Report on solutions and end-user engagement

Rationale

As the CompBioMed CoE develops, its long-term success will be increasingly measured by how well it engages and supports prospective users. Here we offer a first review, where we propose a possible taxonomy for prospective users and prospective use cases, which could guide our offerings. We also review the early solutions made available, and how their developers see them positioned in this taxonomy, in particular in term of provision and support models. Finally, we propose a framework for monitoring our engagement with users. Once this document is discussed, revised, and approved, all relevant actions (for example, concerning user portal(s), software repository, etc.) should be structured around it.

Who are the users?

A first top level of our tentative taxonomy is quite simple: the VPH Institute has proposed all Computational Medicine applications can be divided by end user:

- Digital Patient: these are all predictive medicine solutions that are used by **medical professionals** to do clinical research or to support decisions towards an individual patient; these include diagnostic technologies, prognostic / risk assessment technologies, treatment planning technologies, etc.
- In Silico Clinical Trials: these are all predictive solutions using individualised computer simulation used by **biomedical companies, clinical research organisations, and regulators** in the discovery, development, or regulatory evaluation of a medicinal product, medical device, or medical intervention, both pre-clinical and clinical.
- Personal Health Forecasting: these are all predictive medicine solutions that are used by the **individual citizens/patients** in their daily living to monitor their health status and self-manage their chronic condition, alone or in collaboration with remote medical professionals.

From these three macro-applications, we can derive a number of User Scenarios, that have emerged so far:

- Non-clinical research: all research on the development of computational medicine models which advances the basic science of medicine, as well as computational medicine in support of fundamental research in biology, physiology, biochemistry, and biophysics.
- Clinical research: computational medicine technologies used to refine and empower clinical research, disease mechanisms, observational studies, intervention efficacy assessment, etc.
- Clinical decision support: use of patient-specific models to support the clinical decision on diagnosis, prognosis, or treatment.
- Drug discovery: model mechanisms of disease, support for the discovery of potential druggable targets, support for the elucidation of the mechanism of action.
- Design & optimisation: support for all biomedical products (drugs, devices, combinatory, biological), the design and the optimisation in term of safety, efficacy, synthesis/manufacturability, etc.

- *In silico* preclinical trials: use of individualised computer simulation in the regulatory pre-clinical evaluation of any biomedical product; models can reduce or replace the use of bench, in vitro, or ex vivo tests; they can also be used to reduce, refine, and partially replace animal experimentation (non-animal technologies).
- *In silico* clinical trials: use of individualised computer simulation in the regulatory clinical evaluation of any biomedical product. Models can reduce the number of patients involved, or the duration of a clinical trial, or mitigate the risks involved. Models can also be used to augment clinical trials where conventional approaches are impossible, i.e. in rare diseases.
- Personal Health Forecasting: predictive medicine solutions that are used to advise or monitor individual citizens/patients.

Why use HPC?

Hereinafter we will use the term High-Performance Computing (HPC) in a colloquial sense, indicating anything that accelerates the solution of predictive models by use of computational resources that go beyond the desktop. This may involve classic HPC, Cloud computing, High-throughput Computing, etc.

In our exploration, we have identified a number of reasons why Computational Medicine users need to scale-up:

- Run full order model to:
    - Solve unreducible model
    - Validate reduced-order model
- Scale full order model to:
    - Test convergence
    - Model larger space-time regions
- Run model repeatedly to:
    - Do uncertainty quantification
    - Inform surrogate model
- Combinatory explosion in:
    - Multiscale models
    - Strongly coupled models
- Run model urgently to:
    - Provide time-constrained clinical decision support

Who does what?

If we combine these user motivations with the use case scenarios above, not all possible combinations are relevant. The result is represented in the table below. It should be recognised that this is largely

speculative, and based on the collective experience of what it is a very young sector; we cannot exclude that particular additional use motivations, use scenarios, or combinations of the two that we have now excluded might turn out to be relevant in the future. Till they provide a useful starting point to orient prospective users on which problems we can help them with.

	Solve unreducible model	Validate reduced-order model	Test model convergence	Larger space-time regions	Do uncertainty quantification	Inform surrogate model	Multiscale models	Strongly coupled models
Non-clinical research	X	X	X	X	X	--	X	X
Clinical research	X	--	--	X	X	--	X	X
Clinical decision support	--	--	--	--	X	X	X	X
Drug discovery	--	--	--	X	X	--	X	X
Design & optimisation	--	--	--	X	X	X	X	X
In silico preclinical trial	--	--	--	X	X	X	X	X
In silico clinical trial	--	--	--	X	X	--	X	X
Personal Health Forecasting	--	--	--	--	X	X	--	--

Pilot solutions from core partners

In order to initiate the exploration of what CompBioMed can offer to these prospective users, we queried the core partners involved with WP6: USFD, BSC, UVA, UGE, UPF, and UCL. We asked them to describe their solutions by populating a table with the following information:

Field	**Tab**	**Explanation**	**Values**
Name	Solutions	Short acronym used as unique identifier in all communications to refer to this end-user solution	
Layperson Description	Solutions	Full name and layperson description of the solution from an end-user point of view	
Technical description	Solutions	Technical description of the solution, including key information on solvers, etc.	
Access mode	Solutions	Type of access to the solution offered to end-users: Source indicates Open or licensed access to the source code, via GitHub or similar; Direct indicates direct access to all end-users with an account on the target HPC system; Service indicate and end-user interface that has an HPC backend; Indirect indicates access via consulting service, typically, the user sends the data and gets back the solution	Source, Direct, Service, Indirect
HPC system	Solutions	List of all HPC systems available in CompBioMed that can be used to access the solution	

Version	Solutions	Version of the solution, and/or of the underlying solvers; multiple versions can be listed if their features of HPC systems availability differs	
End-users	Solutions	List end-users that are already using the solution; for each user listed here there should be a match in the Needs tab	
Web site	Solutions	URL to access the solution, or a description of the solution	
Provider	Solutions	Acronym of the partner organisation that provides this solution	
Contact email	Solutions	e-mail of the person(s) to be contacted for more information or access to the solution	
End-user	Needs	Acronym of the End-user institution	
Type	Needs	Type of end-user (Research, Clinical, Industrial)	
Need	Needs	Brief layperson description of the end-user need	
Volume	Needs	Average number of simulations per year once the solution is fully adopted	
Contact	Needs	Name of the end-user presentative that can be contacted for more information of this need, and to offer possible solutions	
email	Needs	e-mail of the end user representative or this need	
Solution	Needs	If the need is already met by one of the solutions provided by CompBioMed, write here the acronym, as indicated in the Solutions tab	

After a first pass, we sent to each group additional requests for clarification. With the tabulated data and the additional replies, we compiled the descriptions reported below. Once this approach is approved and deployed, the list will need to be expanded to include also all the offerings of our associate partners.

CT2S [https://ct2s.insigneo.org]

Computer Tomography to Strength (CT2S) is an on-line service developed by the team of Prof Marco Viceconti at the Insigneo Institute at the University of Sheffield (UK), which allows the prediction of the biomechanical strength of a patient's bone from a clinical CT scan of that bone. The simulation involves stochastic finite element analysis of subject-specific Body-Organ multiscale models generated from CT data. The solution is currently exposed as a service, accessible through a web interface; the backend HPC system currently in use is USFD's own ShARC. The service is currently provided at cost, with a significant discount for non-sponsored clinical studies. USFD is currently exploring the best marketing strategy. The contact person is Shannon Li ct2s-support@insigneo.org.

Current users: MultiSim project (Prof Eugene McCloskey, STHFT); Prof Richard Eastell, STHFT; Dr Jennifer Walsh, STHFT.

Access mode: Service (on-line)

Use scenario: Clinical research, Clinical decision support, In silico clinical trial

HPC Systems: ShARC (Sheffield)

HPC motivation: Do uncertainty quantification; Multiscale model.

OpenBF [https://github.com/INSIGNEO/openBF]

OpenBF is an open-source 1D blood flow solver based on MUSCL finite-volume numerical scheme, written in Julia and released under Apache 2.0 free software license. The software is developed by Alessandro Melis and Alberto Marzo at the Insigneo Institute at the university of Sheffield (UK). The solution is currently exposed as open source software; it is also installed on SURFsara's HPC-Cloud, where it is used for large-scale sensitivity analysis and uncertainty quantification studies. The contact person is Alberto Marzo a.marzo@sheffield.ac.uk.

Current users: Ahmed Mustafa (PhD student at USFD) and Dr Ana Paula Narata (Neuroradiologist at Tours, France) (pre-operative guidance on thrombectomy), Marco Frison (PhD student at USFD).

Access mode: Source; Direct.

Use scenario: Non-clinical research; Clinical research; Clinical decision support; In silico clinical trial.

HPC Systems: ShARC (Sheffield); HPC-Cloud (SURFsara).

HPC motivation: For uncertainty quantification.

Insigneo Bone Tissue Suit [NA]

This is a collection of modelling tools developed by the teams of Prof Marco Viceconti, Dr Shannon Li, and Dr Enrico Dall'Ara, at the Insigneo Institute at the university of Sheffield (UK), with the collaboration of Dr Francesc Levrero Florencio (Oxford), Prof Pankaj Pankaj (Edinburgh) and Prof Lee Margetts (Manchester). Starting from microCT or NanoCT datasets of bone tissue, the suit provides tools for:

- MicroMesh: Automatic generation of Cartesian 8-node hexahedral finite element meshes from microCT data, using both homogenous, or density-based heterogeneous material mapping;
- MicroFE: large-scale micro finite element solver, based on the ParaFEM library for large displacement, large strain, simulations of bone tissue micromechanics;
- BoneDVC: Digital Volume Correlation code that computes the displacement field induced in bone tissue specimens subjected to staged compression.

The Insigneo Bone Tissue Suit will enable a complete modelling and validation cycle on very large-scale datasets generated with Sn-microCT with resolutions of up to 4000^3 voxels. The code is installed, optimised, and is accessible to any user with a valid account and CPU-time on the ShARC and Archer HPC systems. The contact person is Shannon Li xinshan.li@sheffield.ac.uk.

Current users: MultiSim project (Dr Enrico Dall'Ara).

Access mode: Direct.

Use scenario: Non-clinical research; Clinical research; Design & optimisation; In silico preclinical trial; In silico clinical trial.

HPC Systems: ShARC (Sheffield); Archer (EPCC).

HPC motivation: Solve unreducible model; Test model convergence; Larger space-time regions; perform uncertainty quantification.

Alya [https://www.bsc.es/research-and-development/software-and-apps/software-list/alya]

Alya, developed by the team of Mariano Vazquez and Guillaume Houzeaux at the Barcelona Supercomputing Centre, performs cardiac electro-mechanics simulations, from tissue to organ level. The simulation involves the solution of multiscale model using a FEM-based electro-mechanical coupling solver, specifically optimised for the efficient use of supercomputing resources. Alya is available for use to research users on MareNostrum, ARCHER, and Cartesius; for clinical and industrial users, BSC recommends users access it as a service, due to the complexity involved with setting up simulations. To this purpose BSC is setting up a spin-off (ELEM Biotech) that will provide commercial software-as-a-service to biomedical industries based on Alya. The contact person is Mariano Vazquez mariano.vazquez@bsc.es.

Current users: 40 users in BSC, 40 registered users outside BSC

Access mode: Direct; Service.

Use scenario: Non-clinical research; Clinical research; Clinical decision support; Design & optimisation for medical devices; *In silico* clinical trial.

HPC Systems: MN4 and Nord3 (BSC); ARCHER (EPCC); Cartesius (SURFsara).

HPC motivation: Solve unreducible model; Multiscale model; Strongly coupled multiphyisics model.

HTMD [https://www.htmd.org/]

HTMD, developed by the team of Prof Gianni de Fabritiis at the Universitat Pompeu Fabra (ES), is a programmable environment to prepare, execute, visualize and analyse Molecular Dynamic simulations in HPC or HTC systems, including AWS. It is a Python-based programmable environment to perform system preparation and building, execution of simulations with different MD codes using adaptive sampling schemes and generate Markov State models to analyse simulations. The code is now maintained by Acellera; it is distributed commercially, but it remains free for academic users. The contact persons are Adrià Perez and Joao Damas compbiomed@acellera.com.

Current users: 500 registered academic users; 5 commercial users.

Access mode: Source.

Use scenario: Non-clinical research; Drug discovery; Design & optimisation.

HPC Systems: AWS (Amazon).

HPC motivation: Solve unreducible model; Do uncertainty quantification.

Playmolecule [http://playmolecule.org/]

Playmolecule, developed by the team of Prof Gianni de Fabritiis at the Universitat Pompeu Fabra (ES), is an intuitive platform to access a diverse set of web applications for molecular research. It is a repository of free best-in-kind applications with a diverse set of solutions like molecular predictors and modelling tools. Simulations are run on GPUGRID for free or via Amazon AWS; The scalability

is provided by Amazon via acecloud, the cloud interfacing software by Acellera. The contact persons are Adrià Perez and Joao Damas compbiomed@acellera.com.

Current users: Over 6000, with >13% from industry

Access mode: Service.

Use scenario: Non-clinical research; Drug discovery; Design & optimisation.

HPC Systems: GPUDRID, AWS (Amazon).

HPC motivation: Solve unreducible model; uncertainty quantification.

BAC [not open source]

The Binding Affinity Calculator (BAC), developed by the team of Prof Peter Coveney at University College London (UK), is a workflow tool that runs and analyses simulations designed to assess how well drugs bind to their target proteins and the impact of changes to those proteins. It is a collection of scripts which wrap around common molecular dynamics codes to facilitate free energy calculations. Use of ensemble simulations to robust, accurate and precise free energy computations from both alchemical and end-point analysis methodologies. BAC is a fairly complex tool to use, so at the moment the development team at UCL have made it available as part of consulting services or research collaborations. However, EnsembleMD provides user-friendly interfaces to related binding affinity calculation services, which will be made available as an App in the on-line store of associate partner DNAnexus; a beta version is being used by pharma. The contact person is Dave Wright dave.wright@ucl.ac.uk.

Current users: Contracts and projects with pharma companies

Access mode: Indirect.

Use scenario: Non-clinical research; Drug discovery; Design & optimisation.

HPC Systems: MN (BSC); ARCHER (EPCC); Cartesius (SURFsara), SuperMUC (LRZ), Blue Waters (NCSA, UIUC), Titan (ORNL).

HPC motivation: Solve unreducible model; performs uncertainty quantification.

HemeLB [https://github.com/UCL/hemelb]

HemeLB, developed by the team of Prof Peter Coveney at University College London (UK), is a software pipeline that simulates the blood flow through a stent (or other flow diverting device) inserted in a patient's brain. The aim is to discover how different stent designs (surface patterns) affect the stress the blood applies to the blood vessel, in particular in the region of the aneurysm being treated. The pipeline also allows the motion of magnetically steered particles, for example coated with drugs, to be simulated and estimates made as to where they might statistically end up. The HemeLB setup tool voxelises the geometry at the given resolution, and HemeLB (lattice-Boltzmann CFD solver) then simulates the fluid flow within that geometry, using the given velocity-time profiles for each inlet. Once complete, the simulation output is analysed using the hemeXtract utility, which can produce images of cross-sectional flow, or 3D shots of wall shear stress distribution in the geometry using ParaView visualisation software. HemeLB is installed, optimised, and available for use to any user with a valid account and CPU-time on ARCHER, Cartesius, SuperMUC, and

Prometheus. The UCL team also provide consulting to biomedical companies and clinical users. The contact person is Robin Richardson robin.richardson@ucl.ac.uk.

Current users: Open Source software used primarily in academia by approx. 40 users

Access mode: Direct; Indirect.

Use scenario: Clinical research; Clinical decision support; In silico clinical trial.

HPC motivation: Solve unreducible model.

HPC Systems: ARCHER (EPCC); Cartesius (SURFsara); SuperMUC (LRZ); Blue Waters (NCSA, UIUC); Prometheus (PSNC).

HemoCell [http://www.hemocell.eu]

HemoCell, developed by the team of Prof Alfons Hoekstra at the University of Amsterdam (NL), is a high-performance library to simulate the transport properties of dense cellular suspensions, such as blood. It contains validated material model for red blood cells and additional support for further cell types (white blood cells, platelets). The blood plasma is represented as a continuous fluid simulated with an open-source Lattice Boltzmann Method (LBM) solver. The cells are represented as Discrete Element Method (DEM) membranes coupled to the plasma flow through a tested in-house immersed-boundary implementation. HemoCell is computationally capable of handling a large domain size with a high number of cells (> 10^4-10^6 cells). The code is currently installed and optimised for Cartesius, Lisa, and SuperMUC (Leibniz Supercomputing Centre system), and can be used by anyone with a valid account and CPU allocation on any of these systems. The contact person is Gabor Zavodsky g.zavodszky@uva.nl.

Current users: Nanyang Technological University (Singapore); Budapest University of Technology (Hungary).

Access mode: Source; Indirect.

Use scenario: Clinical research, Clinical decision support, In silico clinical trial

HPC Systems: Cartesius (SURFsara); Lisa (SURFsara); SuperMUC (LRZ).

HPC motivation: Solve unreducible model; Multiscale model; Strongly coupled model.

Palabos [http://www.palabos.org]

Palabos is Lattice Boltzmann Method (LBM) solver, available as open source, and massively parallel. The team of Prof Bastien Chopard at University of Geneva (CH) has specialised it to solve a number of relevant biomedical problems, including simulation of blood flow, and bone cement penetration during vertebroplasty. The software has specific features to deal with biomedical problems, reading medical images. The contact person is Jonas Latt, jonas.latt@unige.ch.

Current users: NUMECA for clinical exploitation.

Access mode: Source.

Use scenario: Clinical research, Clinical decision support, In silico clinical trial

HPC Systems: Tested on CADMOS BlueGene/Q (Switzerland), UniGe and Baobab (Switzerland). To be deployed on supported CompBioMed HPC systems.

HPC motivation: Solve unreducible model; Multiscale model; Strongly coupled model.

Palabos - Vertebroplasty Simulator [not yet available for general release]

This solution, currently in its final stage of development, uses Palabos to provide a vertical solution for the pre-operative planning of vertebroplasty. Micro CT images of the damaged vertebral body are converted into an LBM model, which simulates multiple cement injections with different access point and cement volume. The simulation results predict exact filling patterns of the injected cement. Plans of future developments include converting the results into a finite element model, which will predict the increase in biomechanical strength with respect to the untreated vertebra. The contact person is Jonas Latt, jonas.latt@unige.ch.

Current users:	University of Geneva, AO foundation.
Access mode:	Direct.
Use scenario:	Clinical research, Clinical decision support, In silico clinical trial
HPC Systems:	Tested on CADMOS BlueGene/Q (Switzerland), UniGe Baobab (Switzerland). To be deployed on supported CompBioMed HPC systems.
HPC motivation:	Multiscale model; Strongly coupled model; Pore-Scale resolution.

Palabos - Flow Diverter Simulator [not yet available for general release]

This solution, currently in its final stage of development, uses Palabos to provide a vertical solution for the pre-operative planning for the insertion of flow diverters. CT scan images of blood vessels with aneurysms or other anomalies are converted into an LBM model. Different types of flow diverters are numerically inserted to test their impact on the blood flow pattern. Simulation output includes wall shear stress distribtion in the aneurysm to predict the rate of blood clotting. The contact person is Jonas Latt, jonas.latt@unige.ch.

Current users:	University of Geneva.
Access mode:	Direct.
Use scenario:	Clinical research, Clinical decision support, In silico clinical trial
HPC Systems:	Tested on CADMOS BlueGene/Q (Switzerland), UniGe Baobab (Switzerland). To be deployed on supported CompBioMed HPC systems.
HPC motivation:	Multiscale model; Strongly coupled model; Pore-Scale resolution.

Toward a monitoring framework

The proposed taxonomy can let us organise in a coherent way the presentation of the solutions available through CompBioMed to the various user groups. It is responsibility of the individual groups to maintain up to date the information on each service. In addition, it is necessary that each group contributes to the monitoring of the service provision by CompBioMed. To that purpose we propose to maintain two sets of tabulated data that need to be refreshed at least once per year.

For each registered solution, the provider should update periodically a table that lists all academic, clinical or industrial users of their solution; appropriate means are required to capture the statistics about code downloading, to make sure we get a proper list, even only with Anonymous as user name, and IP address as Institution. For clinical and industrial users, we should also log how many patients were analysed with the code, or how much money the company told us they saved by using the solution. For academic users, we can log the number of papers in which the solution is cited.

Solution	CT2S	**Provider**	USFD			
Entry Date	**User name**	**Institution**	**User Type**	**Papers**	**Patients**	**Savings**
31/03/2017	Richard Eastell	STHFT	Clinical	3	100	£-
31/08/2018	Eugene McCloskey	STHFT	Clinical	1	20	£-
31/08/2018	Jennifer Walsh	STHFT	Clinical	0	50	£-

From this detailed table, UCL should periodically update a summery table. This would list each solution, the provider, and the relevant KPIs associated to each.

Solution	**Provider**	**Hospitals**	**Companies**	**Academics**	**Patients**	**Savings**	**Papers**
CT2S	USFD	1	0	3	170	0	3
OpenBF	USFD						
BoneSuit	USFD						
Alya	BSC						
HTMD	Accellera						
PlayMolecule	UPF						
HemeLB	UCL						
BAC	UCL						
HemoCell	UVA						