



Welcome

Prof. Peter V Coveney
Principal Investigator & Comp-
BioMed Coordinator



It is my pleasure to welcome our readers to the New Year and the final stages of our current Centre of Excellence.

Due to the disruption caused by COVID-19, the Centre will continue its existence for a further six months and will end on 31 March 2024.

Our collective focus in the last quarter of 2022 was on a review with the European Commission and the preparation of numerous reports and deliverables which all feed in to the common purpose. The outcome was very positive and we move ahead from here with renewed strength and vigour.

Amongst other achievements, a subset of CompBioMed partners received a DOE INCITE Award for supercomputing allocation on Frontier, the world's first exascale machine. They are located at ORNL in Tennessee as well as on Polaris, a pre-scale exascale machine at ANL near Chicago. This is to set up and run production studies making use of code and associated software developed within our Centre of Excellence (CoE).

Several of our British partners joined forces to successfully land a UK version of our CoE, funded by UKRI, which started on 1 December 2022 and will run for two years (www.compbiomedx.org). Supported by several

of our European partners acting as external collaborators, this project will allow us to continue to support the aims of the CompBioMed2 CoE until the end of 2025. Its kick-off meeting took place at UCL on 13 December 2022.

CompBioMed held our most recent All-Hands Meeting (AHM) at CINECA in Bologna in June 2022. The 2023 version is taking place in Geneva from 21-23 June 2023. We also completed the follow-up to our highly successful *Virtual Humans* IMAX film (see our 'Computational Biomedicine' YouTube channel). The sequel is entitled *The Next Pandemic* and was premiered at the Amsterdam EyeMuseum in December 2022. It was later shown at a Le Studium Conference that same month.

Many events are planned for 2023, including a major public event at the London Science Museum on 29 March (see p3), which will coincide with an exhibition on the development of vaccines for COVID-19. It will include the first IMAX showing of *The Next Pandemic*, along with the launch of the new book, *Virtual You*, co-written with Roger Highfield (see p2).

We look forward to seeing you in person at our forthcoming meetings.

CompBioMed Conference 2023

The third edition of the CompBioMed Conference series will take place in Science Congress Center Munich (Garching, Germany) from 12-14 September 2023, close to CompBioMed partner Leibniz Supercomputing Centre (LRZ). CompBioMed Conference 2023 (CBMC23) will address all aspects of the rapidly burgeoning domain of computational biomedicine, from genome and molecular medicine via cells to organs, whole human and population levels, embracing data driven, mechanistic modelling and simulation, machine learning and combinations thereof. We welcome contributions from academic, clinical and industrial participants alike.

CBMC23 will present an exciting programme of relevant symposia and world-renowned plenary and invited speakers, proposed and secured by our International Organising Committee. Confirmed plenary speakers include Amanda Randles (Duke University), Richard Law (Exscientia), Gunnar Cedersund (Linköping University), and Ines Thiele (National University of Ireland).

LRZ is a supercomputing centre at the Garching campus near Munich, operated by the Bavarian Academy of Sciences and Humanities. LRZ operates the supercomputers SuperMUC and the next generation SuperMUC-NG. Attendees will be able to enjoy a social event on the evening of 13 September 2023 at the oldest brewery in the world, the Bräustüberl Weihenstephan, including a guided tour of the brewery, beer tasting, and a conference dinner.



SuperMUC-NG is LRZ's next generation supercomputer, capable of 26.9 PetaFlop/s peak performance

Registration and abstract submission opens imminently, you can find more details on the conference website: www.compbiomed-conference.org.

We look forward to seeing you in Garching.

Moving Towards FAIR Data In CompBioMed

CompBioMed and DICE (Data Infrastructure Capacities for EOSC, www.dice-eosc.eu) started their collaboration one and a half years ago. In this time, we have setup a federation between the High-Performance Computing centres at SURF and BSC and with UCL. The aim is to not only help researchers with large data transfers between sites, but also facilitate data replication and long-term preservation and publication of data in an open repository to be findable and accessible by the wider community in longer-term.

In the past year, the B2SAFE service has been deployed

at UCL's Advanced Research Computing Centre (ARC) and federated with the SURF B2SAFE instance.

The SURF B2SAFE instance has also been federated with BSC and LRZ and tested performance wise. This federation of HPC sites enables the transfer of large volumes of data from data processing sites. We are currently looking into different use cases to help promote the use of the platform in data and compute workflows. If you are doing multi-site research and see the value in using this platform, do reach out to us.

Virtual You

In *Virtual You*, which will be the first book on the potential impact of human digital twins aimed at a general reader, Roger Highfield and I describe how this technology will make medicine truly personalised and predictive.

Following an introduction by the Nobel Prize-winning biologist, Venki Ramakrishnan, we provide a panoramic account of the mathematics, engineering and science involved in the creation of digital twins to show how disparate fields can work together to turn medicine into a true science, quoting many distinguished figures along the way, from Sir Paul Nurse and Sir Roger Penrose to Craig Venter and Leroy Hood.

One of the greatest challenges is working with a tsunami of medical data about patients available today. We are seeing the first exascale computers capable of a billion billion floating point operations per second, with Frontier in the US being the first machine officially operating at this scale. Moreover, after half a century of innovation, mathematical models have reached sufficient sophistication to simulate large parts of biology.

Indeed, digital twins are already being used to help develop drugs in clinical trials, not least to reduce vivisection, and, in cardiac medicine, where they are already guiding doctors. Although a full virtual body remains some way off, digital twins of organs, cells and biochemical processes are increasingly contributing to medicine.

The idea to write this book was a direct consequence of a partnership with Roger, who is now at the Science Museum, and research at our Centre of Excellence in Computational Biomedicine.

We realised the subject had broad appeal with the *Virtual Humans* IMAX film we premiered a few years ago, where high fidelity simulations performed, rendered and visualised on supercomputers across Europe were composited on Marenostrum IV at Barcelona Supercomputing Center. The movie brought our work to light in a form that could be instantly grasped: a beating digital heart as big as a bus offers an arresting vision of

where medicine is heading.

Virtual You builds on two previous books *The Arrow of Time* and *Frontiers of Complexity*, both widely cited and translated, which Roger and I wrote, and the eruption of the global pandemic from around March 2020 also helped us lay the foundations for the book by underlining the value of digital twins in a pandemic (see Roger's piece).

In the book we outline five steps which are necessary for the creation of virtual human digital twins, and deals with many of the issues we are addressing within our Centre of Excellence, from genome to whole human and beyond to populations, pandemics and in silico trials.

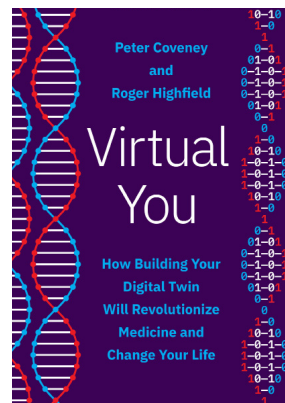
Along the way, we discuss new opportunities raised by quantum computers, which may one day exceed classical machines in some settings, and analogue computing, which is almost entirely forgotten today. The latter is making a comeback because at a stroke it eliminates both the energy problem besetting modern digital computers and overcomes the limitations of using floating point and other discrete number systems, particularly for chaotic dynamical systems.

The importance and limitations of artificial intelligence and machine learning are discussed and their combination with mechanistic models is seen to be the most powerful means of constructing digital twins.

We conclude *Virtual You* with a discussion of the extraordinary new possibilities of ubiquitous digital twins and a debate about the ethical issues, not least how our definition of what it is to be 'healthy' might change.

The book will be published at the end of March 2023, and launched in the IMAX cinema of the London Science Museum on 29 March 2023.

Peter Coveney
University College London and
University of Amsterdam



Event Highlight

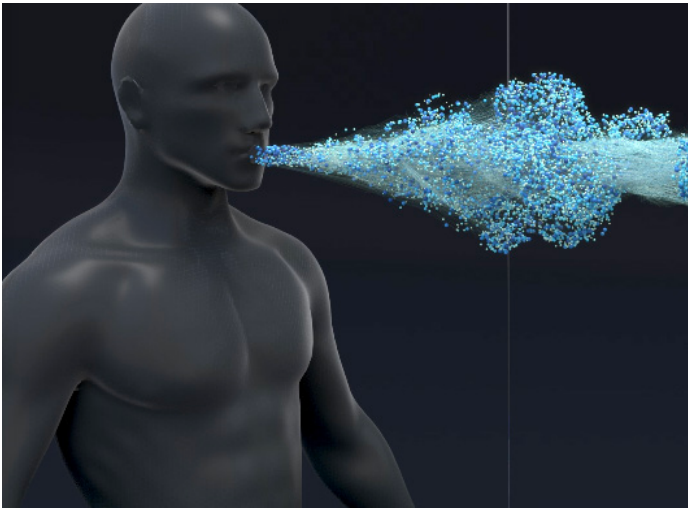
During the VPH Conference 2022 in Porto on September 9th 2022, CompBioMed hosted a special session on "The role of exascale Computing in Computational Biomedicine". The session had four experts from the project consortium talking about examples and case studies benefitting from the power of exascale in their applications.

- Dr Mariano Vazquez from the Barcelona Supercomputing Centre showed how the simulation code Alya can be used on supercomputers to simulate on virtual populations the effect of cardiovascular therapies.
- Prof. Marco Viceconti from the University of Bologna introduced the BoneStrength in silico trials solution to predict the efficacy of a new osteoporosis drug as a guiding example to demonstrate how the simulation of phase III clinical trials does require exascale capabilities.

- A case study shown by Dr Gábor Závodszy highlighted how exascale power enables simulations of red cells interactions in the blood flow and can be used both to solve clinically relevant problems and to falsify causal theories for specific diseases.
- Lastly, the coordinator Peter Coveney showed how exascale supercomputers enable a new generation of high-throughput binding affinity simulations that drastically simplify drug repurposing, e.g., the use of a grid already on the market to treat a different disease.

In the last minutes of the session, a discussion with the audience focused on the opportunities and challenges of using exascale computing in Computational Biomedicine, and confirmed the importance of a European Centre of Excellence on Computational Biomedicine that supports the transition of in silico medicine to become a supercomputing science.

Virtual Pandemics



If you want to see how digital twins might help deal with the next pandemic, come along to see the UK premiere of a new film at the Science Museum in London on 29 March 2023.

Five years after a film about the potential of the virtual human project was unveiled in the museum by the Barcelona Supercomputer Centre, I have worked with the same team - my co-author Peter Coveney, CompBioMed2 and contributors around the world - to showcase how the same approach can be used to curb pandemics. The aim is to celebrate my new book, *Virtual You*, written with Peter Coveney, along with the museum's new COVID-19 vaccine exhibition, *Injecting Hope*. As the new film points out, viruses that circulate in wild animals have caused many serious diseases in people - HIV, of SARS, of MERS, of Ebola and now of COVID-19. Given the global population will surge to around 10 billion by 2050, the resulting changes in land use will prime the planet for more spill-overs of diseases and pandemics.

The film, created by the Barcelona team using the MareNostrum 4 supercomputer, shows how during the global COVID-19 pandemic, supercomputers were used to simulate many features of the outbreak, from the microscopic world of the virus to the spread of infection across cities to how the virus reproduces

to modelling the many possible ways to treat, hinder and halt infection.

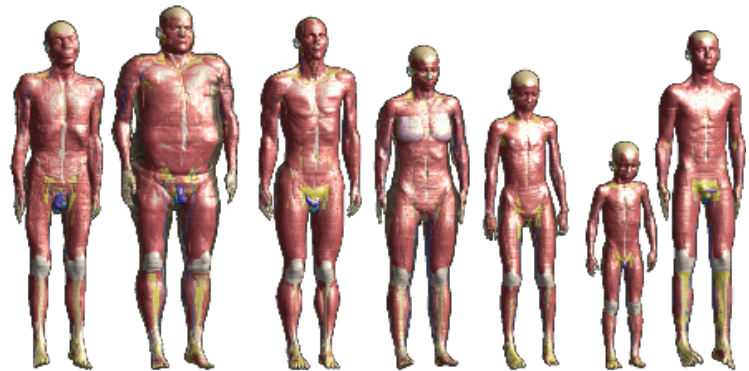
For example, as my co-author has shown, one way to stop infection is to screen drug candidates in supercomputers to find those that can bind to the spike protein of the virus to prevent it from entering human cells.

However, because they borrow our biochemistry to multiply, they are hard to combat without also damaging the cells they have infected, causing side effects. Here, again, computer simulations can provide advance warning of those effects that are potentially dangerous.

Beyond the virus itself, many aspects of a pandemic can be studied in supercomputers: they can make predictions of hospital bed capacity, assess public health measures, such as ventilation, reveal the spread of aerosols, and predict the future course of the pandemic. And supercomputers can, in turn, be used to check the reliability of these pandemic simulation models.

The screening of our new film, which follows a successful viewing at the iconic EyeMuseum in Amsterdam last December, also marks the publication of *Virtual You*. In the words of the book cover, we will tell a general audience "How Building Your Digital Twin Will Revolutionize Medicine and Change Your Life."

Roger Highfield, Science Director, Science Museum Group, Member Medical Research Council



Quicker Results With Higher Fidelity: HPC For Biomedical Apps

Are you struggling because your code cannot run in an acceptable time? Perhaps you are trying to simulate over 1000 virtual patients? Or do you simply need to do some large-scale sensitivity analysis to get your solution certified by the regulatory authority?

Whether you work in academia, in a research hospital, or in a company, there is a good chance that you are starting to make intensive use of the computational biomedicine applications you have developed. As these evolve, you may be discovering performance issues, and we can help with a free service that's open to all.

CompBioMed now offers free support to organisations in their initial steps towards either parallelising their existing computational biomedical applications, or improving the scalability of those applications already parallelised; and thereafter deploy them on HPC platforms.

Our Scalability Service boosts the performance of biomedicine applications via a range of support routes, such as informal discussions about efficient use of parallel platforms in general and code reviews, to porting and profiling applications and suggesting improvements, or working closely with the client and adapting the source code on their behalf.

Please visit the website [1] for more information, where you'll find links to a group email of experts in both HPC and biomedical applications, and access to the public Slack channel #scalability, hosted by "In Silico World" Community of Practice [2], which provides a safe space to share scaling questions. Application routes are also available on the website

The Scalability Service includes a live, useful overview for programmers with ideas on improving the performance of parallel applications for high-end supercomputers [3].

In the field of biomedicine, many applications deal with sensitive data, and clients can be assured that great care is taken when adapting their applications and managing the associated data. The data policies cover data privacy, data security, and research data management. See link [4].

[1] compbiomed.eu/compbiomed-scalability-service

[2] insilico.world/community

[3] compbiomed.eu/rough-guide-to-preparing-software-for-exascale

[4] compbiomed.eu/compbiomed-data-policies

Biomedical Urgent Computing In The Exascale Era



Image shows simulation of the effect on red blood cells of a flow diverter implanted in a brain artery to reduce local pressure.

In the field of personalised medicine, one example of urgent computing is the placement of a stent (or flow diverter) into a vein in the brain: once the stent is inserted it cannot be moved or replaced. Surgeons of the future will use large-scale simulations of stent placement, configured for the individual being treated. These simulations will use live scans to help identify the best stent, along with its location and attitude.

CompBioMed prepares biomedical applications for future exascale machines, where these machines will have a very high node count. Individual nodes have a reasonable mean time to failure; however, when you collect hundreds of thousands of nodes together, the overall mean time to failure is much higher. Moreover, given exascale applications will employ MPI, and a typical MPI simulation will abort if a single MPI task fails, then a single node failure will cause an entire simulation to fail.

Computational biomedical simulations may well employ time- and safety-critical simulations, where results are required at the operating table in faster than real-time. Given we will run these urgent computations on machines with an increased probability of node failure, one mitigation is to employ Resilient HPC Workflows.

Two classes of such workflows employ replication. The first resilient workflow replicates computation, where the same simulation is launched concurrently on multiple HPC platforms. Here the chances of all the platforms failing is far less than any individual; however, such replication can prove expensive, especially when employing millions of cores. The second resilient workflow replicates data, where restart files are shared across a distributed network of data and/or HPC platforms. Then, a simulation that finds its host HPC fail will simply continue from where it left off on another platform.

In both classes, the simulation's results will be available as if no failure had occurred; however, replicated computation will produce results quicker, as the simulation will not have to wait in a second batch system. But by my from personal experience, the time to coordinate multiple HPC platforms to simulate concurrently increases as the cube of the number of platforms. As such, the total turn-around time of the second class (repli-

cated data) is more likely to be faster. Turn-around might be reduced further via the use of batch reservations.

The LEXIS project has built an advanced engineering platform which leverages large-scale geographically-distributed resources from the existing HPC infrastructure, employs Big Data analytics solutions and augments them with Cloud services.

LEXIS has the first class of reliant workflow in its arsenal. EPCC is working with LEXIS using the University of Amsterdam's HemoFlow application, to create an exemplar of the second class: resilience via data replication.

For our workflow, we have created a data network and an HPC network. The data network based on nodes of CompBioMed and of the LEXIS "Distributed Data Infrastructure", relying on the iRODS/EUDAT-B2SAFE system. The HPC network includes five HPC systems at EPCC, LRZ, and IT4I. Both these networks are distributed across countries to mitigate against a centre-wide failure, eg, power-outage. The application is ported to all the HPC systems in advance. The input data resides both on the LEXIS Platform and replicated across the data nodes. The essential LEXIS Platform components are set up across their core data centres (IT4I, LRZ, ICHEC, and ECM-WF), with redundancy built in for failover and/or load balance. These components include an advanced workflow manager, namely the LEXIS Orchestrator System.

The test workflow will progress as follows. The workflow manager submits the simulation at the initial HPC centre. Restart files are created at regular intervals and the workflow manager replicates these across all the data nodes. We will then emulate a node failure: a single MPI task will abort and, as such, will cause the entire simulation to fail. This failure will trigger the workflow manager to restart the simulation on one of the remote HPC platforms, using the latest pre-staged restart files. An automated choice of platform ensures the fastest turn-around as performed by the LEXIS Platform's broker, namely the Dynamic Allocation Module. Once the target HPC platform is known, the latest restart file is staged there from the closest data node, and the simulation will continue.

The staging of data for exascale simulations must naturally consider both the amount of data to be moved and the bandwidth required; however, given biomedical simulations can contain patient-sensitive data, data staging must also ensure the staging follows all relevant legal requirements, including the common FAIR principles of research data management.

Exascale supercomputers bring new challenges and our Resilient HPC Workflow mitigates the low probability but high-impact risk of node failure for urgent computing. These supercomputers are emerging and present an exciting opportunity to realise personalised medicine via ab initio computational biomedical simulations which, in this case, will provide live, targeted guidance to surgeons during life-saving operations.

Gavin J. Pringle
University of Edinburgh

Find CompBioMed online

Our website (www.compbio.med.ucl.ac.uk) is full of all the latest news and information about CompBioMed, including further information on our Partners and Associate Partners, past and future events. We have an active and growing following on Twitter (@bio_comp), a user-forum on LinkedIn (CompBioMed) and

we have our own YouTube channel (YOUTUBE Computational Biomedicine), where you can watch live streaming of events and presentations at previous events and webinars, as well as our *Virtual Humans* film (YOUTUBE <https://youtu.be/1FvRSJ9W734>).

CompBioMed Coordinator:
Peter V. Coveney (p.v.coveney@ucl.ac.uk)
Project Manager:
Hugh Martin (h.s.martin@ucl.ac.uk)

This Centre of Excellence has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823712.

