subject to misclassification or nonspecific medical codes - Data are skewed toward more severe invasive fungal diseases - Can lag by several years
Insurance claims data	- Large sample size - Longitudinal data related to patient encounters with the healthcare system (e.g., costs, diagnosis, treatment, outcomes)	- Subject to misclassification with medical codes - Data related to social determinants of health may be limited - Data skewed to specific insurance types
Laboratory testing data	- Detailed data related to diagnostic testing - Data are often available in close to real-time	- No data related to demographics, socioeconomic status, treatment, or health outcomes - Subject to diagnostic testing practices - Availability and performance characteristics of laboratory diagnostic tests may change over time
Electronic medical record data *	- Large sample size - Detailed and high-quality data - Data are often available in close to real-time	- Access may be difficult, particularly at a national scale or across medical centers - No standardized data structure across systems - Key data may be included as free text responses (e.g., radiology report notes, physician comments), complicating analysis
Health equity indices	- Compilation of national or state-level data on factors specifically associated with social determinants of health - Useful for community-level analysis	- Interpretation is in relative terms as index scores are based on comparative rankings across Census tracts - Analysis requires a fungal disease-related data source to use in conjunction with the index tool - Spatial scale of indices may not correspond to other data - Analysis is at the community level vs. the individual level
Survey data	- Survey questions may be updated and tailored to the needs of the analysis - High-quality data related to demographics and socioeconomic status - Data often available soon after survey completion	- Data are self-reported and subject to recall bias - Data may not be representative of the entire population of interest
Vital statistics data	- Comprise the entire deceased population - Beneficial for analysis related to mortality	- Subject to misclassification or nonspecific cause of death classification - Fungal diseases often only listed in multiple cause of death data - Limited data related to socioeconomic status or social determinants of health

* Electronic medical record data may include insurance claims, administrative hospital data, and laboratory testing data. Considerations listed for those data sources are also applicable in the context of electronic medical records.