The Center for Environmental Medicine and Informatics (CEMI)
A Pre Proposal to the Discovery Challenge Seed Grants Program from Collins et al.

(1) Description of the discovery opportunity: It is now widely recognized that a significant portion of human disease is influenced by environmental exposures. An improved understanding of these links along with rapid translation of such findings into public health policy and clinical practice is critical if we are to reduce the burden of environmentally induced disease. On a national level, leadership is needed in the area of *translational research as applied to environmental medicine* and experts at SUNY ESF, SUNY Upstate Medical University, and Syracuse University are interested in and well-positioned to begin filling this gap. It is crucial to create a formal structure for collaboration centered on the application of big data, artificial intelligence, scientific computing, and informatics to pressing environmental health problems.

With support from the Discovery Challenge, we will create essential infrastructure for a self-sustaining Center for Environmental Medicine and Informatics (CEMI). Such a partnership builds directly on the recent Empire Innovation Program (EIP)-supported faculty lines in the field of Environmental Health, and we will seek an additional EIP hire with relevant expertise early in the seed period. Based on the current work and research interests of collaborators on this proposal, an initial CEMI research portfolio will include (but not be limited to): environmental links to cancer and Parkinson's disease; health disparities in morbidity and mortality and small area variations in environmental quality; as well as the relationships between environmental stressors (e.g., toxicants, discrimination) and cardiovascular disease risk. These conditions frequently involve multiple exposures (the "exposome") with multiple potential mediators (e.g., genomic and other "–omic" profiles) affecting multiple interrelated outcomes. Therefore, strong capabilities in big data analytics, machine learning, and bioinformatics are necessary because of the range, complexity, and fragmentation of the data that must be aggregated and analyzed in support of the research agenda.

During the Discovery Challenge seed period, CEMI will have three specific aims:
- **Assemble available resources to create a cyberinfrastructure system to support research teams studying the interrelationships between environmental, social, and human health correlates**. This will involve leveraging existing Hill campus resources including high performance computing and technical expertise related to the use of cutting edge informatics. This will be paired with the ability to identify, integrate, interpret, store, process, and display large disparate datasets. Emphasis will be placed on enabling the synthesis of knowledge from various sources including observational and simulated environmental, clinical, ecological, spatial, and sociodemographic data. Our focus will include the use of approaches such as machine learning, computational simulation, algorithm development, parameterization, and epidemiological modeling.
- **Facilitate interdisciplinary research collaborations between scientists, students, and environmental medicine clinicians.** It is well known that academic silos are a barrier to effective research collaborations. CEMI will bring together a group of likeminded local scholars/practitioners from varying fields, supporting research and training, and emphasizing the need for study across disciplines. Research topics will have computational components that benefit from the strong technological capabilities that CEMI will offer. Central to our longer term vision (largely post Discovery Challenge funding) is the creation of broader partnerships in which policy makers, community members and private sector participants come together in a translational research paradigm that allows knowledge to be brought rapidly into public health policy and environmental medicine.

- **Secure a sustainable external funding stream for CEMI projects post Discovery Challenge funding**. Building on our second aim, all CEMI collaborations will have an explicit goal of producing preliminary findings and new analytic methods that can be leveraged as we seek more secure, sustainable, and longer-term extramural funds. We will particularly target NIEHS/NIH opportunities for R01-level support. Discovery Challenge funding is critical for gaining momentum toward these ends. Core CEMI collaborators already have a successful history in securing analogous funding.

(2) Programmatic impacts and student positioning: Fundamentally, CEMI is a research initiative focused on building capacity and connection between faculty on the Hill campuses. As such, we have no plans to offer new courses. Instead, fostering this collaboration will link environmental scientists at ESF, medical scientists/clinicians at Upstate, and public health researchers at Syracuse University as we develop new lines of research and establish permanent and sustainable extramural support. CEMI will engage student assistants who will become members of the core team. Students will gain valuable research skills, preparing them for careers in academia or applied fields. We will foster a culture of support for professional development and research inquiry.

(3) A list of agencies, partners, and funding entities: A key goal of CEMI is to establish new and enduring collaborative relationships between Hill researchers interested in using data science and informatics to elucidate relationships between the environment and human health. Essential for this will be securing external funding at levels necessary for a program of national significance. To date, core team members have received active or past grants from major funders such as NSF, NIH, the Bill and Melinda Gates Foundation and others. In our first year, we will seek additional Empire Innovation Program (EIP) funding to recruit an established researcher with expertise in machine learning, software development, and innovative uses of big data in environmental health. We will also seek informatics research funding from agencies such as NSF and NIH within the seed period, giving us traction as we apply for even larger awards (e.g. NIH P Series) following the seed period. Finally, we will explore industry and private foundation partnerships; for example, Toyota has partnerships with environmentally-focused institutions of higher education through which they fund professional degree programs and internships. A similar health-minded multinational corporation or foundation would be a valuable partner.

Core Team Description: Our team consists of faculty at all career phases. Core environmental scientists from ESF (Collins, Leydet, Mirowsky, Razavi, & Newman) offer expertise in big data, industrial toxicant exposure, environmental modeling, health disparities, environmental flows, and environmentally mediated disease experience, with funding from NSF, USDA, DOE, DOD, and NIH/NIEHS. Core public health scientists from Syracuse (Gump, Larsen, & Kmush) offer expertise in epidemiological modeling, high performance computing, health disparities, toxicant exposure in vulnerable populations, with funding from NSF and NIH. Core medical scientists from Upstate (Morley, Middleton, Kuznetsov, & Hobart) offer expertise in health disparities, Parkinson's disease, computational systems biology analytics, sequencing, cancer genomics and transcriptomics analyses integrating with clinical and epidemiological studies and translation, with funding from NYS DOH, HRSA, NIH, and Pfizer.

(4) Expansion of current or creation of new partnerships: Discovery Challenge funding will allow us to formalize the core faculty collaborations needed for assuring an enduring research center. Collaborators will meet regularly, bringing complementary strengths: Syracuse

University partners have immediate access to high performance computing resources and leadership in approaches to studying relationships between the lived experience and disease; ESF-based scholars bring expertise in quantitative socioenvironmental science, environmental toxicology and environmental flows; and Upstate clinical partners will provide a direct link to environmental medicine practice while, at the same time, driving the research agenda toward issues with immediate clinical relevance. CEMI will provide an immediate resource for recent EIP-supported hires at ESF and Upstate. In the next round of funding, we will seek one additional EIP hire with expertise in machine learning and big data approaches as applied to environmental health. When solid external funding has been obtained, we envision providing training and collaborative opportunities to scientists, clinicians, health insurance companies, and policy makers though workshops, short courses and technical training. This will lead to sustained partnerships and will establish us as a nationally recognized leader in computational translational environmental health research.

(5) Use of ESF assets beyond the Syracuse main campus: We will expand collaborations with the NY State Department of Health/Wadsworth Center in Albany and the computational facilities at Brookhaven National Laboratory.

(6) Broader impacts:

Establishment of ESF as a leader in environmental informatics and medicine—This initiative holds the promise of enhancing ESF's reputation by establishing the College as a leader in informatics and environmental medicine. Discovery Challenge funding will provide a foundation to build on. At present, there are few programs in the nation that have developed robust environmental informatics programs with a direct relationship to the clinical practice of environmental medicine. Funding of this proposal will allow ESF and collaborators to distinguish and differentiate themselves in this regard, thus placing them at the forefront of translational environmental health research and practice. It will position ESF to become a successful player in the larger extramural funding arena, introducing a new area of research expertise and, by association, providing greater opportunities for graduate training.

Improvement in the knowledge base for environmental medicine practice by nourishing a translational research context—Creating the proposed formal collaborations will lead to high impact translational environmental health research. With clinicians as partners, CEMI will begin to bridge the traditional boundaries between scientists and clinicians such that fully evaluated and replicated discoveries can make their way into clinical practice, into the community, and back again. As partnerships expand to include policy makers and others, it will speed the rate at which scientific knowledge is embodied in patient care to promote human health.

Deeper understanding of health disparities and the burden of environmentally induced disease—As noted previously, CEMI's initial research portfolio will include studies of: environmental links to cancer and to Parkinson's disease; health disparities in morbidity and mortality; and links between environmental stressors and cardiovascular disease risk. Discovery Challenge funding will allow expansion in these and other areas, thus addressing some of the most perplexing problems at the intersection of environmental exposure and human health.

(7) Investments required: As mentioned, in addition to current EIP hires, we will seek one additional EIP hire in the next funding round with expertise in machine learning and big data approaches. We will use a portion of seed funding as a funding match. We will seek dedicated graduate student lines and associated faculty support.

(1) Feasibility & Sustainability: During the seed period, the Center for Environmental Medicine and Informatics (CEMI) has three aims that are explicitly intended to build the capacity necessary for a self-sustaining Center: (1) assemble available resources for a cyberinfrastructure system that will support research teams studying the interrelationships between environmental, social, and human health correlates; (2) facilitate interdisciplinary research collaborations between scientists, students, and environmental medicine clinicians; and (3) secure sustainable external funding.

While ambitious, these aims are feasible. Core CEMI scientists and clinicians are engaged and active across the Hill campuses. Team members bring significant individual strengths and opportunities for synergistic collaborations. A number are also already linked to current initiatives such as the ongoing EIP cluster search and the Hill collaboration on Environmental Medicine. Many team members have strong individual funding histories (i.e. NSF, NIH/NIEHS).

- Environmental scientists from ESF (Collins, Leydet, Mirowsky, Razavi, & Newman) offer expertise in data-driven socio-environmental systems analytics, and environmental flows.
- Public health scientists from Syracuse (Gump, Larsen, & Kmush) offer expertise in high performance computing, epi-modeling, and toxicant exposures in vulnerable groups.
- Medical scientists from Upstate (Morley, Middleton, Kuznetsov, Hobart, & Rosenbaum) offer expertise in health disparities, Parkinson's Disease and cancer, computational biology, clinical-epidemiological integration (including modeling), and translational research.

Initial milestones toward CEMI self-sustainability include:

- *Pre Year 1*: Develop an EIP proposal (due Jan. 2019, prior to Discovery award notification)
- *Year 1 and 2 of seed funding*:
  - Recruit a new EIP-supported, established, and well-funded faculty member with expertise in health-related big data, artificial intelligence, machine learning, and/or deep learning.
  - Expand CEMI collaborations by recruiting additional scientists and clinicians from the Hill campuses, building a strong culture of collegiality and action-relevant science.
  - Identify, begin, and complete collaborative shovel-ready research projects that will generate high quality publications; solidify research collaborations; and serve as pilot/ongoing research for new/continuing R01-level proposal development.
  - Leverage established core team collaborations in support of NIH funding requests.
- *Year 3 of seed funding*: Secure extramural funding and continue activities of years 1 and 2.
- *Post Year 3*: Submit NIH P-level proposal that will provide long term Center-level funds.

(2) Transformative potential: CEMI has transformative potential based on its interdisciplinary nature, its focus on translational research, and its reliance on informatics as a fundamental tool in tackling environmental health problems. Many such problems lie at the intersection of human-natural systems and thus require an interdisciplinary approach. The fundamental emphasis on informatics will lead to the development of new tools and technologies for understanding complexity. CEMI represents the structure needed to catalyze discovery and support emergent training paradigms for the next generation of environmental medicine researchers and clinicians.

CEMI's translational focus recognizes the need to generate actionable research capable of influencing national health policy and the practice of clinical medicine, thus moving discovery rapidly from "bench to bedside." This proposal reflects current funding priorities communicated by relevant funding agencies (NIH). As CEMI Hill campus partners become more prominent, they will gain influence in shaping the national environmental medicine research agenda.