eIUS: Radiotherapy Experience Report

In the text: Some barriers in '{{…}}' to provide more contextual information.

Interviewee profile (group interview; four interviewees;)

Two of the four interviewees work at a department of medical physics providing radiotherapy treatment (high energy radiation treatment of cancer) at a UK hospital.

The other two interviewees work at a computer science department at a UK university, one is also working at a UK e-Science centre with a specific "background (.) on using clusters and grid for parallel type computation".

All collaborate in a research project which explores the use of the computational power of the Grid in this context using Monte Carlo simulations.

Time spent in research

One interviewee working at the department of medical physics specifies the time in research as 100%, the other with 5% (95% clinical management). One of the interviewees at the computer science department is spending 100% time in research.

Research area/Research question(s)

Radiotherapy is a high energy radiation (using X-rays) treatment of cancer. The verification of the accuracy of this treatment is of particular importance in radiotherapy physics in order to target the tumour in the desired way, not affecting the organs or tissue in (very close) proximity. Intensity modulated radiotherapy (IMRT) is a method to split the beams into many smaller 'beam-lets', making the procedure more precise. Besides this core research the interviewees are also collaborating in a research project which explores the use of the computational power of the Grid in this context using simulations.

"[Research] questions [are] related to computational physics numerical methods applied to radiotherapy, we use software to calculate the dose given to patients; one of the crucial question here is how to calculate (..) [this] better. There has been an intense use of Monte Carlo software in (..) our [institute] and we have been using this software for the calculation of the dose as well as for the verification."

"The first aim [in terms of the model] is to have (.) [the procedure] more accurate."

GRID aspects: "(..) what we are trying to do: those calculation techniques for radiotherapy which are understood to be more accurate but are very computer intensive and we are not too worried about the technologies used to achieve our results but we want to get our results as quickly as possible for good clinical reasons. (..) we are using distributed computing to perform those calculations. The other dimension we wanted to do look at (..) [was developing prototypes for] national facilities, we didn't want to have [only] a local solutions for ourselves but to have something that other users [clinical medical physicists] in other clinical centres could use in the future." **For developing a system facilitating the Grid** this means: "When we were thinking about the target audience we wanted a system which an expert user could go and use (..), quickly cut[ting] through the scripts and all the rest of it, but an inexpert user could use, too." "Basically I created a number of general use profiles [based on portlets] that people can use." "(..) you can create scripts that do things, but the idea [is] if people who have the skills and create scripts (.) [those] can be available to others as a pull down menu thing, and (..) profiles (..)." "And if I share profiles with others they can download that, have a look at it and simplify it as it suits them." It is the plan that "the clinicians will be using something that they are familiar with, ideally we are going to have another pull down menu built into the commercial package, and then we need to prove it to the vendors." "(..) sometimes the system is there e.g., **CONDOR**, that's an existing system, all you have to do is to install it and get the portal to feed the job (.) straight into CONDOR's queue, and then trying to put the job over to the GRID, if you want to use one resource it's fairly straightforward but if you want to use all the different resources in the **NGS** and in the sensible way where you run jobs fast as well, you need a **resource broker** (..)" – the latter not being trivial to use. GLOBUS is not used at the moment (because Windows is used by most scientists, with GLOBUS currently not available), but this is planned.

"**Research and applications** are really the same at this research centre, the question is how it is going to benefit the patient. The fundamental and basic research is done at the university but here there is only applied research. Research development, implementation, there must be at least a fair probability of successful implementation in a reasonable time scale (e.g., some small scope projects are as short as a month, on average research projects are done during a PhD project). The research is tied up to patient needs or clinical needs or their perception of future clinical needs, they have to be careful they aren't driven purely by the new technologies. Introduction of new treatment techniques that are going to benefit patients are influencing the research. E.g. the Science and Technology Facilities Council (STFC) had a study day (.)[at a] Health Care Forum recently, they tried to discuss the technological innovations, and the theme on this was uptake in health care (..)."

Research Lifecycle

Literature Review (including deciding on use of technology) – Start of the research process

After identifying the research question the next step is the "literature review, we look to see what the state of the art technology is, looking at MEDLINE, PUBMED, all the usual suspects, they are just normal sort of publications. We get our data from clinical activities related to patients, other materials, e.g. software we get have [either] been developed elsewhere and before, like (..) [by] the National Research Council of Canada, some US labs, and also we develop our own software, too."

Data collection/statistical simulation process – Development/use of software – goes over into data analysis in assessing the simulation results

"There are two phases then, calculation & verification. Calculation is based on the patient's physiology data (CT dataset, X-ray tomography, a slice by slice scanning technique that allows 3D mapping) that allows you to calculate the density of the tissues through which the radiation is going, you can do this both to visualize things and calculate the density. That will feed into the model. Software will calculate (..)

[what] dose the patient needs (..) [as a] target. It also tells you where to focus in the body of the patient and allows you not to expose the part of body that you do not want to be touched by radiation. (..) We are using [the] full Monte Carlo approach."

"The first aim [in terms of the model] is to have (.) [the procedure] more accurate. Full Monte Carlo is used to make it more accurate, it is a statistical technology to get a real time simulation of the patient's body when you hit it with the field of electrons. With Monte Carlo software the more events you run the more accurate results you get; [each] event is a radiation particle tracking from the source of the radiation through the patient and depositing the energy (..). It is a statistical method, so you have to calculate the errors; you repeat [the] event simulations until you are happy with the error. Each event is governed by the probability and it is independent from another event. It is the number of simulations that matter. There is a software called CONDOR, it is a distributed software that helps to run simulation on PCs, like a CampusGrid type distributed thing. But its level of complexity and subtlety is below what gird technology can offer; [the] grid uses a lot (.) [more] computation."

Another way to verify the radiation procedure: "(.) if you treat a patient with a photon field, you have a dose distribution, you touch a **device called portal imager**, you can use this device to get the dose too, and you can calculate the dose in the imager and measure it, during the radiation, compare what you have got with what you deliver, second stage is to use the information you have to retrace back the dose and have an idea of what the dose is you delivered to the patient."

Software used/developed/languages: "Anything, Fortran, C++, traditional codes are mostly Fortran based presented in the National lab in the US primarily, what the computer scientists call legacy code and in physics they all well-defined for clinical use. {{We have also developed a modern Code called Geant, but in terms of application it had an awful lot of inaccuracy in it and it had limitations in fundamental physics.}} We do not use any commercial codes, because there are so many good free codes out there, there is no point. More free Codes are EGS, NRC, MTMP, PENELOPE. When we use these codes we download them on our system rather than on any other hosted application system, but these softwares have different plus points and different problems using them, e.g. [in the case of the free] MTMP you are not given the source code, (..) [only the] 250 MB executable (..); MTMP and Geant are good for neutron transport and all the other particles outside electron or photons, and one just runs the binary on one's system (..). A lot of people are using MTMP because it's easy to use, it doesn't need significant programming (..), and it's only good on the local file system (..)."

Source data (kind of): For running the Monte Carlo models it is important to "match our calculations with measurements that we are taking under standard condition." To this end they bring over "machine from the 'bunker'; 'bunker' is a terminology for a big room with some machine grounded in it which gives out radiation and patients are treated in those rooms, (..) [which are] shielded, because we do not want to give too much background radiation to everyone else." The calibrations of this machine is then checked and adapted to the simulations. "Sometimes we do simulation to calculate, sometimes we do simulation for verification of the calculation (..)."

Use of visualisation at the data collection stage (as input by the doctors): "We use a commercial vendor; (.) the doctor identifies where the tumour is, (..) [provided as a] CT dataset; it is sorted anatomically, the clinician would be drawing on slice by slice base cases and the volume they want to treat and (.) become rendered in 3D, that's built into the treatment planning system, that has not been evolved over the past decades in terms of graphical techniques. (..) Different types of images give you different type of info to use to calculate the radiation dose." Besides **CT (Computed Tomography)**, **Magnetic Resonance Imaging (MRI)** is used commonly today.

Output after input given by the doctors through CT/MR images: "The output is another layer of the dose distribution of the energy released to the patient, and then we can see if the dose distribution fits with the target or not. Then the doctors need to approve the plan for the patients, so they have to look at the plans and then these are discussed with the planner and the physicists and medical doctors."

Data analysis/simulation/visualisation

For **data analysis and visualisation** "we use **Matlab**, another way to analyse data is using **DICOM** which is physical digital imaging and communication in medicine and (..) [a] standard in medicine for [recording] medical information about patients. [An] American college of a radiologist first promoted it and standardised it (.). There are (.) other (.) similar software that we use as part of our planning treatment system, e.g., NUC, OMP. {{[The] visualization aspect is something we are interested in from the point of view that a number of national clinical trails in radiotherapy (..) [developed] some test datasets (..) at a particular centre; now the current situation is that the data is developed in a local centre and then transferred (.) as CD-ROMs; it would be nice to have more modern infrastructure, that sort of information could be kept in a central repository and visualize remotely and potentially clinicians or research staff could look at how things were, or worked out centrally or interact with the central distributed centre, so participating centre could communicate differently.}} It is what we do with the data that matters, e.g. in the radiotherapy process the commission is about how to define volume that can be treated, it's not completely definitive and there are many debates about these."

On the research side (includes collaboration): "The key is that sometimes the calculations methods are different from the results that the clinical system has put out, so there are methods to compare those calculations, datasets and those distributions. e.g. gamma maps (describes the difference between two maps), histograms. Potentially these are discussed with colleagues, particularly if we find out that there is a big difference between Monte Carlo calculations and calculations from **TPS [=conventional treatment planning system]**, and then we go back to the initials and discuss the case, since all of our calculations are respective calculations it won't benefit that specific patient but it will benefit others."

GRID point of view (enabling Monte Carlo): "In most clinical techniques what you need is a trail, where you do large calculations in parallel and compare and analyse the results, that way you prove one method over the other one. Monte Carlo is adding a marginal value to the methods, up to now it was difficult to prove it because of computationally prohibited, but now Monte Carlo can be tested with the computational power available, and as the technology is maturing."

Collaboration

“Collaboration is internal as well as external within Europe, Switzerland, we are part of European network Monte Carlo planning, we consult our local colleagues first but since the community is a relatively small and we are scattered around the world, for instance north America and Europe, we exchange information through conferences and meetings rather than close contacts like phone, etc, because you need a visual compare. We have used Access Grid, particularly (.) [one interviewee] as part of his PhD with colleagues from northern Spain.”

Dissemination

Important journals: ‘British Journal of Radiology’; ‘Physics’; ‘Medicine and Biology Journal’; ‘Medical Physics’; ‘Royal Journal of Oncology’;

Interdisciplinary publications: “One of the interesting things with our colleagues in the Grid research project, is publishing with computer science people, it’s the nature of the work, it’s not really computer science and it’s not really radiology, it’s a classic problem. On the analytical work (.) [one of the interviewees] published a work on histogram analysis in radiotherapy oncology, he managed to make it relevant to radiology, algorithm per se would not allow us to publish in medical journals.”

Publishing/linking to data: “There have been some discussions about publishing the codes, but how is it possible to get it peer reviewed? Some 2000-10.000 lines of codes, so we make it available on the website rather than get it published in journals. Most of the work published in the end is related to our field, rather than publishing the algorithms unless there is a request for a piece of code. An example would be a model that (.) [one of the interviewees] and one PhD student developed, [i.e.] modules for some these big Monte Carlo codes, like the bean codes, the model there, is to give back to the authors of the code, that will give a better framework to the authors and they can incorporate it in their code; then its distributed to the community that way and they have their own validation methods, and they give it back to the originator of the code. It will be open source. {{(..) but we are not able to replace the code that is out there, it needs to be validated, needs a lot of validation work (..) and it takes a while, a few weeks to validate it.”}}

GRID source code: Has not been published so far as “no one has asked for it!”

Other important elements about/in the research:

{{“One of the problems we face is **security**, the level of security that is imposed on the hospital are quite high. e.g. if (.) [another interviewee] wants to share a screen with his colleagues from Switzerland, he cannot do it because of the firewalls and security that information technology department of the hospital has imposed on them. So we need to rely on computers from the university network and use it as a shared repository, and this is not a problem that goes away easily, if anything security is going to increase over the coming years. Also there is a problem that you cannot have a machine that is on two networks, you can have only one network on your machine, you have to transfer files physically between them. If we go ahead with getting Monte Carlo calculation based using National Grid service, one of the problems that we face is patients’ security and data confidentiality.”}}