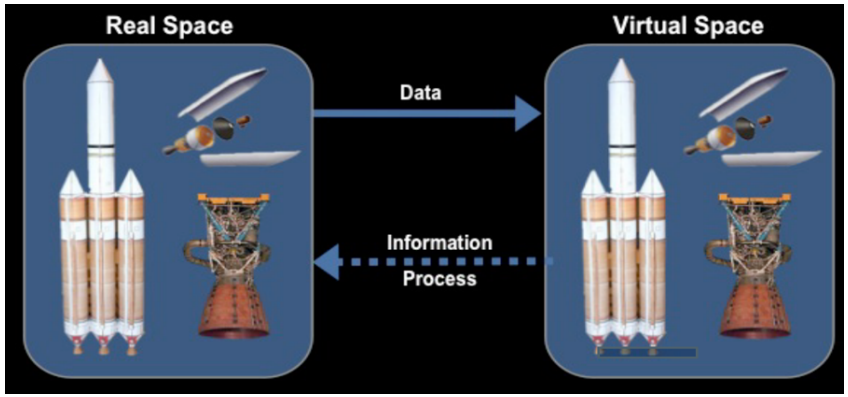
Statistical Challenges of Cardiac Digital Twins

Richard Wilkinson

University of Nottingham

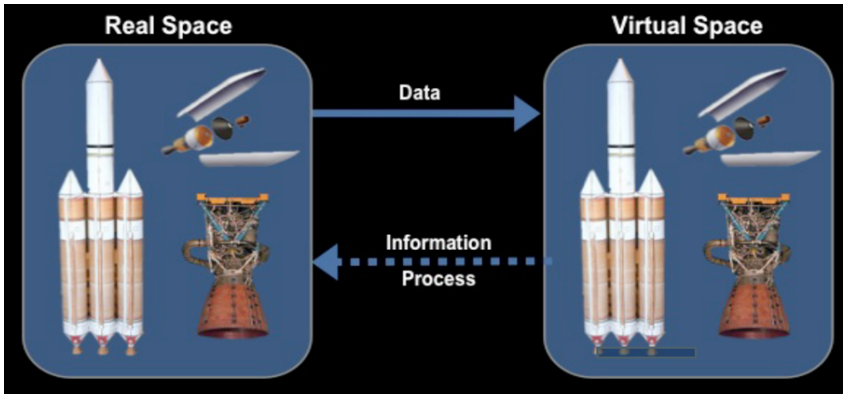
Digital twins

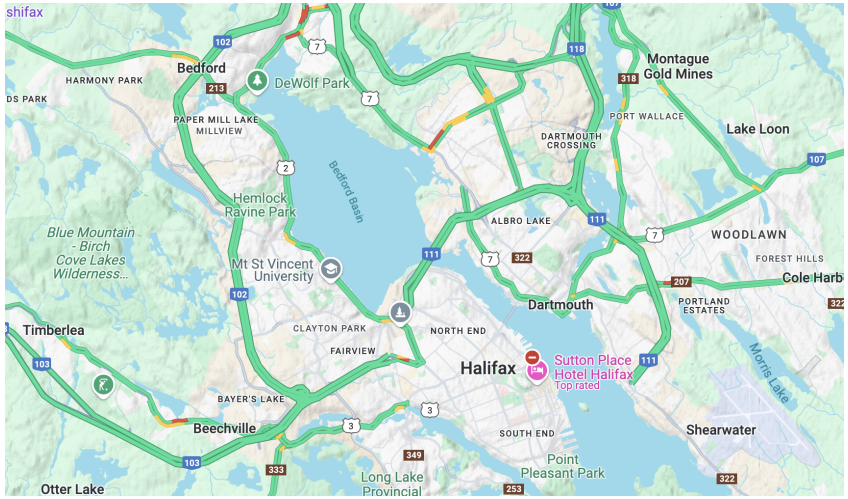
A set of virtual information constructs that mimics the structure, context and behaviour of an individual or unique physical asset, that is dynamically updated with data from its physical twin throughout its life-cycle that informs decisions that realise value.



Digital twins

~~*A set of virtual information constructs that mimics the structure, context and behaviour of an individual or unique physical asset, that is dynamically updated with data from its physical twin throughout its life-cycle that informs decisions that realise value.*~~





A model of an individual, informed by data, that influences decisions.

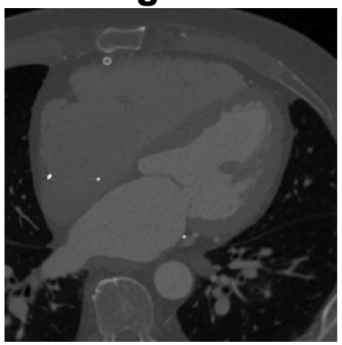
Cardiac physiology

With Steve Niederer, Richard Clayton, Sam Coveney, Cesare Corrardo, Chris Lanyon, Fay Frost, Mariya Mamiwajala, Marina Strocchi, ...

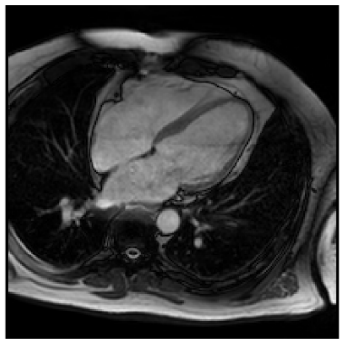
Aim: move from treatment based on guidelines derived from heterogeneous patient groups, to treatment tailored to individual patients based on their data.

Imaging

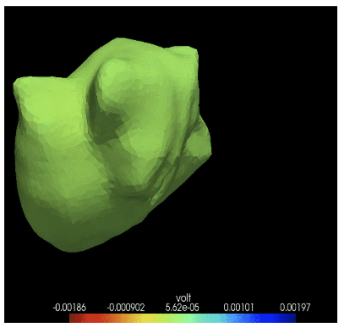
ECG-gated CT



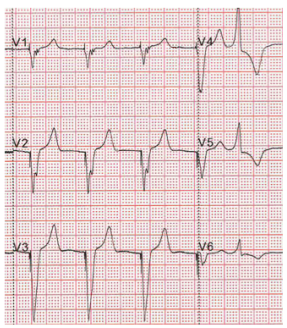
Cardiac MRI



Atrial voltage

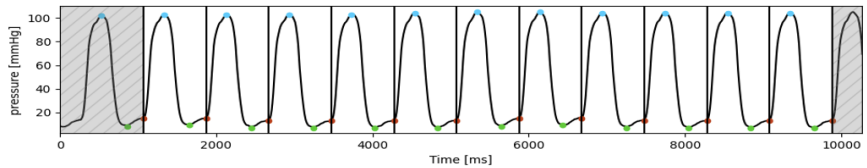


ECGs

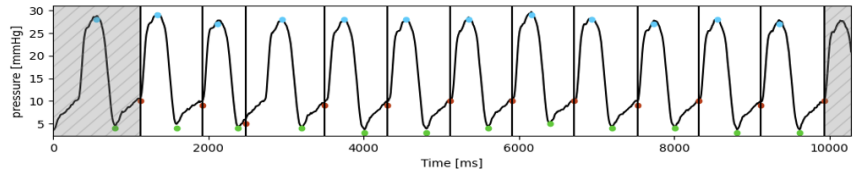


Pressure measurements

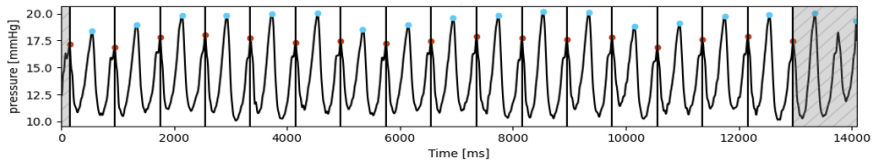
Left ventricle



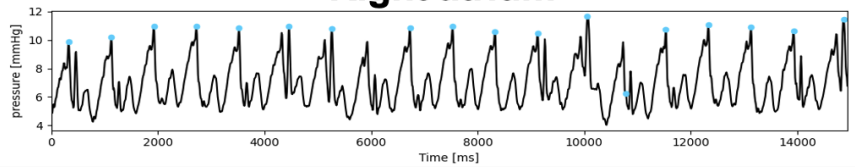
Right ventricle



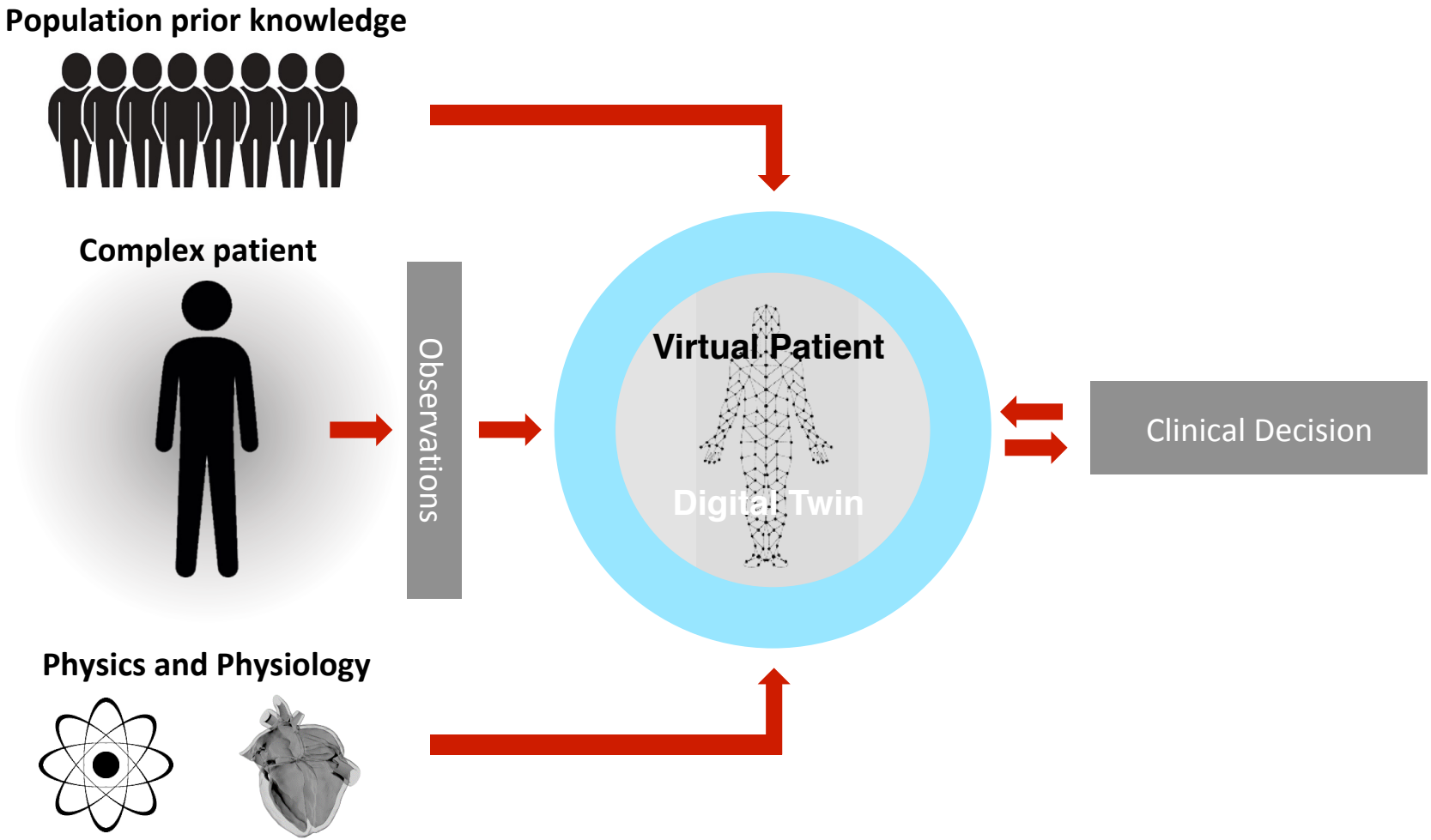
Left atrium



Right atrium

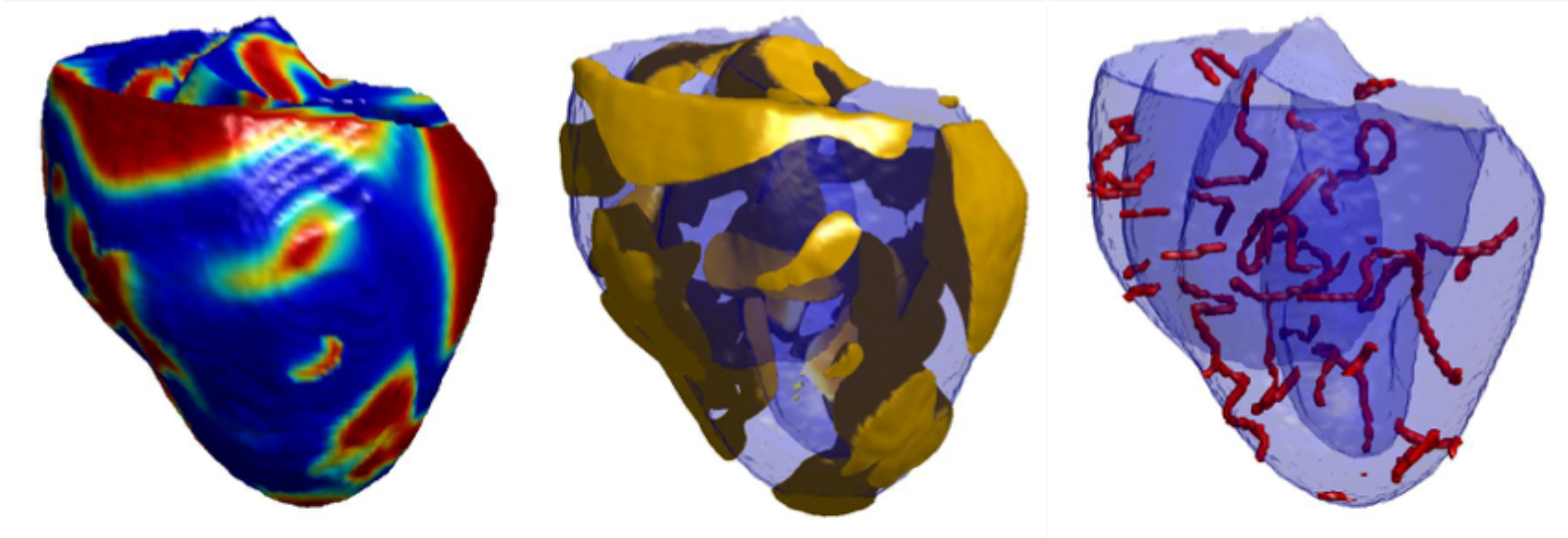


Cardiac digital twin



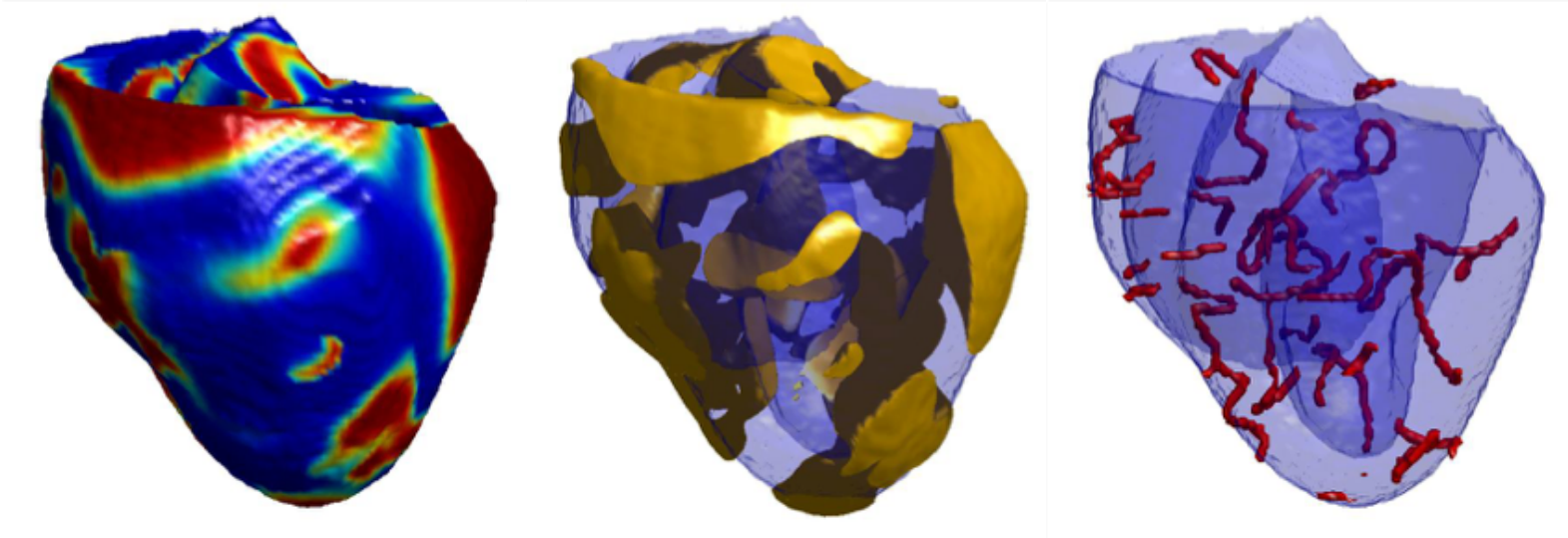
But how confident are we in our prediction

Example 1: Atrial fibrillation



Atrial fibrillation (AF) - rapid and uncoordinated electrical activation (arrhythmia) leading to poor mechanical function.

Example 1: Atrial fibrillation



Atrial fibrillation (AF) - rapid and uncoordinated electrical activation (arrhythmia) leading to poor mechanical function.

- Affects around 1,000,000 people in UK.
- Catheter ablation removes/isolates pathological tissue that sustain/initiate AF.
- 40% of patients subsequently experience atrial tachycardia (AT).

Patient Specific Cardiac Models

Aim: predict whether ablation will successfully treat an AF, by infering reentry pathways, and guiding the surgical ablation to treat for both AF and AT in a single procedure.

- Each intervention: 6% risk of major complication; cost $\sim$£10k.

Patient Specific Cardiac Models

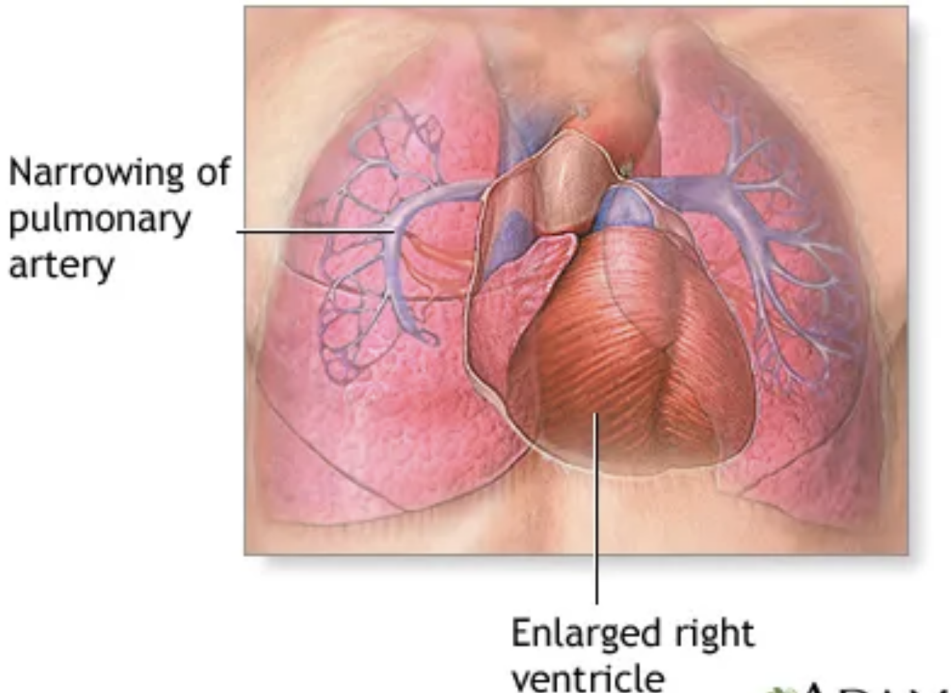
Aim: predict whether ablation will successfully treat an AF, by infering reentry pathways, and guiding the surgical ablation to treat for both AF and AT in a single procedure.

- Each intervention: 6% risk of major complication; cost $\sim$£10k.

Cardiac models at forefront of personalised modelling

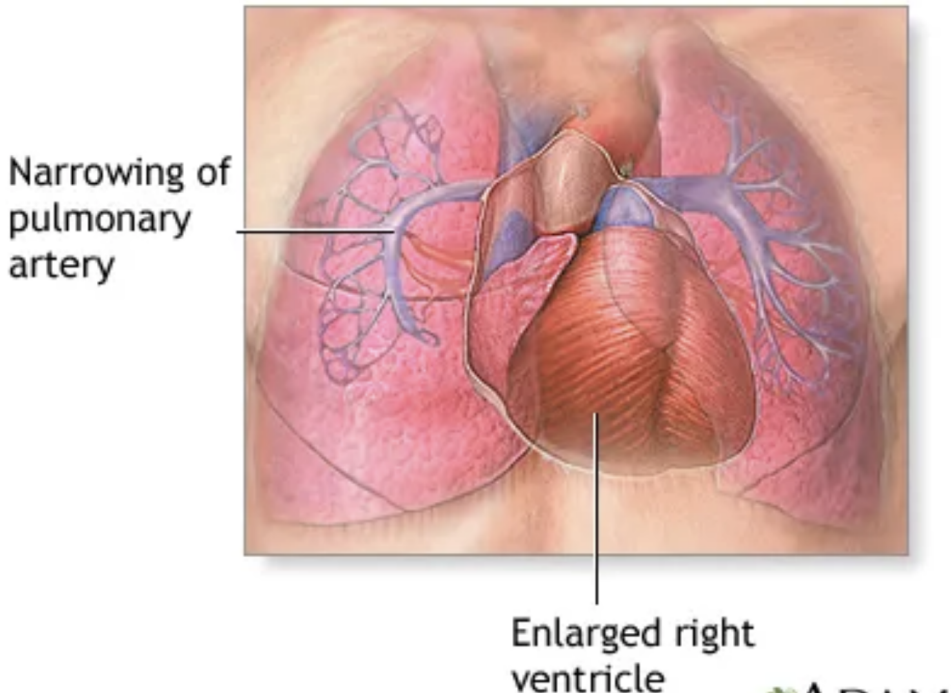
- Models are deterministic but clinical diagnosis is rarely definitive
  - uncertainty quantification/statistics challenge
- aim to consider costs and benefits across all potential outcomes weighted by their probability.

Example 2: Pulmonary arterial hypertension (PAH)



- Blood vessels in the lungs narrow, making the right heart work harder to pump blood.
- The heart compensates and the right ventricle grows, often leading to heart failure.
- $\sim$ 8000 PH patients in the UK (diagnosis is hard) affecting people of all ages.
- Life expectancy is $\sim$ 2 years if untreated.

Example 2: Pulmonary arterial hypertension (PAH)



- Blood vessels in the lungs narrow, making the right heart work harder to pump blood.
- The heart compensates and the right ventricle grows, often leading to heart failure.
- $\sim$ 8000 PH patients in the UK (diagnosis is hard) affecting people of all ages.
- Life expectancy is $\sim$ 2 years if untreated.

Example 2: Pulmonary arterial hypertension (PAH)

CDT-Net 2024-2029

Working with the UK pulmonary hypertension centres, we aim to develop digital twins of PAH patients.
Our aim is to determine if digital twins are useful clinical tools.

Example 2: Pulmonary arterial hypertension (PAH)

CDT-Net 2024-2029

Working with the UK pulmonary hypertension centres, we aim to develop digital twins of PAH patients.
Our aim is to determine if digital twins are useful clinical tools.

- Augment the usual data collected on patients with data from pressure sensors implanted into the right ventricle giving real time measurements.
- Use mechanistic models to estimate internal physiological condition.
- Use machine learning models to link routine clinical variables with inferred physiological state, and forecast future disease progression and guide treatment.

Statistical challenges

For a given patient, we want to select a model from our class of models $f(\theta, \omega)$ where

- ω are directly observable parameters specific to the patient such as geometry (ie for the computational mesh)
- θ are patient specific model parameters, eg diffusion parameters, which may be spatially varying ($\theta(x)$ for $x \in \omega$).

Statistical challenges

For a given patient, we want to select a model from our class of models $f(\theta, \omega)$ where

- ω are directly observable parameters specific to the patient such as geometry (ie for the computational mesh)
- θ are patient specific model parameters, eg diffusion parameters, which may be spatially varying ($\theta(x)$ for $x \in \omega$).

Given data D we want to solve the inverse problem

$$D = f(\theta, \omega) + e$$

to estimate

$$\pi(\theta, \omega \mid D) \propto \pi(\theta, \omega)\pi(D \mid \theta, \omega)$$

Statistical challenges

For a given patient, we want to select a model from our class of models $f(\theta, \omega)$ where

- ω are directly observable parameters specific to the patient such as geometry (ie for the computational mesh)
- θ are patient specific model parameters, eg diffusion parameters, which may be spatially varying ($\theta(x)$ for $x \in \omega$).

Given data D we want to solve the inverse problem

$$D = f(\theta, \omega) + e$$

to estimate

$$\pi(\theta, \omega \mid D) \propto \pi(\theta, \omega)\pi(D \mid \theta, \omega)$$

Many of the statistical challenges familiar from UQ, but (cardiac) DTs also present new challenges.

Statistical challenges

In practice we need to be pragmatic

- Complex simulator and limited computational resource
- Large number of unknowns θ, ω, f
- Sparse noisy data
- Misspecification/discrepancy

Statistical challenges

In practice we need to be pragmatic

- Complex simulator and limited computational resource
- Large number of unknowns θ, ω, f
- Sparse noisy data
- Misspecification/discrepancy

$$\mathbb{P}(\text{Event}|D) = \int \mathbb{P}(E|\theta, \omega, f)\pi(\theta, \omega, f|D)d\theta d\omega df$$

where

$$\pi(\theta, \omega, f|D) \propto \pi(D|\theta, \omega, f)\pi(\theta)\pi(\omega)\pi(f)$$

Statistical challenges

In practice we need to be pragmatic

- Complex simulator and limited computational resource
- Large number of unknowns θ, ω, f
- Sparse noisy data
- Misspecification/discrepancy

$$\mathbb{P}(\text{Event}|D) = \int \mathbb{P}(E|\theta, \omega, f)\pi(\theta, \omega, f|D)d\theta d\omega df$$

where

$$\pi(\theta, \omega, f|D) \propto \pi(D|\theta, \omega, f)\pi(\theta)\pi(\omega)\pi(f)$$

We need to characterize variability at the

- population level $\pi(\theta), \pi(\omega)$ etc
- individual level $\pi(\theta, \omega, f, ...|D)$ – may need to be partially done in real time
- and the physics/simulator $\pi(D|\theta, \omega, f)$

Surrogate models

If f is slow/costly to evaluate standard methods such as MCMC are impracticable.

Surrogate models

If f is slow/costly to evaluate standard methods such as MCMC are impracticable. We can use surrogate models/emulators of f, e.g.

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot))$$

which are trained on a small ensemble of simulator evaluations $C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$

- Currently run $\sim$1000 simulations for each new patient. Cost of £4-16k per patient.

We can then use the surrogate to estimate parameters etc

Surrogate models

If f is slow/costly to evaluate standard methods such as MCMC are impracticable. We can use surrogate models/emulators of f, e.g.

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot))$$

which are trained on a small ensemble of simulator evaluations $C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$

- Currently run $\sim$1000 simulations for each new patient. Cost of £4-16k per patient.

We can then use the surrogate to estimate parameters etc
Note that this adds an additional uncertainty

$$\pi(f|C)$$

Other methods: NNs (e.g. PINNs), polynomial chaos, ROM, POD etc.

Compact representation

If θ is high dimensional, we need to find a subset or transformation of the parameters $A\theta$ that we can estimate

- mesh used to simulate atrial electro-physiology has $\sim 30,000$ nodes, with 5 spatially varying parameters

Compact representation

If θ is high dimensional, we need to find a subset or transformation of the parameters $A\theta$ that we can estimate

- mesh used to simulate atrial electro-physiology has $\sim 30,000$ nodes, with 5 spatially varying parameters

Typical methods

- Global sensitivity analysis: select a subset of the most important parameters (re contribution to variance).
- Basis expansions

$$\theta = \sum_{i=1}^{k} z_i \psi_i$$

where $k \ll dim(\theta)$ and ψ_i are basis vectors to be chosen

  - Imaging data, random projection, PCA/KL, active subspace methods...

Compact representation

If θ is high dimensional, we need to find a subset or transformation of the parameters $A\theta$ that we can estimate

- mesh used to simulate atrial electro-physiology has $\sim 30,000$ nodes, with 5 spatially varying parameters

Typical methods

- Global sensitivity analysis: select a subset of the most important parameters (re contribution to variance).
- Basis expansions

$$\theta = \sum_{i=1}^{k} z_i \psi_i$$

  where $k \ll dim(\theta)$ and ψ_i are basis vectors to be chosen
  - Imaging data, random projection, PCA/KL, active subspace methods...

Given the cost of forward evaluation, how should we choose A so that θ is identifiable?

- Trade-off with dimension: accuracy, emulation, and identifiability.

Non-identifiability

The huge number of parameters, sparse data, and limited computational power mean we can't hope to estimate everything.
How can we identify non-identifiabilities?

Non-identifiability

The huge number of parameters, sparse data, and limited computational power mean we can't hope to estimate everything.
How can we identify non-identifiabilities?

- Difference between training and prediction tasks. We use data D

$$D = h_1 f(\theta, \omega) + e$$

to estimate $A\theta$.
But suppose our prediction task is then

$$h_2 f(\theta, \omega)$$

How should we choose projection A?

Fast and/or cheap inference

We want to calibrate in (close to) real time

- Catheter ablation: every additional 10mins of surgery increases stroke risk by x%

Fast and/or cheap inference

We want to calibrate in (close to) real time

- Catheter ablation: every additional 10mins of surgery increases stroke risk by x%

Even using a surrogate (which can be trained prior to surgery), MCMC can be too expensive to use in-procedure.

- We need cheaper approximate inference methods.

Fast and/or cheap inference

We want to calibrate in (close to) real time

- Catheter ablation: every additional 10mins of surgery increases stroke risk by x%

Even using a surrogate (which can be trained prior to surgery), MCMC can be too expensive to use in-procedure.

- We need cheaper approximate inference methods.

Distinguish between

- Case-based inference: for each new dataset D, run a separate optimization to approximate $\pi(\theta|D)$.
- Amortized inference: global upfront training (before data collected) using simulations, so that inference at test time is rapid.

Case based inference

For each new dataset, D, solve the inference problem (e.g. via MCMC).

- Kalman sampling methods:
  - Small ensemble of particles $\{\theta_i^t\}_{i=1,\ldots,n}$. At each iteration $(t = 1, \ldots, T)$, forward simulate, then adjust using a Kalman update.
  - Compute mean and variance for a Gaussian approximation of $p(\theta|D)$.

Case based inference

For each new dataset, D, solve the inference problem (e.g. via MCMC).

- Kalman sampling methods:
  - Small ensemble of particles $\{\theta_i^t\}_{i=1,\ldots,n}$. At each iteration $(t = 1, \ldots, T)$, forward simulate, then adjust using a Kalman update.
  - Compute mean and variance for a Gaussian approximation of $p(\theta|D)$.
- Variational inference: instead of sampling, find variational approximation $q_\phi(\theta)$ to the posterior
  - E.g., mean field approximation $q_\phi(\theta) = N(\mu, \text{diag}(\sigma^2))$

Case based inference

For each new dataset, D, solve the inference problem (e.g. via MCMC).

- Kalman sampling methods:
  - Small ensemble of particles $\{\theta_i^t\}_{i=1,\ldots,n}$. At each iteration $(t = 1, \ldots, T)$, forward simulate, then adjust using a Kalman update.
  - Compute mean and variance for a Gaussian approximation of $p(\theta|D)$.
- Variational inference: instead of sampling, find variational approximation $q_\phi(\theta)$ to the posterior
  - E.g., mean field approximation $q_\phi(\theta) = N(\mu, \text{diag}(\sigma^2))$
  - Solve

$$\arg\min_\phi KL(q_\phi(\theta)||p(\theta|D)) = \arg\min \mathbb{E}_{q(\theta)} p(D, \theta) - \mathbb{E}_{q(\theta)} \log q(\theta)$$

  - Can be minimized using stochastic gradient descent within a variational auto-encoder (VAE) framework

Amortized inference

Train a model that predicts $p(\theta|D)$ for any D:
Large upfront cost, rapid test time inference.

- Conditional VAE. Assume

$$q_\phi(\theta|D) = N(m_\phi(D), s^2_\phi(D))$$

where m_ϕ and s^2_ϕ are pre-trained neural networks.

Amortized inference

Train a model that predicts $p(\theta|D)$ for any D:
Large upfront cost, rapid test time inference.

- Conditional VAE. Assume

$$q_\phi(\theta|D) = N(m_\phi(D), s^2_\phi(D))$$

  where m_ϕ and s^2_ϕ are pre-trained neural networks.

- Neural posteriors. Eg use a normalizing flow:
  - Find invertible f such that

$$\theta \sim p(\theta|D) \iff f(\theta; D) \sim N(0, I)$$

    then $f^{-1}(z; D) \sim p(\theta|D)$ when $z \sim N(0, I)$.
  - Model f as an invertible NN with easily computable Jacobian.
  - Can include an additional summary network $S : D \mapsto \mathbb{R}^p$ to learn optimal summary $p(\theta|S(D))$

Scalable DTs

At the moment, we create a new surrogate model for each new patient, e.g. estimating ω from imaging data

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot)) \text{ trained with } C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$$

How can we reduce this cost?

Scalable DTs

At the moment, we create a new surrogate model for each new patient, e.g. estimating ω from imaging data

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot)) \text{ trained with } C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$$

How can we reduce this cost?

- Learn a statistical shape model $\omega = \sum_{i=1}^{M} z_i \phi_i$ for small M, e.g. via PCA and include z in the inputs to the surrogate.

Scalable DTs

At the moment, we create a new surrogate model for each new patient, e.g. estimating ω from imaging data

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot)) \text{ trained with } C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$$

How can we reduce this cost?

- Learn a statistical shape model $\omega = \sum_{i=1}^{M} z_i \phi_i$ for small M, e.g. via PCA and include z in the inputs to the surrogate.
- Learn latent features l_i for each patient eg using GP-LVM type ideas:

$$f(\theta, l) \sim GP(m(\theta, l), k)$$

Scalable DTs

At the moment, we create a new surrogate model for each new patient, e.g. estimating ω from imaging data

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot)) \text{ trained with } C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$$

How can we reduce this cost?

- Learn a statistical shape model $\omega = \sum_{i=1}^{M} z_i \phi_i$ for small M, e.g. via PCA and include z in the inputs to the surrogate.
- Learn latent features l_i for each patient eg using GP-LVM type ideas:

$$f(\theta, l) \sim GP(m(\theta, l), k)$$

- Learn the discrepancy from a set of reference heart simulations to the new heart

$$f(\cdot, \omega') = f(\cdot, \omega^r) + \delta(\cdot)$$

Scalable DTs

At the moment, we create a new surrogate model for each new patient, e.g. estimating ω from imaging data

$$f(\cdot, \omega) \sim GP(m(\cdot), k(\cdot, \cdot)) \text{ trained with } C = \{\theta_i, f(\theta_i, \omega)\}_{i=1}^{n}$$

How can we reduce this cost?

- Learn a statistical shape model $\omega = \sum_{i=1}^{M} z_i \phi_i$ for small M, e.g. via PCA and include z in the inputs to the surrogate.
- Learn latent features l_i for each patient eg using GP-LVM type ideas:

$$f(\theta, l) \sim GP(m(\theta, l), k)$$

- Learn the discrepancy from a set of reference heart simulations to the new heart

$$f(\cdot, \omega') = f(\cdot, \omega^r) + \delta(\cdot)$$

- Learn diffeomorphism: hearts are topologically equivalent. If $\omega' = T\omega^r$, can we learn a T' from T such that $f(\cdot, \omega') = T' f(\cdot, \omega^r)$?

Not clear *a priori* which approach, if any, will work best.

Networked Digital Twins

CDT-Net 2024-2029

Suppose we have digital twins of 1000s of patients.

- How we we learn informative priors? Stratified priors?
- How do we develop models of disease progression?
  - If patient A and B are similar but A has more advanced disease, can we learn about progression for B
- How do we transfer knowledge through the network?
- How do we cheaply initialize new twins? How do we exploit the scale of the network to minimize costs?

Physics-informed models

Building knowledge into data-models

How can we incorporate relatively simple physics into data-models?

$$\frac{\partial u}{\partial t} = \nabla.(p_1 u) + \nabla.(p_2 \nabla u) - p_3 u + g$$

Physics-informed models

Building knowledge into data-models

How can we incorporate relatively simple physics into data-models?

$$\frac{\partial u}{\partial t} = \nabla.(p_1 u) + \nabla.(p_2 \nabla u) - p_3 u + g$$

Suppose we want to infer forcing function g in the linear system

$$\mathcal{L}u = g \text{ given observations } d_i = \langle h_i, u \rangle + e \qquad i = 1, \ldots, n$$

for example by solving constrained optimization problem

$$\min_g (d - Hu)^\top (d - Hu) \text{ subject to } \mathcal{L}u = g$$

or finding the Bayesian posterior

$$\pi(g|d)$$

where $g(x) \sim GP(m(x), k(x, x'))$.

Adjoint aided inference

$\mathcal{L}u = g$. Observations $d_i = \langle h_i, u \rangle + e_i$

Introduce n adjoint systems $\mathcal{L}^* v_i = h_i$

where $\mathcal{L}^*$ is the adjoint operator of $\mathcal{L}$ (automatable).

Adjoint aided inference

$\mathcal{L}u = g$. Observations $d_i = \langle h_i, u \rangle + e_i$

Introduce n adjoint systems $\mathcal{L}^* v_i = h_i$

where $\mathcal{L}^*$ is the adjoint operator of $\mathcal{L}$ (automatable).

Then

$$\langle h_i, u \rangle = \langle \mathcal{L}^* v_i, u \rangle = \langle v_i, \mathcal{L}u \rangle = \langle v_i, g \rangle$$

Adjoint aided inference

$\mathcal{L}u = g$. Observations $d_i = \langle h_i, u \rangle + e_i$

Introduce n adjoint systems $\mathcal{L}^* v_i = h_i$

where $\mathcal{L}^*$ is the adjoint operator of $\mathcal{L}$ (automatable).

Then

$$\langle h_i, u \rangle = \langle \mathcal{L}^* v_i, u \rangle = \langle v_i, \mathcal{L}u \rangle = \langle v_i, g \rangle$$

If $g(x) = \sum z_i \phi_i(x)$ is a linear model, then

$$\langle h_i, u \rangle = \sum_i z_i \langle v_i, \phi_i \rangle$$

$$D = \Phi z + e$$

i.e., an unconstrained linear model in z. Thus exact inference for g possible at zero additional cost.

Adjoint aided inference

$\mathcal{L}u = g$. Observations $d_i = \langle h_i, u \rangle + e_i$

Introduce n adjoint systems $\mathcal{L}^* v_i = h_i$

where $\mathcal{L}^*$ is the adjoint operator of $\mathcal{L}$ (automatable).

Then

$$\langle h_i, u \rangle = \langle \mathcal{L}^* v_i, u \rangle = \langle v_i, \mathcal{L}u \rangle = \langle v_i, g \rangle$$

If $g(x) = \sum z_i \phi_i(x)$ is a linear model, then

$$\langle h_i, u \rangle = \sum_i z_i \langle v_i, \phi_i \rangle$$

$$D = \Phi z + e$$

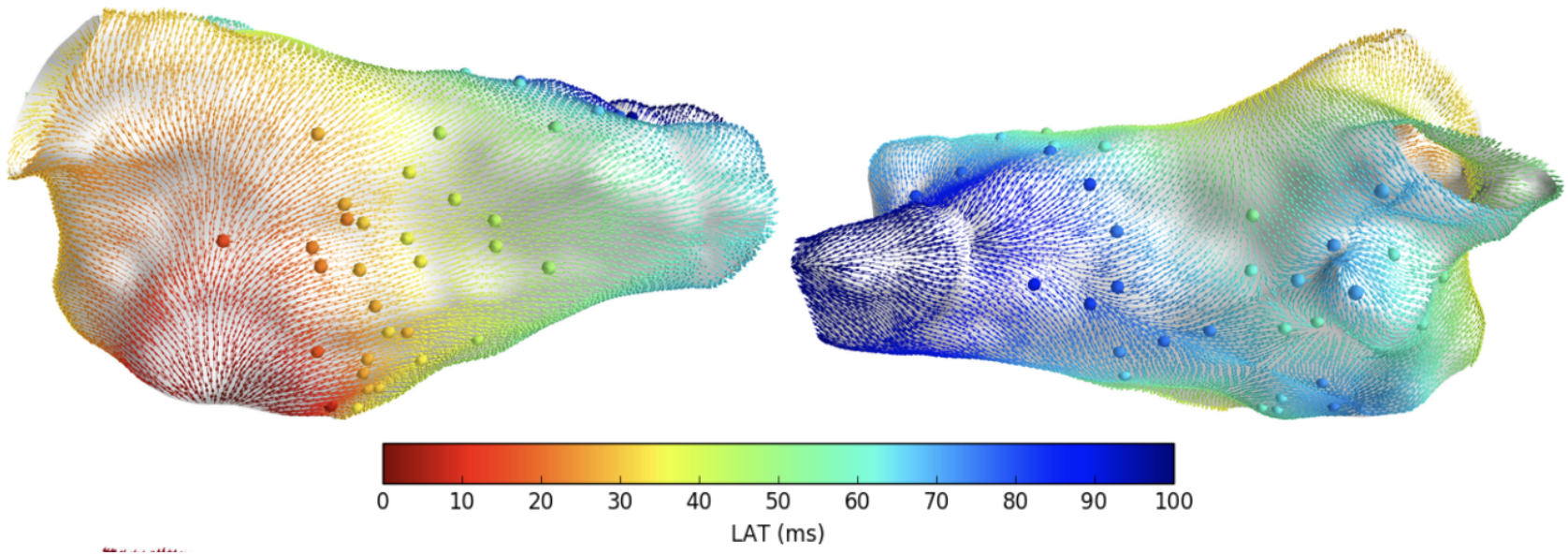
i.e., an unconstrained linear model in z. Thus exact inference for g possible at zero additional cost.

- Many possible basis expansions of GPs, e.g. Mercer, random Fourier features, Laplace etc.
- Computational cost is n (#data points) adjoint solves.
- Method is sequential: each additional data point just requires one additional adjoint sovle.

Manifold valued data

We want to estimate local activation times at all locations on the atria (the *LAT map*)

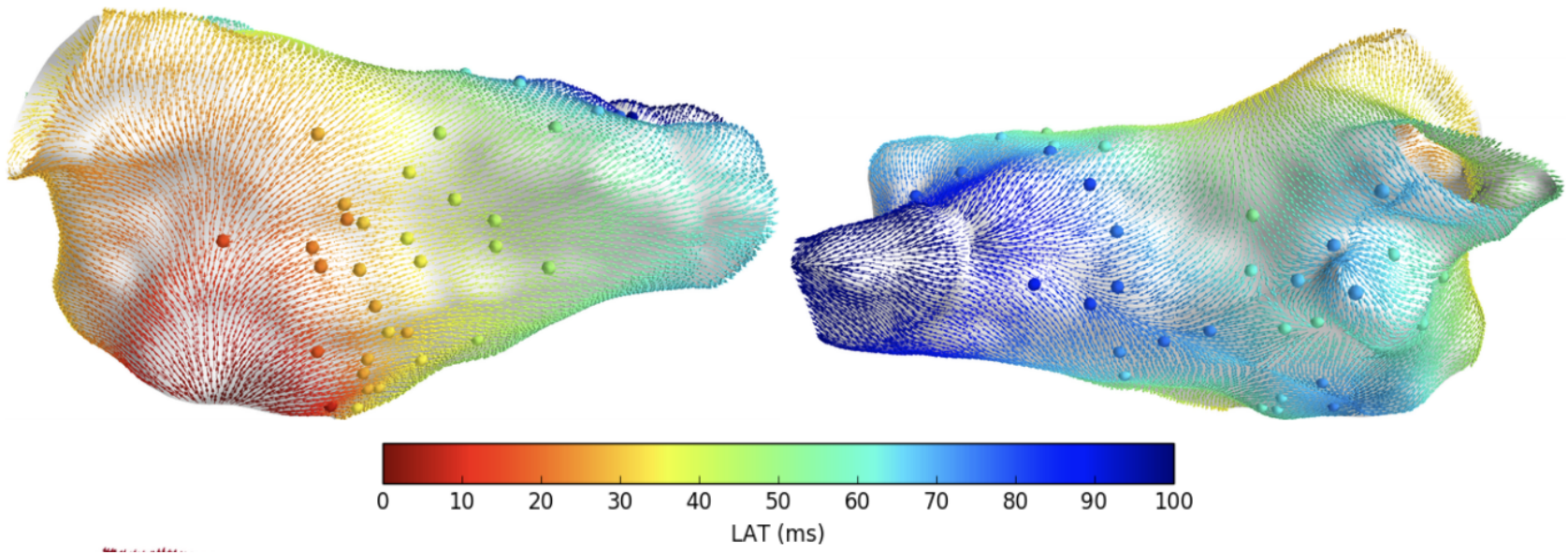
- Typically, only able to measure LAT at a small number (∼10s) of locations on the atrium.



Manifold valued data

We want to estimate local activation times at all locations on the atria (the *LAT map*)

- Typically, only able to measure LAT at a small number ($\sim$10s) of locations on the atrium.



How can we interpolate to other locations $x \in \omega$?

GP interpolation

We want to model

$$LAT(x) \sim GP(m(x), k(x, x'))$$

but standard approaches won't work when the domain is an atrial manifold ω

- Typically covariance is a function of the Euclidean distance between two points i.e. $k(x, x') \equiv k(\|x - x'\|_2)$,

GP interpolation

We want to model

$$LAT(x) \sim GP(m(x), k(x, x'))$$

but standard approaches won't work when the domain is an atrial manifold ω

- Typically covariance is a function of the Euclidean distance between two points i.e. $k(x, x') \equiv k(\|x - x'\|_2)$,

We want the interpolation to take into account distance on the manifold travelled by electrical wave.

- Defining a valid positive definite covariance function on the manifold is hard!

Laplacian basis functions

Coveney *et al.* Phil. Trans. Roy. Soc. 2020

There is a duality between stationary covariance functions, and spectral densities (Wiener-Khinchin):

$$S(\rho) = \int k(r) \mathrm{e}^{-i\rho r} \mathrm{d}r$$

Laplacian basis functions

Coveney *et al.* Phil. Trans. Roy. Soc. 2020

There is a duality between stationary covariance functions, and spectral densities (Wiener-Khinchin):

$$S(\rho) = \int k(r) \mathrm{e}^{-i\rho r} \mathrm{d}r$$

Solin and Sarkka (2019) showed that if we use the Laplacian eigenbasis

$$\begin{aligned} -\nabla^2 \phi_j(x) &= \lambda_j \phi_j(x) \qquad & x \in \omega \\ \phi_j(x) &= 0 & x \in \partial\omega \end{aligned}$$

then

$$f(x) = \sum z_k \phi_k(x) \qquad \text{with } z_k \sim N(0, S(\sqrt{\lambda_j}))$$

is a GP with spectral density S.

This allows us to

- specify a GP in terms of its spectral density, bypassing the need to explicitly define a covariance function
- work directly with processes on the atrial manifold

This allows us to

- specify a GP in terms of its spectral density, bypassing the need to explicitly define a covariance function
- work directly with processes on the atrial manifold

Note that

$$k(x, x') = \sum S(\sqrt{\lambda_j})\phi_i(x)\phi_i(x')$$

and that unlike many other expansions (e.g., Mercer, RFF), the basis functions don't change if the hyper-parameters of the GP change (so we only need compute them once).

This allows us to

- specify a GP in terms of its spectral density, bypassing the need to explicitly define a covariance function
- work directly with processes on the atrial manifold

Note that

$$k(x, x') = \sum S(\sqrt{\lambda_j})\phi_i(x)\phi_i(x')$$

and that unlike many other expansions (e.g., Mercer, RFF), the basis functions don't change if the hyper-parameters of the GP change (so we only need compute them once).
Truncating the sum gives us an approximate low rank GP

$$k(x, x') \approx \sum_{i=1}^{M} S(\sqrt{\lambda_j})\phi_i(x)\phi_i(x'), \quad f(x) \approx \sum_{i=1}^{M} w_k\phi_k(x)$$

for which inference can be done in $O(M^3)$ operations.

Computing conduction velocities

Interest lies in conduction velocities, which are the inverse of the LAT gradient. The Laplacian eigen expansion allows us to compute these

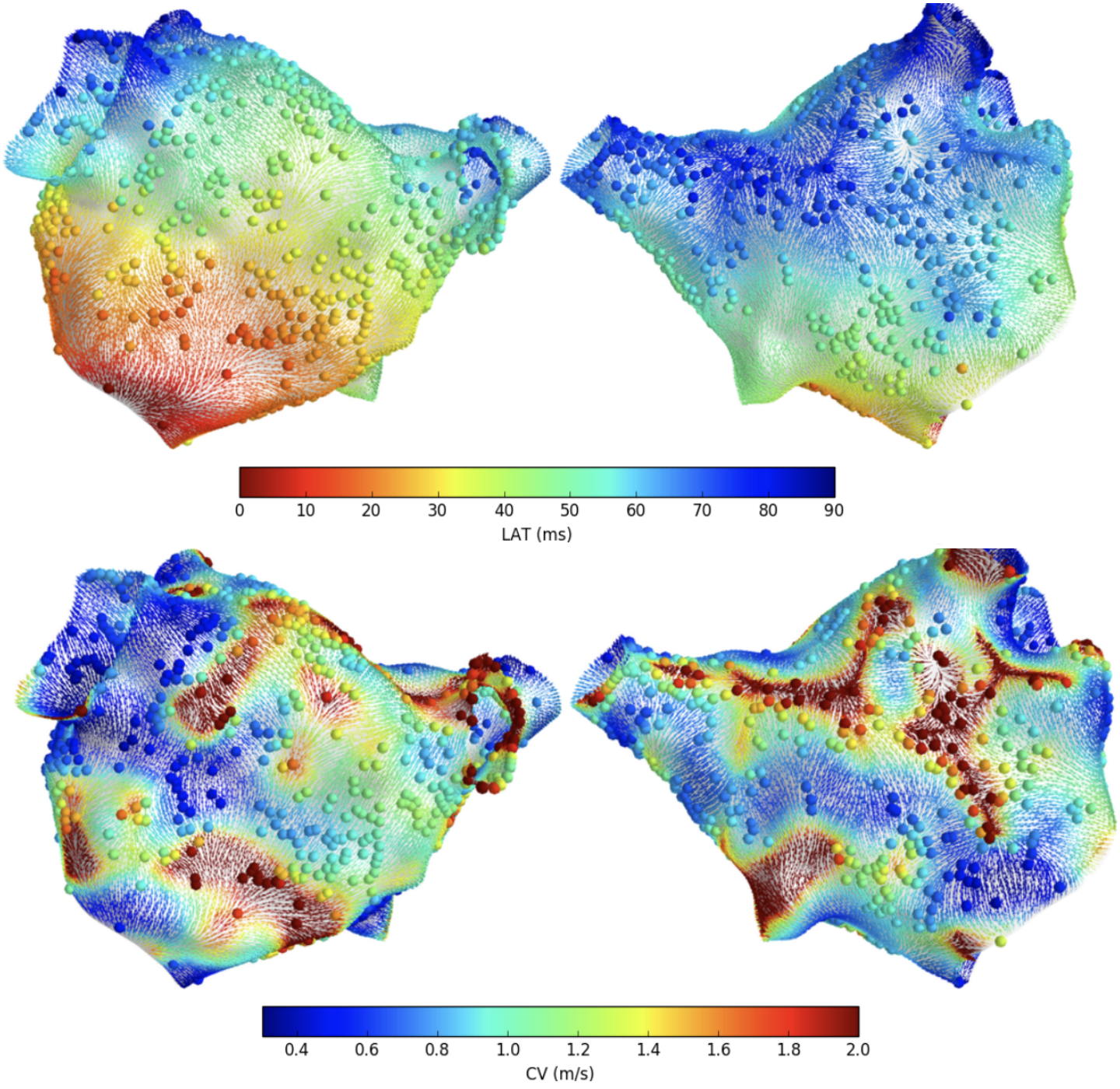
$$\mathbb{E}\left[\frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{x}^*} \mid \mathcal{D}\right] = \frac{\partial \mathbf{k}_*^{\mathbf{T}}}{\partial \mathbf{x}^*}(\mathbf{K}+\boldsymbol{\Sigma})^{-1}\mathbf{y}$$

$$\mathbb{V}\left[\frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{x}^*} \mid \mathcal{D}\right] = \tau^2 \left.\frac{\partial^2 k(\mathbf{x_a}, \mathbf{x_b})}{\partial \mathbf{x_a} \partial \mathbf{x_b}}\right|_{\mathbf{x_a}=\mathbf{x_b}=\mathbf{x}^*} - \frac{\partial \mathbf{k}_*^{\mathbf{T}}}{\partial \mathbf{x}^*}(\mathbf{K}+\boldsymbol{\Sigma})^{-1}\frac{\partial \mathbf{k}_*}{\partial \mathbf{x}^*}$$

where

$$\frac{dk(x,x')}{dx} = \sum_{i=1}^{M} S(\sqrt{\lambda_j})\frac{d\phi_i}{dx}(x)\phi_i(x')$$

allowing us to compute variance estimates of the estimated conduction velocities...


0
10
20
30
40
50
60
70
80
90
LAT (ms)
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
CV (m/s)

Other topics

- Geometric uncertainty
  - Heart is never still, segmentation of MRI/CT image imperfect, images are obtained in unnatural situations.
  - Data are collected from an uncertain geometric location.
  - Need manifold valued models etc.
- Design
  - What data should we collect from the patient?
  - What simulations should we perform with expensive simulators?
- Model discrepancy
  - How can we use the network of DTs to learn the model error?
- Multi-fidelity/multi-level methods
  - If we have models $f_1, f_2, \ldots$, of varying costs and accuracies, how do we make the most accurate predictions we can within some given computational budget?
- Modular models
  - Can we calibrate model components independently before coupling?

Conclusions

Digital twins run into many of the key mathematical/statistical challenges arising in UQ, as well as motivating new challenges

Conclusions

Digital twins run into many of the key mathematical/statistical challenges arising in UQ, as well as motivating new challenges

- At present, DTs aren't used to guide therapy.
  - We can currently build DTs for a single patient, but at great expense
  - Need to scale and speed up this process
- The huge number of uncertain parameters and cost of the simulations will mean we need to compromise:
  - find regularities in the problem to allow us to reduce dimension sufficiently in order to make inference possible
  - learn strong population structured prior distributions
  - develop fast method to approximately infer parameters.

Conclusions

Digital twins run into many of the key mathematical/statistical challenges arising in UQ, as well as motivating new challenges

- At present, DTs aren't used to guide therapy.
  - We can currently build DTs for a single patient, but at great expense
  - Need to scale and speed up this process
- The huge number of uncertain parameters and cost of the simulations will mean we need to compromise:
  - find regularities in the problem to allow us to reduce dimension sufficiently in order to make inference possible
  - learn strong population structured prior distributions
  - develop fast method to approximately infer parameters.
- Newton Institute programme on *Representing, Calibrating and Leveraging Uncertainty* May-August 2025 with 3 workshops.

Workshops

RCLW01
Uncertainty in multivariate, non-Euclidean, and functional spaces: theory and practice
6 May 2025 to 9 May 2025

RCLW02
Calibrating prediction uncertainty : statistics and machine learning perspectives
2 June 2025 to 6 June 2025

RCLW03
Accelerating statistical inference and experimental design with machine learning
23 June 2025 to 27 June 2025

References

- Niederer *et al.* 2021 Nat. Comp. Sci. [Review of DTs]
- Gahungu *et al.* 2022 NeurIPS. [Adjoint aided inference]
- Holden *et al.* 2019 Nat. Clim. Change; Holden *et al.* 2019 Geo. Mod. Dev.; Turner *et al.* 2023 J. Roy. Soc.; Wilkinson 2010 [High dimensional emulation]
- Lok Lei *et al.* 2020 Phil. Trans. A [Discrepancy in electrophysiology models]
- Coveney *et al.* 2020 Phil. Trans. A; Coveney *et al.* 2019 IEEE Trans. Bio. Eng. [GP interpolation on manifolds]
- Corrado *et al.* 2020 Med. Im. Anal.; Coveney *et al.* 2023 [Quantifying geometric uncertainty]
- Coveney *et al.* 2022 Nat. Sci. Reps.; Coveney *et al.* 2021 Front. Elec. Phys. [Calibrating electrophysiology]
- Smith *et al.* 2023 J. Roy. Stat. Soc C, [Sensor calibration]
- Strocchi *et al.* 2023 PLOS Comp. Bio [Sensitivity analysis for whole heart model]
- Breaz *et al.* 2024 ICASSP [Randomized Maximum Likelihood]