Journal of Physics: Conference Series

Using computing models from particle physics to investigate dose-toxicity correlations in cancer radiotherapy

To cite this article: A Drew *et al* 2017 *J. Phys.: Conf. Ser.* **898** 072048

Using computing models from particle physics to investigate dose-toxicity correlations in cancer radiotherapy

A Drew[1], P J Elwood[1], K Harrison[1], M A Parker[1], H L Pullen[1], M Romanchikova[2], E Silvester[1], A D Sultana[1], M P F Sutcliffe[3], S J Thomas[2] and P L Yeap[1]

[1] Cavendish Laboratory, University of Cambridge, UK
[2] Department of Medical Physics and Clinical Engineering, Addenbrooke's Hospital, UK
[3] Department of Engineering, University of Cambridge, UK

E-mail: `harrison@hep.phy.cam.ac.uk`

Abstract. A system has been developed to provide flexible, efficient and robust processing of radiotherapy planning and treatment data collected in the VoxTox project, which investigates differences between planned and delivered dose, and dose-toxicity correlations. This paper outlines the system requirements and implementation, highlighting the use made of software tools and computing models developed for experiments at the Large Hadron Collider. Experience with VoxTox data processing is summarised.

1. Introduction

Curative radiotherapy treatment of cancer is planned with the aim of delivering a lethal dose of ionising radiation to a tumour, while minimising damage to nearby healthy organs. Organ movements and shape changes, over a course of treatment typically lasting four to eight weeks, can result in actual doses being different from planned. The UK-based VoxTox project, started in 2012, aims to compute actual doses, at the level of millimetre-scale volume elements (voxels), and to correlate with short- and long-term side effects (toxicity). This is an interdisciplinary study, bringing together researchers in engineering, mathematics, medical physics, oncology and particle physics. The initial focuses are prostate cancer, and cancers of the head and neck. Results may suggest improved treatment strategies, personalised to individual patients.

This paper outlines the data-processing requirements of the VoxTox project, highlights the use of tools and computing models from particle physics to satisfy these requirements, and describes system performance.

2. Planning and treatment data

The VoxTox study analyses anonymised planning and treatment data from consented patients treated using the two TomoTherapy Hi-Art machines (TomoTherapy, Madison, WI, USA) at Addenbrooke's Hospital, Cambridge, UK.

Before treatment commences, a patient undergoes a computed-tomography (CT) scan, performed using kilovoltage (kV) X-rays. The kV CT scan typically covers a region 30 cm to 42 cm in length, in 3 mm slices with 512×512 pixels of $1.074\,\text{mm} \times 1.074\,\text{mm}$. This scan provides images of reasonably high resolution and contrast of the tumour site and surrounding area. A clinician manually outlines in the scan the tumour and relevant neighbouring structures, and specifies constraints on the dose to each area. The data points that define the outlines are known as structures sets. Healthy organs to which dose constraints are applied, and which are investigated in VoxTox, include the rectum for patients with prostate cancer, and the spinal cord and parotid glands for patients with cancers of the head and neck.

A treatment plan is computed that best matches the dose constraints. The plan and resulting dose field are reviewed, and are recomputed as needed, with constraints modified. The prescribed dose is delivered in fractions, following standard guidelines [1], in treatment sessions held daily, except at weekends. Prostate cancer is treated either in 37 fractions of 2 Gy or in 20 fractions of 3 Gy. For cancers of the head and neck, there is some variation in prescription, but 30 fractions of 2 Gy is common.

To maximise dose at the required depth within a patient, treatment is delivered using beams of megavoltage (MV) X-rays. TomoTherapy machines [2, 3] allow beam shaping via a multi-leaf collimator (intensity-modulated radiotherapy), and allow use of the MV treatment beam to perform a CT scan that guides patient positioning (image-guided radiotherapy). The MV CT scans typically cover a region 6 cm to 24 cm in length, in 6 mm slices containing 512×512 pixels of $0.754\,\mathrm{mm} \times 0.754\,\mathrm{mm}$. The images from these scans are relatively low contrast, and in a clinical setting are currently used only to estimate tumour location at treatment time.

Following the completion of treatment, data for a patient's kV CT planning scan, MV CT guidance scans, treatment plan and calculated dose field are archived on the hospital system, but with the kV CT scan downsampled to 256×256 pixels of $2.148\,\mathrm{mm} \times 2.148\,\mathrm{mm}$. The standard method for accessing the data is via a graphical interface, and the process can be time consuming. Within VoxTox, command-line tools have been developed that allow automation of data extraction from the archive, and export to files in the format of Digital Imaging and Communications in Medicine (DICOM) [4].

3. Data-processing model

3.1. Requirements

The exported DICOM files for patients recruited to VoxTox are anonymised, then transferred for processing to the Linux cluster of the High-Energy Physics (HEP) Group at the Cavendish Laboratory, University of Cambridge. From the project start date to the end of December 2016, data have been transferred for 489 patients treated for prostate cancer, and for 249 patients treated for cancers of the head and neck. There is an average of 470 files and 230 MB of data per prostate patient, and an average of 860 files and 440 MB of data per head-and-neck patient, for whom CT scans tend to be longer. The totals to the end of December 2016 correspond to 738 patients, 446 000 files and 220 GB of data.

The planning and treatment data are supplemented by the results of toxicity questionnaires, completed by patients at various time points. Purpose-written software is able to parse responses, most of which are numerical gradings of severity, and maps to standard toxicity scoring systems, following human-readable rules, provided by oncologists. To obtain good statistics on long-term side effects, the aim is to track as many of the recruited patients as possible for at least two years after completion of treatment.

The VoxTox data-processing scheme (Fig. 1) encompases centralised tasks, to produce derived data and data summaries relevant to all collaboration members, and end-user tasks, where individual researchers carry out specialised studies. Centralised tasks include automated structure outlining, or autocontouring, on the MV CT scans, and the calculation of delivered dose based on the density information contained in these scans. End-user tasks include evaluations of autocontouring uncertainties, investigations of differences between dose fields at planning time and at treatment time, and searches for correlations between dose and toxicity.

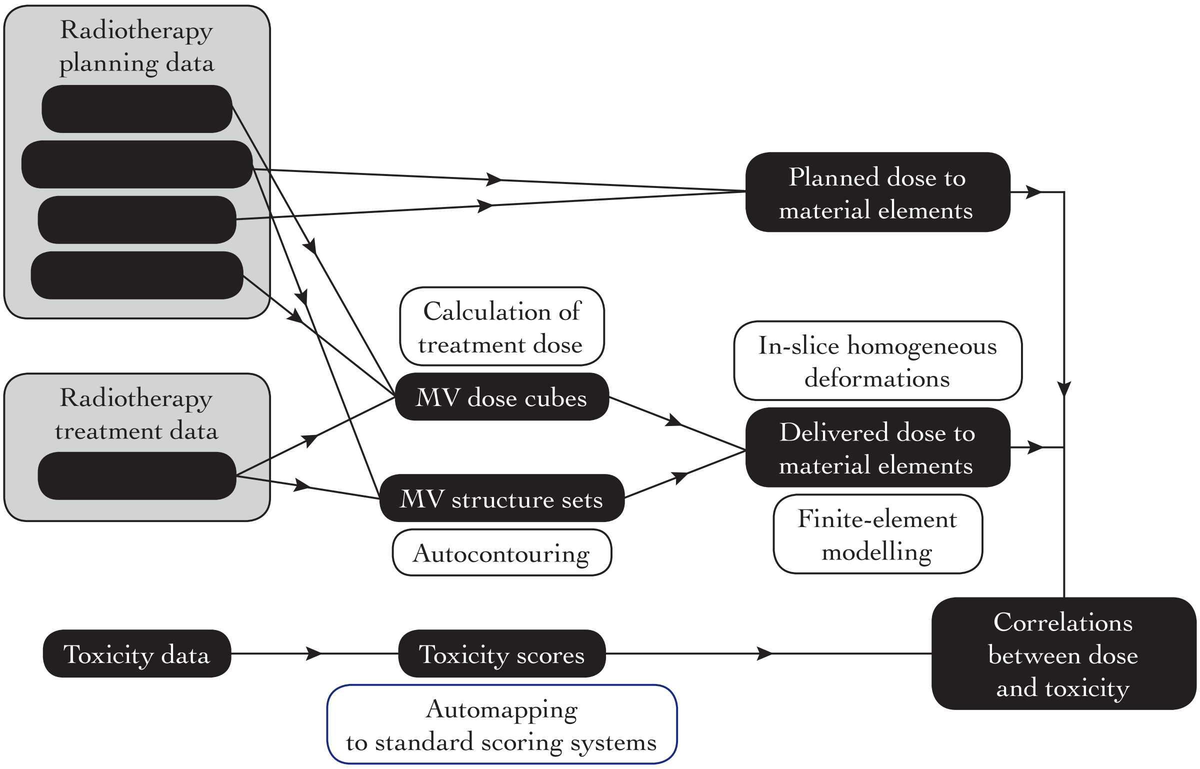

Figure 1. VoxTox data-processing scheme. Radiotherapy treatment and planning data are used to derive information at the level of voxels, or material elements, about planned and delivered doses to anatomical structures of interest. Searches are then carried out for correlations between dose and toxicity.

The data-processing model needs to cope with the different types of task, needs to handle the complexity of the VoxTox data, needs to support users from a variety of disciplines, and needs to allow use of legacy software. Solutions for the processing of petabyte-scale data volumes have been developed to meet the computing challenges [5] of experiments at the Large Hadron Collider (LHC). Some of the resulting software tools and computing models have been adopted for working with the more modest data volumes of VoxTox. Based on the LHC example, these solutions would readily scale for use in multi-centre follow-on studies, where data volumes might be increased by one or two orders of magnitude.

3.2. Software framework

To allow for a wide range of computing tasks to be carried out efficiently, by diverse user groups, two of the LHC experiments, ATLAS and LHCb, have developed a C++ software framework, known as Gaudi [6]. This maintains a distinction between data objects, which are processed, and algorithm objects, which perform processing. The framework also provides services to simplify common operations. Applications are built as ordered sets of algorithm objects, which may be passed configuration parameters at run time. A simplified version of the Gaudi architecture has been implemented, in Python, as the VoxTox software framework.

The primary data object in the VoxTox software framework is the patient object, which links to objects representing associated data subtypes, for example the patient's MV CT guidance scans. On disk, each patient's data are stored in a separate folder, where they're organised according to type and acquisition time. This organisation is reflected in the structure of the

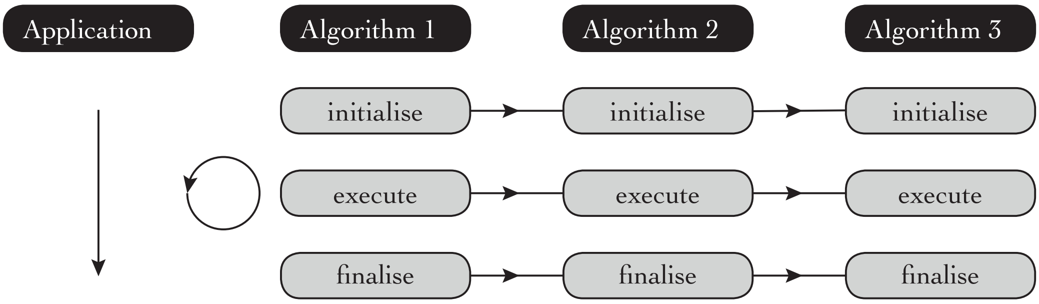

Figure 2. Processing model of Gaudi-inspired VoxTox software framework. An application is built as an ordered set of algorithms. Initialise methods are called before processing starts, execute methods are called once for each patient, and finalise method are called when processing is completed.

patient object. All data objects carry with them the location of the file where the data are stored, and a timestamp. During processing, the file data are loaded into memory only when needed. Locations are currently physical file paths, but it would be straightforward to change to using logical file paths, with lookup via a catalogue.

Mirroring the Gaudi model, each VoxTox algorithm object has an initialise method, called before patient processing commences; an execute method, called once for each patient; a finalise method, called after patient processing has completed (Fig. 2). Services provided by the framework include manipulations of DICOM data, conversions between different coordinate systems, for example from the voxel coordinates of a CT scan to spatial coordinates in the treatment setup, and histogramming, using the Python bindings to ROOT [7].

3.3. Job-management system

In the model of the LHC experiments, a researcher might develop an analysis on a local machine, progresses to testing on a batch system, then carry out a full-scale analysis on a computing grid [5]. To allow transparent use of the different processing systems, to monitor the computing jobs for a full-scale analysis, and to record job details, the ATLAS and LHCb experiments have developed an extensible job-management system, known as Ganga (Gaudi and Grid Alliance) [8]. This job-management system is used directly in VoxTox.

In Ganga (Fig. 3), a computing job is defined, via a Python interface, in terms of the application to be run, the data to be processed, and the system where processing is to be performed. It's optionally possible to specify a rule for splitting the job into sub-jobs that can run in parallel. For VoxTox, Ganga has been extended through the addition of components with built-in knowledge of the software framework and of patient data. Jobs can be split based on the number of patients per sub-job, or based on the number of MV CT scans per sub-job.

4. Experience with VoxTox data processing

4.1. Autocontouring

For a small number of patients, VoxTox oncologists have manually drawn outlines of structures of interest on all image slices of all MV CT scans. Subsets of these scans have been used to train autocontouring algorithms. Other, non-overlapping, subsets, have then been used to evaluate algorithm performance. An algorithm for autocontouring the rectum has been developed [9] that uses two-dimensional Chan-Vese segmentation [10], and is implemented in Matlab. An algorithm for autcontouring the spinal cord has been developed [11] that uses the Elastix toolkit [12] to register a patient's kV CT scan to each MV CT scans, and to apply the registration transform

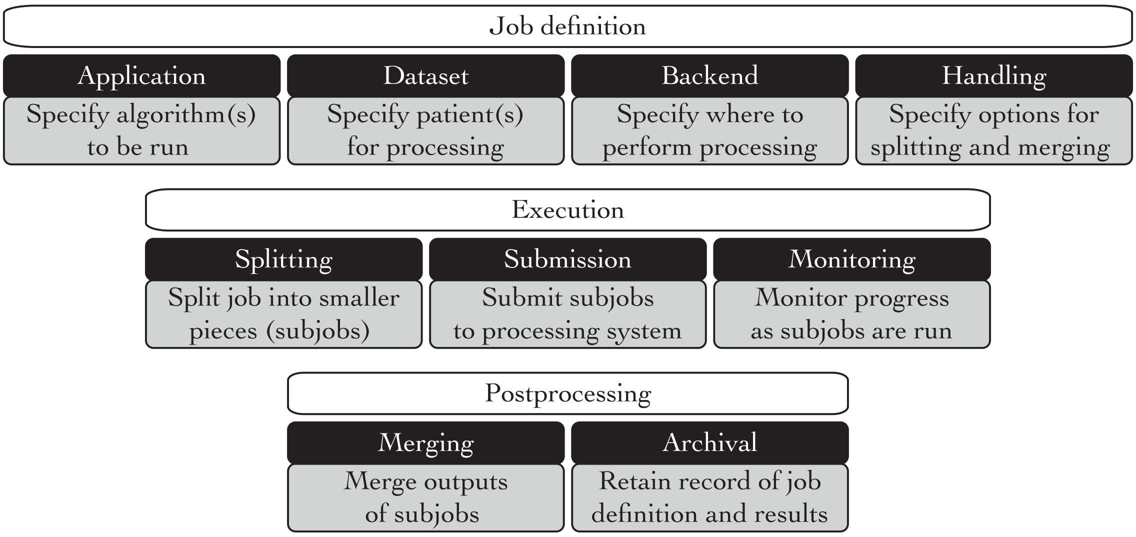

Figure 3. Ganga job-management system, as used in VoxTox. A job is defined in terms of the application to run, the data to process, the processing system, and the rules for job splitting and merging. Job splitting and sub-job submission are performed to start execution, then sub-job progress is monitored. At completion, sub-job outputs are merged. The job definition and results are retained until explicitly deleted.

to the outlines of the spinal cord drawn manually on the kV CT scan. Although both algorithms use third-party software, they have readily been integrated in the VoxTox software framework, and are run via Ganga. The processing time is 1 or 2 minutes per MV CT scan on the Cambridge HEP machines. Possibilities for improving autocontouring accuracy are being investigated, but results in many cases are already good (Fig. 4).

4.2. Calculations of dose fields

A Matlab application that uses a ray-tracing method to calculate dose fields relative to a patient's kV CT scan and treatment plan was developed [13] at Addenbrooke's Hospital, to enable checking of results from the proprietary software of the treatment system. This application has been extended in VoxTox, to allow calculation also of dose fields relative to MV CT scans [14], and the application has been wrapped as a VoxTox algorithm. With standard settings, the processing time per scan is about 6 hours for prostate patients, and is around double this for head-and-neck patients.

4.3. End-user tasks

The VoxTox data-processing system has been used for a variety of end-user tasks. These have included: studies of consistency in structure outlining in individual oncologists (intra-observer variability), and among different oncologists (inter-observer variability); investigations of the delivered dose to the rectal wall, and of the uncertainties in the dose estimate that arise from uncertainties in the rectal outline; exploration of machine-learning techniques to search for correlations between dose and toxicity. The demands of the different tasks have helped develop and test the VoxTox software framework, which has emerged as flexible, robust, and generally straightforward to use.

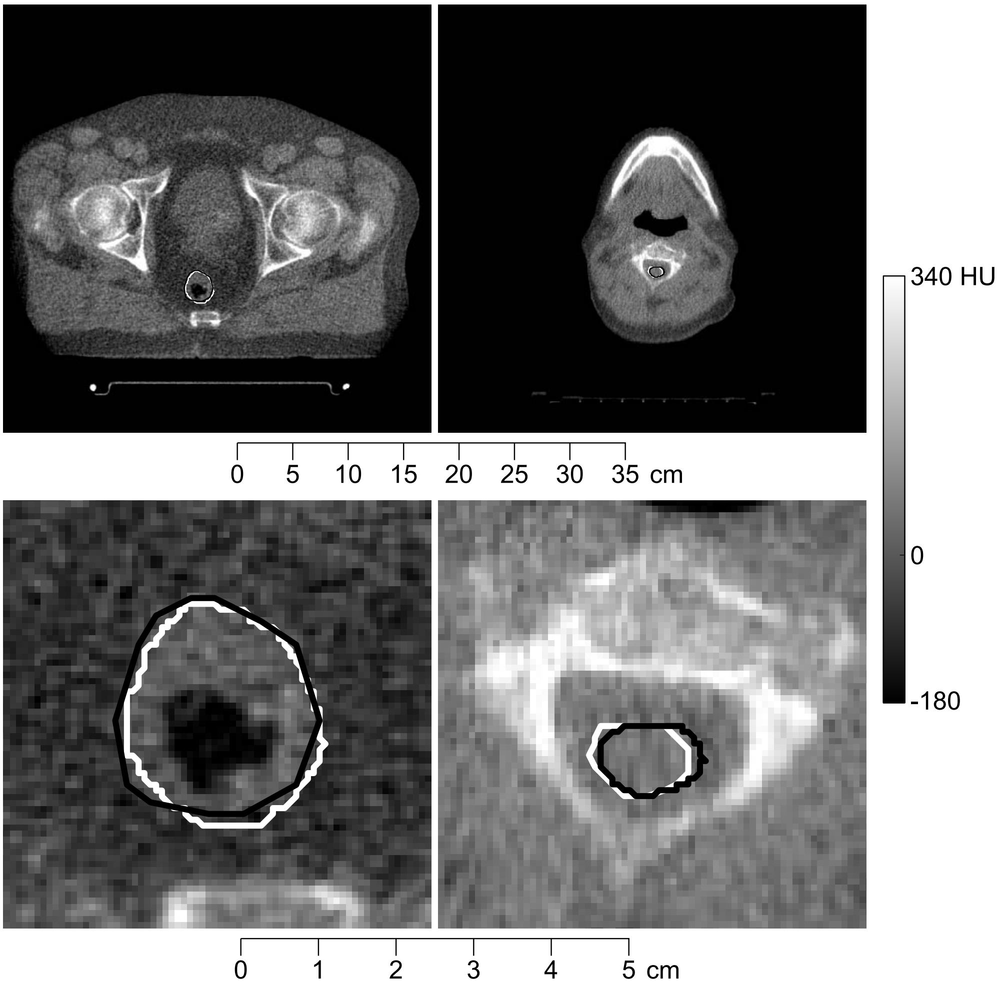

Figure 4. MV CT images, showing axial slices, for two different magnifications: (left) at the level of the femoral heads, for a patient with prostate cancer, with outlines of the rectum superimposed; (right) at the level of the jaw, for a patient with head-and-neck cancer, with outlines of the spinal cord superimposed. Structure outlines are: (black) as drawn manually by oncologists; (white) as obtained by autocontouring algorithms. The images show radiodensities in Hounsfield units.

4.4. Improving voxel tracking

In developing algorithms for autocontouring structures on MV CT scans, problems arise in that the true locations of the structure boundaries aren't known, only the estimates from oncologists, and these estimates tend to be available only for limited numbers of scans. To overcome these problems, possibilities are being investigated for producing synthetic data, where exact structure boundaries are known. One approach considered is to use Geant4 [15], in combination with the Gamos framework [16], to simulate scan recording. A pilot study has been undertaken, where organ volumes were created as tessellated solids, from surface meshes obtained by applying Delaunay triangulation to outlines drawn on a patient's kV CT scan. Appropriate material properties were assigned to the volumes, which were then placed in a simulated TomoTherapy setup. There are challenges with the time needed to simulate adequate numbers of photons, and with modelling parts of the patient for which no organ outlines are available, but initial results (Fig. 5) are promising.

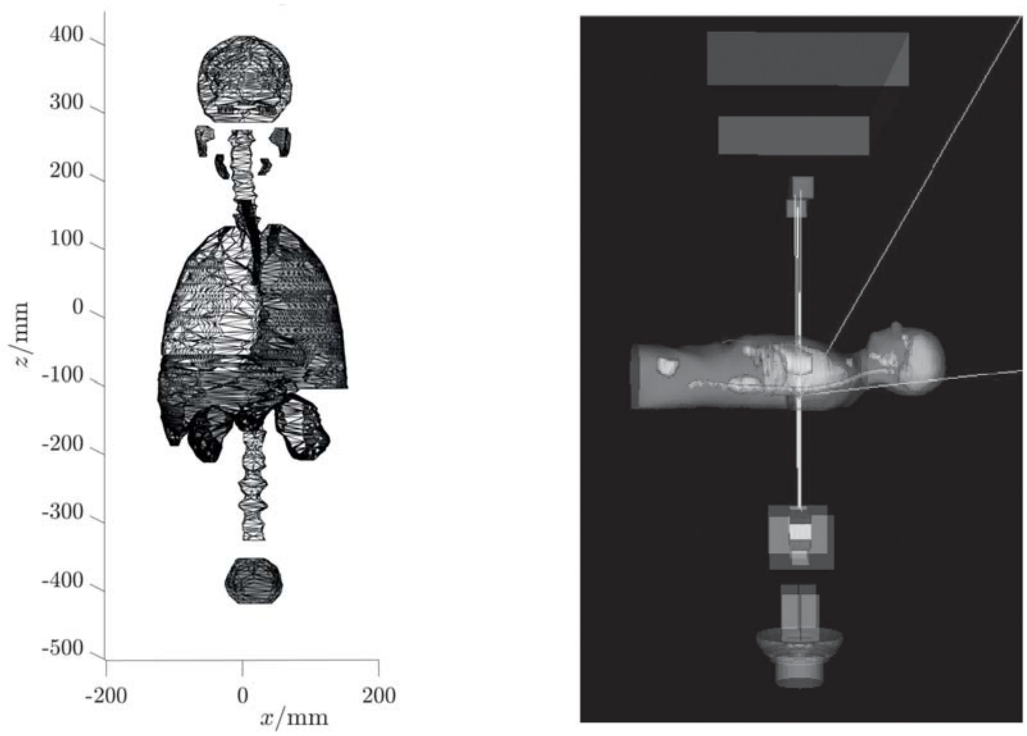

Figure 5. Use of Geant4 to simulate an MV CT scan: (left) representation of internal organs as tesselated solids; (right) virtual patient in the simulated TomoTherapy setup.

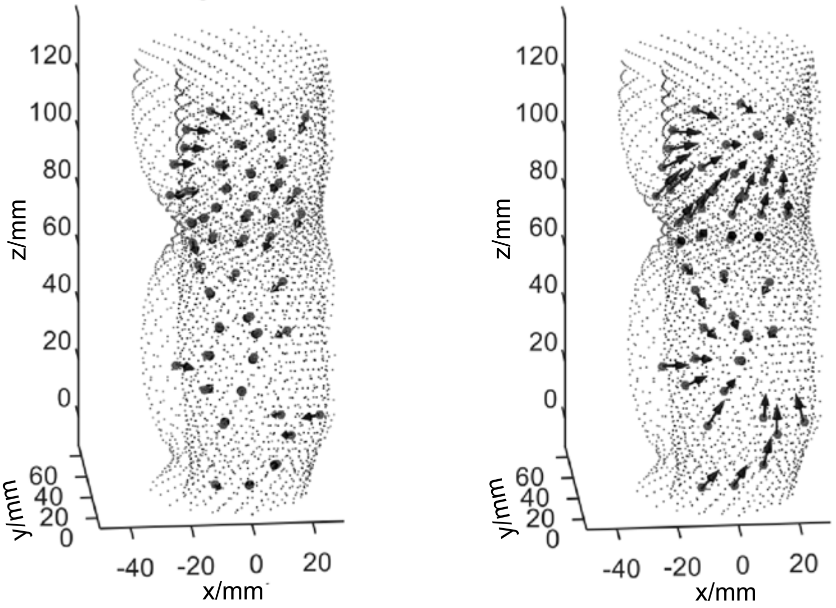

Figure 6. Movements of surface points between an initial inflated state and a final deflated state: (left) as measured experimentally for a physical model of the rectum; (right) as predicted by a finite-element model. The surfaces corresponding to initial and finial states are superimposed, and arrows show the directions of point movements.

Another issue is that organ outlines obtained from examination of MV CT scans provide information on changes in shape over a course of treatment, but provide limited information on the movement of the voxels that make up an organ. Finite-element models may improve on this. To investigate the potential, a simple physical model of the rectum has been created, using a section of a bicycle inner tube, with constraints applied. The movement of surface points when the model undergoes deformation has been compared with the predictions of a finite-element model (Fig. 6). There's imperfect agreement between the finite-element model and the observed behaviour, but the preliminary investigation has been instructive in defining a method for making comparisons.

5. Conclusions

Computing models for petabyte-scale data analysis in experiments at the Large Hadron Collider have been used in developing a data-processing system for the VoxTox study, which investigates correlations between actual dose and toxicity in radiotherapy treatment of cancer. This system provides an efficient solution for the current analysis work, involving 220GB of data, and would scale for use in follow-on studies, with significantly higher data volumes.

References

[1] The Royal College of Radiologists 2016 *Radiotherapy Dose Fractionation, 2nd Edition* (London: The Royal College of Radiologists)

[2] Mackie T R 2006 History of tomotherapy *Phys. Med. Biol.* **51** R427–R453

[3] Burnet N G *et al* 2010 Practical aspects of implementation of helical tomotherapy for intensity-modulated and image-guided radiotherapy *Clin. Oncol. (R. Coll. Radiol.)* **22** 294–312

[4] Mildenberger P, Eichelberg M and Martin E 2002 Introduction to the DICOM standard *Eur. Radiol.* **12** 920–927

[5] Messina P 2004 Challenges of the LHC: the computing challenge *Eur. Phys. J.* C **34** 57–60

[6] Barrand G *et al* 2001 Gaudi — a software architecture and framework for building HEP data processing applications *Comput. Phys. Commun.* **140** 45–55

[7] Antcheva I *et al* 2009 ROOT – a C++ framework for petabyte data storage, statistical analysis and visualization *Comput. Phys. Commun.* **180** 2499–2512

[8] Mościcki J T *et al* 2009 Ganga: a tool for computational-task management and easy acces to grid resources *Comput. Phys. Commun.* **180** 2303–2316

[9] Sutcliffe M P F *et al* 2015 Auto-contouring of the rectum on megavoltage computed tomography images *Cambridge University Engineering Department Technical Report* CUED/C-MICROMECH/TR.100

[10] Chan T F and Vese L A 2001 Active contours without edges *IEEE Trans. Image Process.* **10** 266–277

[11] Yeap P L 2016 Automatic contour propagation using deformable image registration to determine delivered dose to organs-at-risk in head-and-neck cancer radiotherapy *MPhil thesis* (University of Cambridge)

[12] Klein S *et al* Elastix: a toolbox for intensity-based medical image registration *IEEE Trans. Med. Imag.* **29** 196–205

[13] Thomas S J *et al* 2011 Dose calculation software for helical tomotherapy, utilizing patient CT data to calculate an independent three-dimensional dose cube *Med. Phys.* **39** 160–167

[14] Thomas S J *et al* 2016 Recalculation of dose for each fraction of treatment on TomoTherapy *Br. J. Radiol.* **89** 20150770

[15] Agostinelli S *et al* 2003 Geant4 – a simulation toolkit *Nucl. Instrum. Meth.* A **506** 250–303

[16] Arce P *et al* 2014 Gamos: a framework to do Geant4 simulations in different physics fields with an user-friendly interface *Nucl. Intrum. Meth.* A **735** 304–313

[17] Delaunay B 1934 Sur la sphère vide *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na* **6** 793–800