



RESEARCH!AMERICA — RESPONSE TO NOT-RM-26-004

Due September 19, 2026

SUBMITTER NAME (50 char limit)

Eleanor Dehoney, Research!America

SUBMITTER EMAIL (100 char limit)

edehoney@researchamerica.org

TITLE OF SUBMISSION (50 char limit)

Identifying the Infectious Roots of Chronic Disease

BOX 1 — A CRITICAL CHALLENGE OR EXCITING EMERGING OPPORTUNITY (500 words)

Infection is a significant and under-recognized cause of chronic disease in Americans. A systematic review found that five agents, HIV, hepatitis C virus (HCV), *Helicobacter pylori*, hepatitis B virus, and *Chlamydia pneumoniae*, are collectively implicated in the etiology of 37 distinct chronic conditions (1). Chronic HCV infection is illustrative: up to two-thirds of chronically infected patients present with manifestations outside the liver, including vasculitis, lymphoma, cardiovascular disease, insulin resistance and type 2 diabetes, and kidney disease (2).

Multi-organ involvement is also characteristic of acute infection. SARS-CoV-2 acutely causes myocarditis, acute kidney injury, thrombosis and stroke, and neurologic injury. NIH's RECOVER Initiative found new cases of myalgic encephalomyelitis/chronic fatigue syndrome at roughly fifteen times pre-pandemic levels (3). Comparable post-acute syndromes follow Epstein-Barr virus, enterovirus, influenza, *Giardia*, dengue, and Ebola (4).

A few examples illustrate the range of infections and conditions involved. Chagas disease is estimated to affect over 300,000 people in the U.S., many of whom are undiagnosed for months, and is a known cause of heart failure, arrhythmia, and sudden cardiac death (5). Congenital cytomegalovirus occurs in over five of every 1,000 U.S. live births and is a leading nongenetic cause of sensorineural hearing loss in American children (6, 7). Three to four percent of all cancers diagnosed in U.S. adults are attributable to infectious agents (8, 9).

The central obstacle is that the initiating infection is rarely observed. For most chronic conditions with a suspected infectious contribution, the causal agent has not been identified, and a study cannot be designed around an exposure not yet linked to the outcome. Where specimens have been available, results have been striking. Serial serum samples banked from more than 10 million U.S. service members implicated Epstein-Barr virus as a leading environmental cause of multiple sclerosis, with a 32-fold increase in risk following

seroconversion and none after other viruses (10). That repository was not assembled to study multiple sclerosis. The answer came from a federal collection that happened to exist, and from the service members who supplied it.

The opportunity is to determine, systematically rather than incidentally, which infections cause chronic disease in Americans. Each pairing established converts a condition of unknown origin into one with a documented cause that can be screened for, vaccinated against, or eliminated before disease develops. Of those infection-attributable cancers, 84 percent are caused by agents that can already be vaccinated against or treated (9). The same is possible for conditions whose causes remain unidentified.

1. Orrskog S, et al. PLoS One. 2013;8(7):e68861. doi:10.1371/journal.pone.0068861
2. Cacoub P, Saadoun D. N Engl J Med. 2021;384(11):1038-1052. doi:10.1056/NEJMr2033539
3. Vernon SD, et al. J Gen Intern Med. 2025. doi:10.1007/s11606-024-09290-9
4. Choutka J, et al. Nat Med. 2022;28(5):911-923. doi:10.1038/s41591-022-01810-6
5. Nunes MCP, et al. Circulation. 2018;138(12):e169-e209. doi:10.1161/CIR.0000000000000599
6. Kenneson A, Cannon MJ. Rev Med Virol. 2007;17(4):253-276. doi:10.1002/rmv.535
7. Lieu JEC, et al. JAMA. 2020;324(21):2195-2205. doi:10.1001/jama.2020.17647
8. Islami F, et al. CA Cancer J Clin. 2024;74(5):405-432. doi:10.3322/caac.21858
9. Volesky-Avellaneda KD, et al. JAMA Oncol. 2023;9(12):1678-1687. doi:10.1001/jamaoncol.2023.4273
10. Bjornevik K, et al. Science. 2022;375(6578):296-301. doi:10.1126/science.abj8222

BOX 2 — RESOURCES, TOOLS, OR KNOWLEDGE NEEDED (499 words)

Three things are needed:

1. Knowledge of which infections are linked to which chronic conditions. The pairings established to date have largely emerged by chance, from specimens collected for unrelated reasons or from outbreaks large enough to make an association visible. For most chronic conditions in which an infectious contribution is suspected, including several with substantial prevalence in the United States, the responsible agent remains unidentified.
2. Specimens and clinical information captured at or near the time of infection. This is the clearest point at which the exposure can be documented and the pathogen identified, allowing for causal inference.
3. Common protocols that allow existing collections to be used together. NIH has already supported the collection of acute-phase specimens across many cohorts and many Institutes, Centers, and Offices (ICOs). Those collections were designed for different purposes, annotated according to different conventions, and are in practice largely invisible to investigators outside the ICO that funded them. Their combined scientific value is considerably greater than the sum of their individual contributions. Consistently annotated, these collections become a valuable resource ready for modern computational and AI analysis. Such analysis permits comparisons no individual collection can support: whether the same agent precedes conditions studied by

different Institutes, whether unrelated agents converge on shared mechanisms, whether patterns too sparse to detect in one cohort become visible across many.

No other entity can realize that value. The collections are federally funded and scattered across NIH, and what is needed is a coordinated infrastructure that no individual ICO has an incentive to build, since the benefit accrues to research beyond its own portfolio. Foundations and consortia each build around a single condition, which is how the fragmentation arose. The Common Fund is uniquely positioned to coordinate across ICOs.

BOX 3 — WHAT MAKES THIS PARTICULARLY TIMELY (499 words)

Several developments make this a timely area for Common Fund consideration.

Chronic illness following infection has become a defined scientific category. Previously separate literatures on ME/CFS, post-treatment Lyme disease, post-Giardia irritable bowel syndrome, post-dengue and post-Ebola illness, and Long COVID are increasingly understood as instances of a common phenomenon, with shared features and plausible shared mechanisms in chronic inflammation and immune dysregulation (1, 2). Questions that were previously pursued separately can now be pursued together.

NIH's RECOVER Initiative has established a scientific foundation and identified what is missing. RECOVER demonstrated that infection produces durable, multi-system chronic illness (3). Because it enrolled participants largely after infection had occurred, the initiating exposure was not documented, which is the limitation the multiple sclerosis finding escaped only because serum had been banked before anyone knew what question it would answer (4). Together, they indicate both the importance of the question and the material required to answer it.

Every outbreak has the potential to leave a wave of chronic disease in its wake, and the specimens that would allow the next one to be understood are being collected now. COVID-19 demonstrated the cost of learning this after the fact: the acute emergency was met with extraordinary speed, while the chronic consequences were investigated years later, in patients whose initiating infection could no longer be characterized. New pandemic threats are emerging more quickly than they ever have (5). Whether the response to the next one begins with usable specimens or without them is being determined by collection decisions made today.

The cost of characterizing pathogens has fallen sharply. Sequencing and serologic methods that were prohibitive a decade ago are now routine, making retrospective analysis of banked specimens feasible at a scale previously impossible. The principal remaining barrier is inconsistency in how specimens and associated data are collected, annotated, and described. In global SARS-CoV-2 genomic surveillance, the most heavily resourced sequencing effort ever mounted, 63 percent of shared sequences lacked patient demographic information, and over 95 percent lacked clinical information (6).

Advances in artificial intelligence have increased the value of well-structured data. NIH has identified artificial intelligence and real-world data among its priorities and has published a strategic plan for data science to guide this work (7, 8). Data intended for such use must be

organized accordingly, and retrofitting consistency onto existing collections becomes more difficult with time.

1. Choutka J, et al. Nat Med. 2022;28(5):911-923. doi:10.1038/s41591-022-01810-6
2. Furman D, et al. Nat Med. 2019;25(12):1822-1832. doi:10.1038/s41591-019-0675-0
3. Vernon SD, et al. J Gen Intern Med. 2025. doi:10.1007/s11606-024-09290-9
4. Bjornevik K, et al. Science. 2022;375(6578):296-301. doi:10.1126/science.abj8222
5. Marani M, Katul GG, Pan WK, Parolari AJ. Proc Natl Acad Sci USA. 2021;118(35):e2105482118. doi:10.1073/pnas.2105482118
6. Struelens MJ, et al. Front Sci. 2024;2:1298248. doi:10.3389/fsci.2024.1298248
7. NIH. Advancing NIH's Mission Through a Unified Strategy. nih.gov
8. NIH Strategic Plan for Data Science. datascience.nih.gov

BOX 4 — OTHER CONTENT (500 words)

On behalf of Research!America, a nonprofit, nonpartisan alliance representing 300 member organizations, thank you for this opportunity to submit ideas for Common Fund programs at NIH.

Relevance across NIH Institutes, Centers, and Offices. A single pathogen routinely produces disease across many organ systems, and the same specimens and clinical data therefore bear on questions belonging to many ICOs at once. Chagas cardiomyopathy and *C. pneumoniae* in atherosclerosis fall to NHLBI; infection-attributable cancers to NCI; EBV in multiple sclerosis and post-treatment Lyme disease to NINDS; HCV-associated kidney disease and post-Giardia IBS to NIDDK; congenital CMV and childhood hearing loss to NICHD and NIDCD. NIAID, NIA, NIAMS, NEI, and the data science components of NLM and ODSS have comparable stakes.

Because no individual ICO holds responsibility for the question as a whole, existing collections are designed and annotated to answer one Institute's question rather than many. This is the same structural problem the Common Fund has addressed through the Common Fund Data Ecosystem and Bridge to Artificial Intelligence, and it is a principal reason the answers obtained to date have come as much by chance as by design.

Category. This concept fits within catalytic data resources, with a secondary contribution to transformational science and discovery.

Relevance to NIH priorities. The concept is responsive to NIH priorities on chronic disease, artificial intelligence, and real-world data (1, 2). It also bears on the Make Our Children Healthy Again Strategy (3). That Strategy identifies diet, chemical exposure, physical inactivity, and chronic stress as drivers of childhood chronic disease. Infection belongs alongside them: congenital and early-life infection is a documented cause of conditions children carry for life, including the hearing loss caused by congenital CMV described above, and unlike most drivers of chronic disease an infectious cause, once identified, can be vaccinated against or treated. Several of the infections at issue, among them Chagas, dengue, and chikungunya, are best characterized where transmission is highest, which in some cases is outside the United States. The knowledge would be held and used by American researchers, and the beneficiaries are

American patients, including the service members, travelers, and Americans living abroad who encounter these infections first.

Research!America stands ready to support your efforts to identify and launch Common Fund programs that clarify the relationship between infection and chronic disease in the American population. Thank you for considering our submission.

CITATIONS

1. NIH. Advancing NIH's Mission Through a Unified Strategy. <https://www.nih.gov/about-nih/nih-director/statements/advancing-nihs-mission-through-unified-strategy>
2. NIH Strategic Plan for Data Science. <https://datascience.nih.gov/nih-strategic-plan-data-science>
3. MAHA Commission. The MAHA Strategy: Make Our Children Healthy Again. September 9, 2025. <https://www.whitehouse.gov/wp-content/uploads/2025/09/The-MAHA-Strategy-WH.pdf>