



Synthetic Health Data Challenge Phase I Informational Webinar

Stephanie Garcia, MPH | ONC

Crystal Kallem, RHIA, CPHQ | Clinovations GovHealth

Casey Thompson, MSN, RN-BC | Clinovations GovHealth

February 2, 2021

The Office of the National Coordinator for
Health Information Technology



Introductions



**Stephanie Garcia,
MPH**

The Office of the National Coordinator
for Health Information Technology



**Crystal Kallem,
RHIA, CPHQ**

Clinovations
Government + Health



**Casey Thompson,
MSN, RN-BC**

Clinovations
Government + Health

Objectives

- Learn about ONC's efforts to build Patient-Centered Outcomes Research (PCOR) data infrastructure and capacity
- Understand the value of synthetic health data and the use of Synthea™, an open-source synthetic patient generator
- Understand the structure and requirements of the Synthetic Health Data Challenge
- Learn about tools and resources available to those interested in participating in the Synthetic Health Data Challenge

ONC Vision, Mission, and Priorities



Vision

- A health system that uses information to engage individuals, lower costs, deliver high quality care, and improve individual and population health



Mission

- Improve the health and well-being of individuals and communities through the use of technology and health information that is accessible when and where it matters most

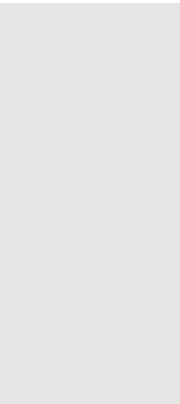
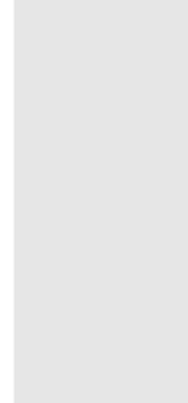


Priorities

- Interoperability of health information
- Provider burden reduction through increased usability of health information technology (health IT)

At the Intersection of Research and Care Delivery

- Lead ONC's scientific efforts and activities
- Recommend scientific policy to the National Coordinator
- Promote and lead activities that spur innovation, support patient-centered outcomes research, and advance precision medicine



<https://healthit.gov/topic/scientific-initiatives>



Patient-Centered Outcomes Research (PCOR)

- Produce evidence to inform patients', families', and their health care providers' health care decisions
- Patient-Centered Outcomes Research Trust Fund (PCORTF) managed by the Assistant Secretary for Planning and Evaluation (ASPE) supports intradepartmental projects that build data infrastructure and capacity for PCOR

[https:// www.healthit.gov/topic/scientific-initiatives/building-data-infrastructure-support-patient-centered-outcomes-research](https://www.healthit.gov/topic/scientific-initiatives/building-data-infrastructure-support-patient-centered-outcomes-research)

ONC Synthetic Health Data Project

ONC Synthetic Health Data Project

Accelerate PCOR by furthering the development of a synthetic data engine, by:

- Enhancing Synthea, an open-source synthetic patient generator developed by The MITRE Corporation
- Supporting the use of Synthea
- Validating the realism of synthetic health records generated by Synthea

<https://www.healthit.gov/topic/research-evaluation/synthetic-health-data-generation-to-accelerate-patient-centered-outcomes>



Why Synthetic Data?

- Clinical data are critical for the conduct of PCOR
- Cost, patient privacy concerns, or other legal restrictions associated with using real patient data
- Address barriers to using real patient data and speed development and testing of health innovations



PCOR Priority Use Cases

Opioids



**Complex
Care**



Pediatrics

Synthetic Health Data Challenge

Prize competition that invites a wide array of innovators, researchers, and technology developers to **create and test innovative and novel solutions** to cultivate Synthea and the synthetic data it generates

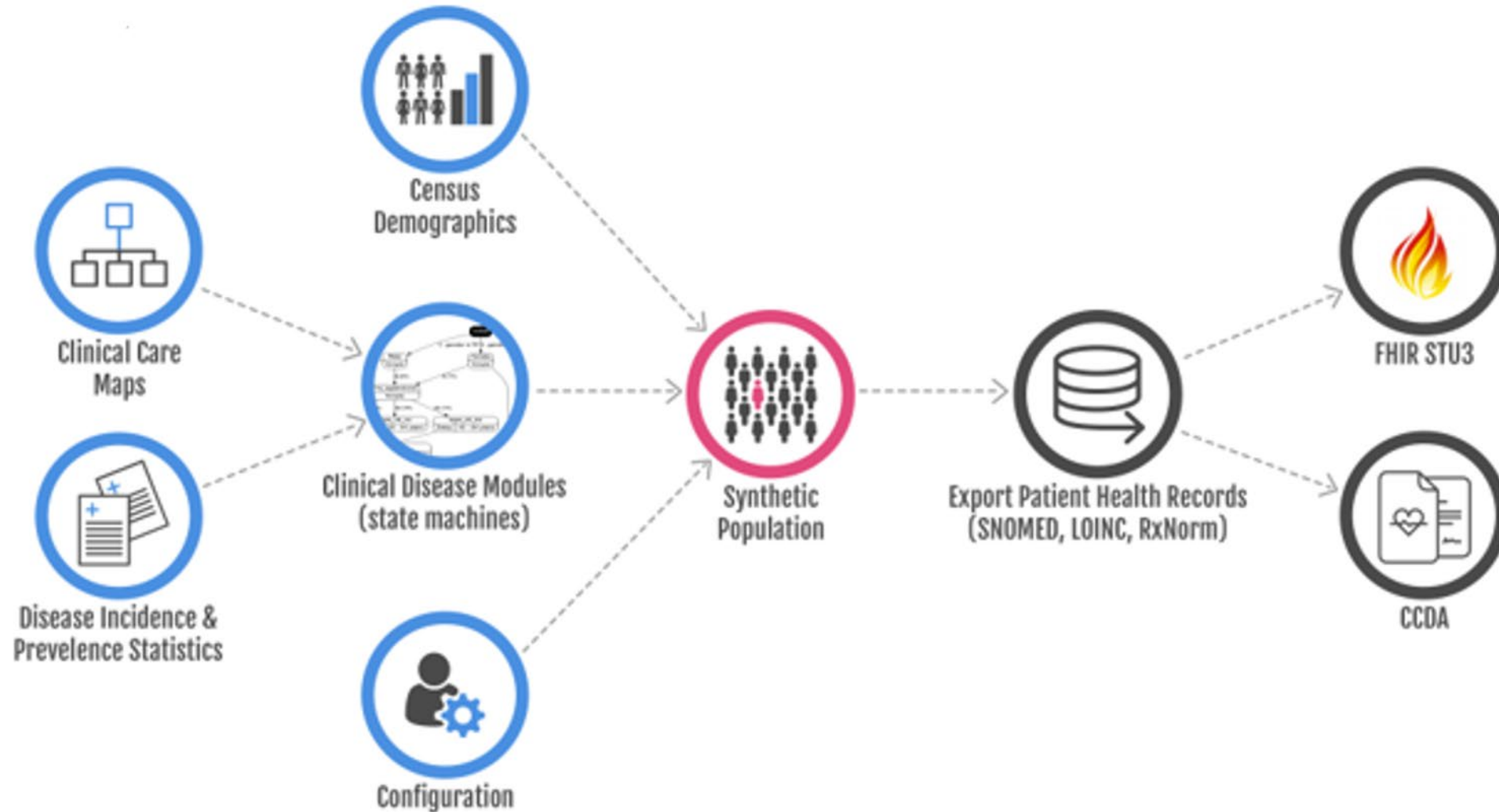
- Advance novel uses of synthetic data for PCOR research
- Validate the realism of Synthea-generated synthetic health records

About Synthea™

- Open-source synthetic patient generator developed by the MITRE Corporation
- Uses publicly available statistical data
- Models the medical history of synthetic patients over a patient's lifetime
- Outputs realistic (but not real) patient data and associated health records in a variety of formats
- Is free from cost, privacy, and security restrictions and can be used for a variety of purposes in academia, research, industry, and government

<https://synthea.mitre.org/about>

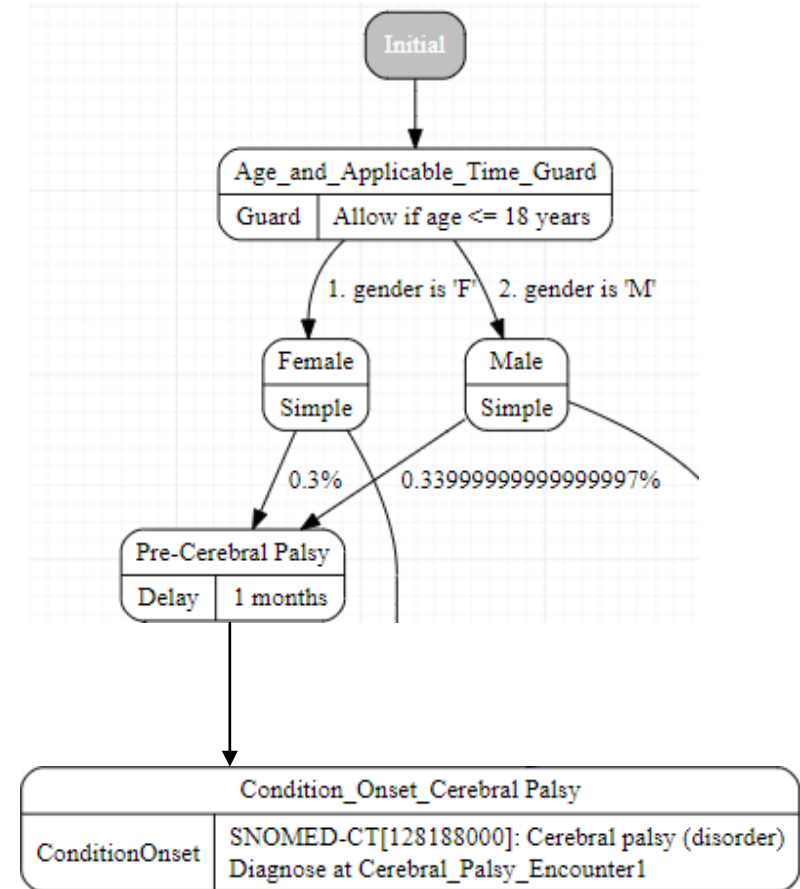
Synthea Overview



<https://github.com/synthetichealth/Synthea/wiki>

Synthea Modules

- Contains a generic framework for defining modules using JSON known as the Generic Module Framework (GMF)
- Each module describes progression of states and the transitions between them to model a clinical disease process
- States trigger conditions, encounters, medications, and other clinical events
- Synthea Module Builder is web interface to support the design of Synthea modules



<https://github.com/synthetichealth/synthea/wiki/Generic-Module-Framework>

Limitations of Synthea-Generated Synthetic Data

- Assumes all clinicians act perfectly according to the modeled standard of care
- Focuses on clinical treatments by providers performed within a clinical setting
- May not provide value when analyzing quality outcomes
- The realism of the generated data must be validated

Walonoski J, Kramer M, Nichols J, Quina A, Moesel C, Hall D, Duffett C, Dube K, Gallagher T, McLachlan S. Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. Journal of the American Medical Informatics Association. 2018 Mar 1;25(3):230-8. <https://doi.org/10.1093/jamia/ocx079>

Newly Developed Synthea Modules

Prescribing Opioids for Chronic Pain and Treatment of Opioid Use Disorder

- CDC Guidelines for Prescribing Opioids for Chronic Pain
- American Society of Addiction Medicine National Practice Guideline for Use of Medications in Treatment of Addiction Involving Opioid Use

Cerebral Palsy

- Models the treatment of sialorrhea in cerebral palsy
- American Academy for Cerebral Palsy and Development Medicine Clinical Pathway for Sialorrhea

Spina Bifida

- Models myelomeningocele, the most severe form of spina bifida
- Spina Bifida Association Care Coordination Guidelines

Sepsis

- Models the treatment of sepsis patients
- Surviving Sepsis Campaign Clinical Care Guidelines for Hour-1 Bundle



Challenge Structure and Requirements

Challenge Structure



Phase I: Proposals for Innovative Models

- Written proposal describing proposed solution
- Proposals are invited from teams or individuals
- Proposals must include methodology and intended outcomes



Phase II: Prototype/ Solution Development

- Approved proposals proceed to Phase II
- Solutions designed and tested
- Final paper describing the solution
- Video demonstration
- Evidence of validation
- Non-proprietary source code



Winning Solutions

- Total cash prize pool: **\$100,000**
- Solutions will be judged by a panel based on criteria
- Up to six (6) winners will be selected (first, second, and third place winners)

Entry Categories: Select One



Category I: **Enhancements to Synthea**

- Development or enhancement of Synthea modules
- Development of solutions that enhance or address limitations of Synthea



Category II: **Novel Uses of Synthea Generated Synthetic Data**

- Novel uses of Synthea generated data for research and technology development

Phase I Proposals Must:

- Comply with all Phase I submission requirements (registration form, cover page, written proposal)
- Specify one entry category (select one of two options)
- Use Synthea
- Include a synthetic data validation component
- Address one or more of the following use cases: opioids, pediatrics, complex care
- Agree to submit non-proprietary source code as part of Phase II

Phase I Proposal Due Date

Phase I Submission Period
closes on
March 2, 2021 5:00 PM ET

Late submissions will
not be accepted.

Phase II Solution Packages Must Include:

- Cover page, including solution title, organization, contact information, and challenge category
- Final paper describing the solution
- Five (5) minute video that demonstrates the solution and evidence of validation, uploaded to YouTube
- Supplemental documentation or attachments
- Non-proprietary source code

**Phase II
Proposal Due Date**
Phase II Submission Period
closes on
July 13, 2021 5:00 PM ET

Late submissions will
not be accepted.

Evaluation Criteria: Phase II Solutions



Criterion 1

Impact and Innovation

15 points



Criterion 2

Functionality and Implementation

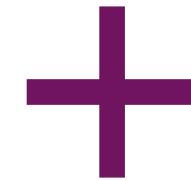
20 points



Criterion 3

Validation

15 points



Bonus Points

Evaluation Criterion 1



Impact and Innovation
15 points

- Novel, groundbreaking, and/or a creative application of an existing approach
- Enhances Synthea and/or Synthea generated synthetic data
- Impacts the field of research
- Provides new insights on the use of synthetic data for research
- Encourages the use of Synthea and/or Synthea generated synthetic data

Evaluation Criterion 2



**Functionality and
Implementation**
20 points

- Enhances Synthea functionality and/or addresses known limitations of Synthea
- Supports others with reproducing and/or reusing the solution
- Functions and aligns with the defined objectives of the solution

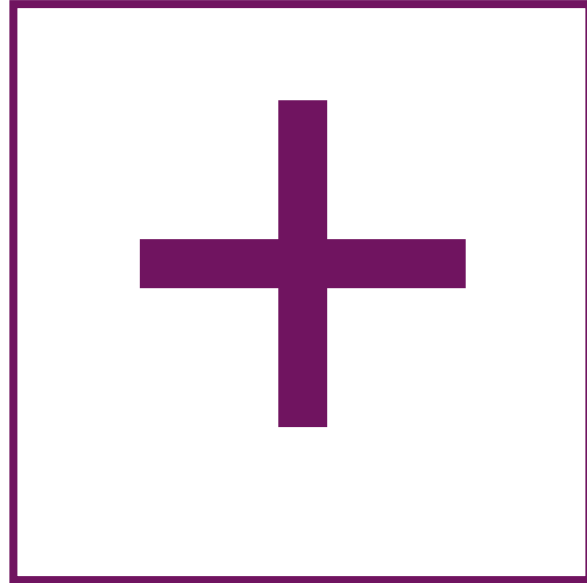
Evaluation Criterion 3



Validation
15 points

- Supports and/or improves validation capabilities of Synthea
- Improves the clinical relevance of Synthea generated synthetic data
- Supports others with reproducing and/or reusing the validation methods

Bonus Points



Up to 5 points

May be awarded for demonstrating or providing recommendations for the following, in addition to meeting challenge requirements:

- Overcoming limitations of Synthea
- Addressing validation anomalies
- Other recommendations for solutions

Prize Breakdown

Total Cash Prize Pool: **\$100,000**



First place
winner(s) will receive
\$25,000 - \$50,000

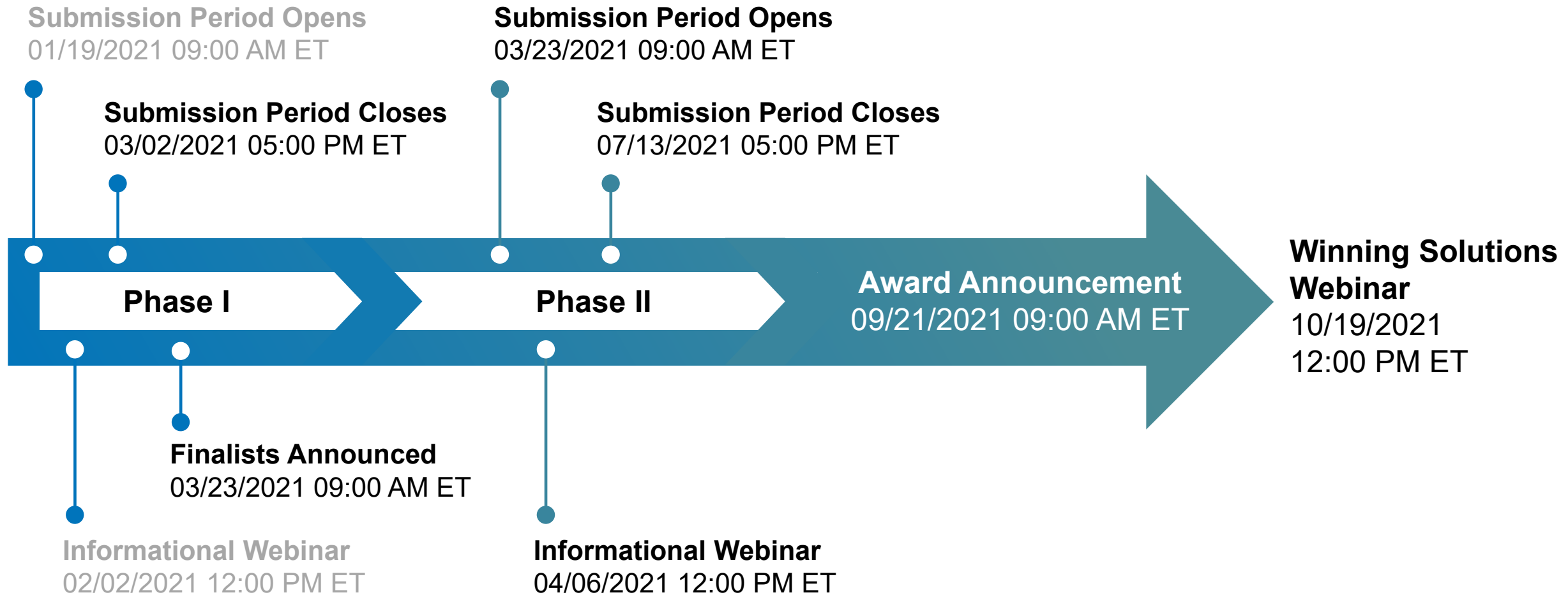


Second place
winner(s) will receive
\$15,000 - \$30,000



Third place
winner(s) will receive
\$10,000 - \$20,000

Key Dates



Resources

Challenge website: <https://www.challenge.gov/challenge/synthetic-health-data-challenge>

- Frequently Asked Questions (FAQs)
- Technical Guidance and Tips
- Modules and Companion Guides
- Synthetic Health Data Challenge Registration Form
- Phase I Informational Webinar Recording and Materials
- Support and Entry Submissions: SyntheticDataChallenge@GovHealth.com

Synthea Wiki: <https://github.com/synthetichealth/synthea/wiki>

Q&A



The Office of the National Coordinator for
Health Information Technology

Thank You

SyntheticDataChallenge@govhealth.com



Phone: 202-690-7151



Health IT Feedback Form:

<https://www.healthit.gov/form/healthit-feedback-form>



Twitter: @onc_healthIT



LinkedIn: Search “Office of the National
Coordinator for Health Information Technology”



**Subscribe to our weekly eblast
at [healthit.gov](https://www.healthit.gov) for the latest updates!**