Digital Twins for Predictive Cancer Care: an HPC-Enabled Community Initiative

Paul Macklin, Ph.D.

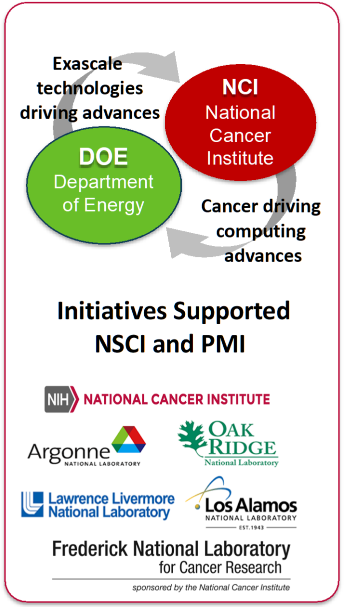
Intelligent Systems Engineering
Indiana University

November 17, 2019

History: JDACS4C and ECICC

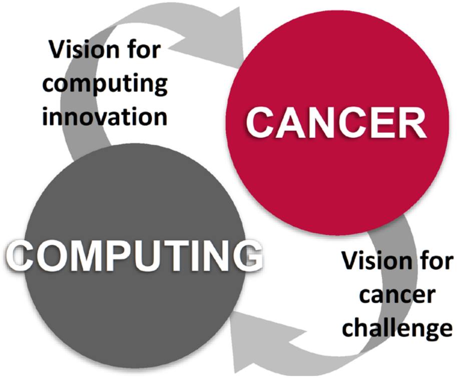
NCI-DOE Collaboration: Joint Design of Advanced Computing Solutions for Cancer (JDACS4C)

DOE-NCI partnership to advance exascale development through cancer research



Envisioning Computational Innovations for Cancer Challenges (ECICC) Scoping Meeting

March 6-7, 2019, Livermore Valley Open Campus, Lawrence Livermore National Lab



Build a Predictive Oncology Community

Join the online community!

https://nciphub.org/groups/cicc

Core Team

Emily Greenspan
National Cancer Institute
- cross-cutting cancer domain knowledge
- policy expertise
- Interagency collaborations & multidisciplinary teams

Michael Cooke
Department of Energy
- cross-cutting computing domain knowledge
- policy expertise

Amy Gryshuk
Lawrence Livermore National Lab
- cross-cutting cancer domain knowledge
- Interfacing HPC and biosciences at LLNL
- Interagency collaborations & multidisciplinary teams

Jonathan Ozik, Nicholson Collier
Argonne National Lab
- HPC domain knowledge
- large-scale model exploration on HPC

Tanveer Syeda-Mahmood, IBM
- artificial intelligence & machine vision
- big data analytics
- clinical decision support tools

Ilya Shmulevich
Institute for Systems Biology
- dynamical cancer modeling
- bioinformatics
- systems biology

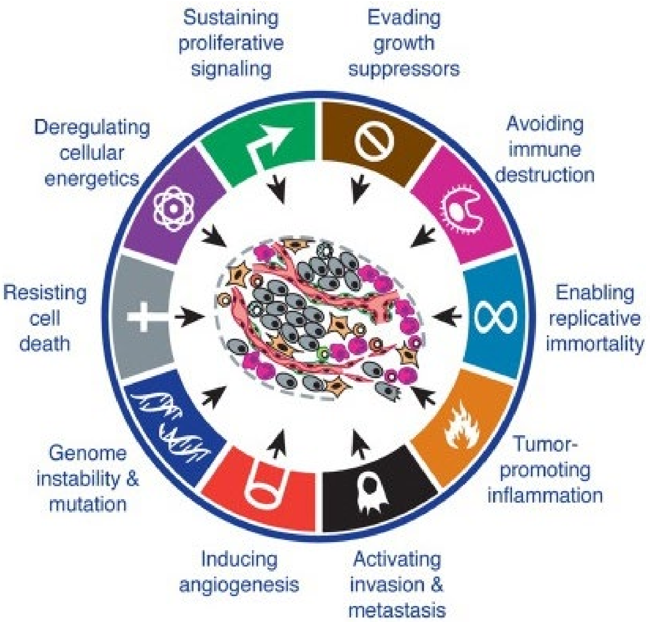
Tina Hernandez-Boussard
Stanford University
- clinical bioinformatics
- population science
- quality of care

Paul Macklin
Indiana University
- dynamical cancer models (+/- HPC)
- open source communities
- multidisciplinary teams

The need for dynamics in clinical planning

- Cancer is a **complex multiscale dynamical system:**
  - Individual cell processes and dynamics
  - Interactions between heterogeneous cells (competition and cooperation!)
  - Physical constraints (e.g., oxygen diffusion, mechanical barriers)
  - Treatments can cause **adverse systems effects:** toxicity, resistance, long-term effects
- **Precision medicine is** ultimately grounded in patient **stratification:**
  - Find the prior patients who best match my patient (e.g., by genetic profiling)
  - Treat my patient according to best practice for similar patients
- Precision medicine jumps to the endpoint and **oversimplifies the disease:**
  - Ignores multiscale dynamics and evolution
  - Cannot account for system dynamics that drive toxicity
- **Precision medicine** matches patients to prior treatment plans.
  It **cannot explore treatment variations**.
- Stratification treats the individual like "typical similar patients." It neglects a patient's personal values, access to care, and support structures.

We need ***predictive*** **medicine** to treat individual disease dynamics.



Source: Hanahan & Weinberg (2011)
DOI: 10.1016/j.cell.2011.02.013

What is *predictive medicine?*

- **Precision medicine** to date has focused on precisely matching cancer patients to the "right" treatment, based on precise individual profiling.
  - Which prior patients does this patient best match?
  - What worked best for those best matched patients?
  - **Entirely based on observables and prior measurements.**

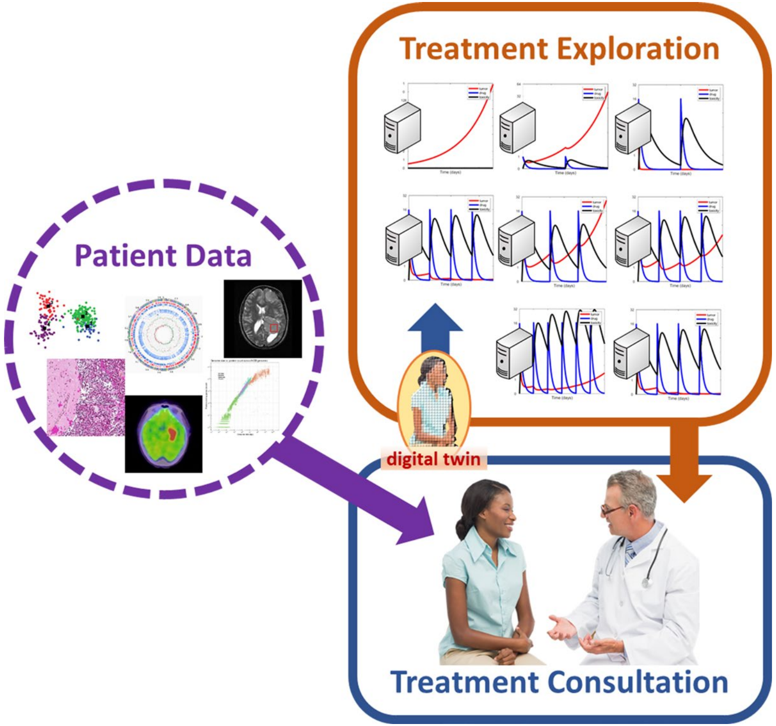
- ***Predictive* medicine** aims to predict the disease dynamics for an individual patient, based on precise individual calibration.
  - What is the expected disease course without treatment?
  - What is the expected response to a proposed treatment schedule?
  - **Integrates observables and dynamical theory.**

What is a digital twin?

- A ***digital twin*** **is synchronized digital replica** of a physical system. The digital twin is used to monitor, model, and control the real-world counterpart.
- Digital twins are used to monitor industrial devices, **fine-tune performance**, plan tasks, predict faults, and optimize maintenance schedules.
- Digital twins can be used for **virtual experiments:**
  - *What if* I run the engine hotter? *What if* I push my next service back?
- In medicine, a ***digital twin*** is a patient-tailored model that can:
  - evaluate potential therapeutic plans;
  - help choose a plan to meets personalized objectives;
  - benchmark clinical performance (virtual control);
  - continuously integrate new data and knowledge to refine treatment plans.

Digital twins could help us plan cancer care

1. Patient and clinicians discuss **treatment goals** and preferences
2. Use patient data to build a **digital twin**
3. Use HPC to try **thousands of treatment plans** on the virtual twin
4. Patient and clinicians **explore the results**:
   - Predicted response
   - Side effects
   - Long-term effects
5. **Choose** a plan
6. **Benchmark** progress against digital twin



Early community progress on key components

New technologies for patient profiling

- functional and molecular imaging
- intravital imaging (live microscopy within a patient)
- whole-slide, highly-multiplexed digital pathology
- liquid biopsies (e.g., circulating tumor cells)
- genomic profiling
- single-cell profiling (e.g., scRNA-seq for immunotyping)
- patient-derived cell cultures, organoids, and assays
- radiomics (deep learning-augmented analysis of radiology)
- fitness trackers & wearables
- implantable sensors …

Each of these technologies gives new light on a patient's health state, but it is challenging to coherently *fuse* these together to plan treatment.

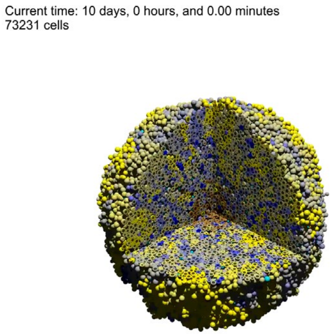
Early progress: calibrated virtual patients

There are many notable virtual cancer models for individual patients. Some examples include:

Patient tumor organoids

Macklin group, 2008-present

- Calibration of agent-based models to measurements in patient's pathology
- Recent work on multiscale organoid models
- Now researching AI-assisted coarse graining and surrogate models



Ghaffarizadeh et al. (2018)

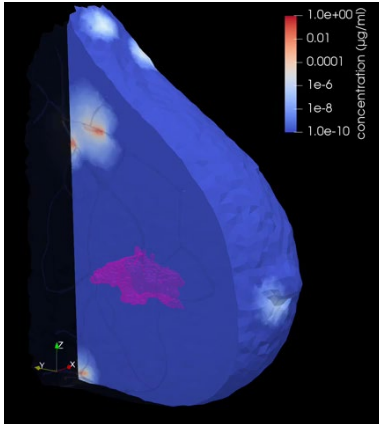
DOI: [10.1371/journal.pcbi.1005991]

Breast Cancer

Yankeelov group, 201x-present

- Calibrates PDE models of breast cancer and blood vessels to patient's imaging
- Simulates drug distribution and tumor response on HPC

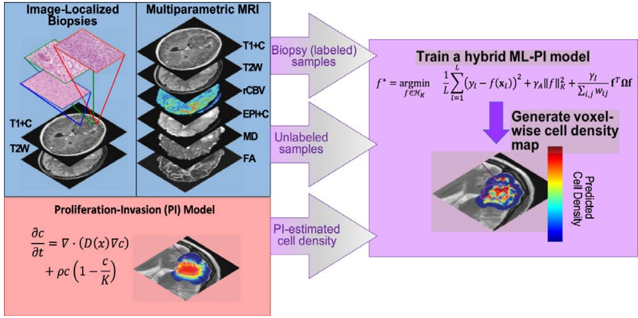
YouTube: [source here]



Glioblastoma multiforme (GBM)

Swanson group, 2000-present

- Calibrate PDE models to patient's MRIs
- Simulates a "virtual control" to benchmark patient progress, or test radiotherapies.
- Most recent work integrates simulations with machine learning

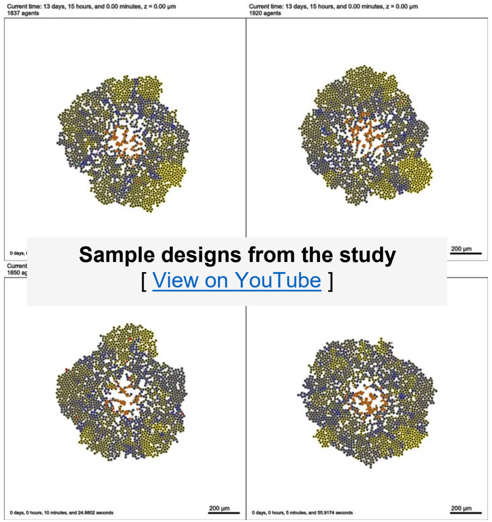


Guy et al. (2019)

DOI: [10.1038/s41598-019-46296-4]

Early Progress: HPC-driven therapy exploration

- Therapeutic planning is **exploration** in a **high-dimensional treatment space**
- We need to connect domain expertise and resources across disciplines
  - Detailed patient simulations
  - High-performance computing
  - Artificial intelligence
- Recent work combined PhysiCell + EMEWS
  - Optimize an **immunotherapy model** over 6 design parameters
  - Assess impact of clinical and biological treatment constraints
  - Use AI to choose simulations (**Cut needed simulations by 1000x**)
  - Use AI to interpret results



Sample designs from the study
[View on YouTube]

Rapid hypothesis exploration during **digital twin construction**:

- Combine simulations + HPC + AI to **rapidly test and refine** the digital twin platform

Rapid treatment exploration after **digital twin deployment:**

- AI-guided simulations on HPC to **intelligently explore treatment space**

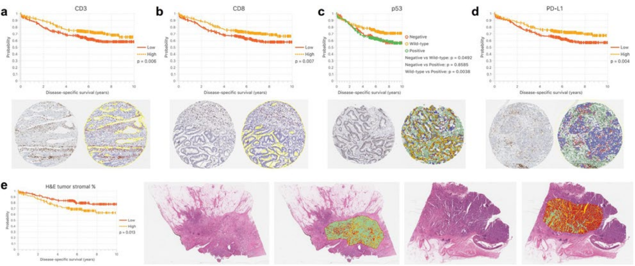
Try this model yourself!
nanohub.org/tools/pc4cancerimmune

Early progress: AI-augmented workflows

Machine learning is increasingly being used to augment research and clinical workflows. For example:

Digital pathology

- **Deep learning** for automated image **segmentation** and **annotation**



QuPath applied to colon cancer
DOI: 10.1038/s41598-017-17204-5

- **CNNs** for **virtual immunostaining**
  e.g., 10.1038/s41523-018-0084-4

- Deep Learning for **feature extraction** and **biomarker discovery**

Clinical support

- **Natural language processing** for **cancer staging** from path. reports
  e.g., 10.3233/SHTI190515

- **NLP** to assess **adherence** to clinical guidelines (for bone scan use)
  e.g., 10.1016/j.jbi.2019.103184

- **NLP** on electronic health records (EHRs) to assess **pain management**
  e.g., 10.1371/journal.pone.0210575
  e.g., 10.1177/1460458219881339

- **Regression** and **clustering** for post-operative pain **trajectory analysis**
  e.g., 10.1177/1460458219881339

- **NLP** and **regression** analysis to assess **quality of care**
  e.g., 10.5334/egems.307

Simulation workflows

- **Bayesian parameter estimates** for simulations, model inference, and UQ
  e.g., 10.1007/s00285-018-1208-z
  e.g., 10.1098/rsif.2018.0943

- **Deep neural networks** to auto-tune simulation parameters
  e.g., https://arxiv.org/abs/1910.14620

- **Surrogate models** to accelerate parameter sweeps & **optimization**
  e.g., 10.1016/j.biosystems.2019.05.005
  e.g., https://arxiv.org/abs/1910.01258
  e.g., 10.1089/cmb.2018.0168

- **CNNs**, **RNNs**, or **autoencoders** to replace/accelerate **sub-components**
  e.g., https://arxiv.org/abs/1910.01258
  e.g., https://arxiv.org/abs/1910.07291

What’s next?
Where is this heading?

ECCIC progress to date and next steps

Cancer Challenges & Advanced Computing MicroLabs

- virtual interactive events, ongoing

***1st MicroLab:* June 11, 2019:** [https://ncihub.org/groups/cicc/pastmeetings/sept25thmicrolab]

- Discussed ideas and challenges relating 4 Cancer Challenge Areas
  - *Synthetic Data Generation*
  - *Hypothesis Generation using Machine Learning*
  - ***Digital Twin Technology***
  - *Adaptive Cancer Treatments*

***2nd MicroLab:* Sept 25, 2019:** [https://ncihub.org/groups/cicc/pastmeetings/sept25thmicrolab]

- Developed use cases and persona through the lens of the 4 Cancer Challenge Areas

ECICC Ideas Lab: *planned for June 2020*

- 5-day **immersive event** to develop innovative research proposals

$$\{\text{predictive oncology}\} \cap \{\text{advanced computing}\}$$

- Call for applications will be forthcoming

Building a national forecasting resource …

- No single group, single organization, or single discipline has all the pieces to build, validate, and deploy digital twins.
- We need to ***combine our efforts*** to build a **national resource** that can continuously be improved. ***We hope you will join us!***

ECICC Community:
nciphub.org/groups/cicc

Dynamical multiscale models

Integrated artificial intelligence (AI)

State-of-the-art patient data

Critical HPC and data infrastructure

Clinical trial and practice expertise

Usable tools for research and clinical care

The digital twins will need help from many areas.
Please join the community to pitch in!

Dynamical multiscale models
- Molecular-scale networks
- Cellular behaviors and heterogeneity
- Whole-body cancer cell trafficking
- Whole-body drug kinetics & response
- Efficient multiscale coupling

State-of-the-art patient data
- Genomic profiling
- Single-cell profiling
- Novel bioengineered cultures
- Lifestyle data / telemetry
- (Molecular) functional imaging

Clinical trial and practice expertise
- Determine best use cases for prototypes
- State-of-the-art, multi-site trial protocols

ECICC Community:
nciphub.org/groups/cicc

Integrated artificial intelligence (AI)
- Sensor / data fusion
- Hypothesis generation and testing
- Patient calibration / data assimilation
- Model acceleration (e.g., via surrogates)
- Model analysis (including validation, UQ)

Critical HPC and data infrastructure
- HPC-accelerated machine learning
- High-throughput model exploration
- Secure data storage

Usable tools for research and clinical care
- Securely connect patient data
- Connect researchers and clinicians to data, models, and compute resources
- Clearly present predicted data (UX, HCI!)