# Validated respiratory drug deposition predictions from 2D and 3D medical images with statistical shape models and convolutional neural networks

# Validated respiratory drug deposition predictions from 2D and 3D medical images with statistical shape models and convolutional neural networks

Josh Williams[1,2*], Haavard Ahlqvist[1], Alexander Cunningham[1], Andrew Kirby[3], Ira Katz[4], John Fleming[5,6], Joy Conway[5,7], Steve Cunningham[8], Ali Ozel[1‡], Uwe Wolfram[1,9‡]

**1** School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK
**2** Hartree Centre, STFC Daresbury Laboratory, Daresbury, UK
**3** Royal Hospital for Children and Young People, NHS Lothian, Edinburgh, UK
**4** Consultant, Meudon, France
**5** National Institute of Health Research Biomedical Research Centre in Respiratory Disease, Southampton, UK
**6** Department of Medical Physics and Bioengineering, University Hospital Southampton NHS Foundation Trust, Southampton, UK
**7** Respiratory Sciences, Centre for Health and Life Sciences, Brunel University, London, UK
**8** Centre for Inflammation Research, University of Edinburgh, Edinburgh, UK
**9** Institute for Material Science and Engineering, TU Clausthal, Clausthal-Zellerfeld, Germany

‡These authors share last authorship
* josh.williams@stfc.ac.uk

## Abstract

For the one billion sufferers of respiratory disease, managing their disease with inhalers crucially influences their quality of life. Generic treatment plans could be improved with the aid of computational models that account for patient-specific features such as breathing pattern, lung pathology and morphology. Therefore, we aim to develop and validate an automated computational framework for patient-specific deposition modelling. To that end, an image processing approach is proposed that could produce 3D patient respiratory geometries from 2D chest X-rays and 3D CT images. We evaluated the airway and lung morphology produced by our image processing framework, and assessed deposition compared to *in vivo* data. The 2D-to-3D image processing reproduces airway diameter to 9% median error compared to ground truth segmentations, but is sensitive to outliers of up to 33% due to lung outline noise. Predicted regional deposition gave 5% median error compared to *in vivo* measurements. The proposed framework is capable of providing patient-specific deposition measurements for varying treatments, to determine which treatment would best satisfy the needs imposed by each patient (such as disease and lung/airway morphology). Integration of patient-specific modelling into clinical practice as an additional decision-making tool could optimise treatment plans and lower the burden of respiratory diseases.

## 1 Introduction

Respiratory diseases such as asthma, cystic fibrosis and chronic obstructive pulmonary disease (COPD) affect over one billion people worldwide [1]. These diseases are typically treated with inhaled therapeutics which are delivered to the affected tissue to prevent symptoms such as breathlessness from arising, or relieve them when they do arrive. Many patients' symptoms are worsened by a combination of poor inhaler technique and a lack of adherence to their treatment plan [2]. Electronic health (e-health) measures such as smart inhalers [2] are being introduced to improve adherence and quality of care [3]. Smart inhalers can use sensors to track how often patients use inhalers and also provide estimates on technique or inspiratory flowrate based on acoustic sensors [2, 4]. The value of e-health monitoring

for patients could be increased significantly by linking measurements such as inhalation flowrate with patient-specific physical models that accurately predict how the inhaled drug has deposited in patient lungs.

Patient-specific modelling of drug deposition can be done with computational particle-fluid dynamics (CPFD) which solves the governing equations of air and drug particle transport during inhalation [5–7]. Patient-specific CPFD modelling requires 3D representations of the patient's respiratory system, usually obtained from clinical computed tomography (CT) scans [8]. Typically the airways are segmented using manual approaches, which can take several hours [9], and is therefore unsuitable for usage in future e-health monitoring frameworks. Tschirren et al. [10] showed that manually correcting airway segmentations obtained from semi-automatic methods can also take several hours. Such a long and labour-intensive process is a substantial bottleneck in applying patient-specific deposition models to a large number of respiratory patients. Fully-automatic segmentation methods have recently become available, which are promising for rapid segmentation of lungs and airways from volumetric CT using convolutional neural networks (CNNs) [11–14]. CNNs apply varying levels of convolution, where the convolution kernels are learned by evaluating the difference between the CNN output compared to a 'ground truth' human-verified result [15]. CNNs have been widely used in medical image processing and are used here to segment airways from 3D images such as clinical CT. However, CT images are more costly to acquire than chest X-rays and are therefore not optimal for imaging a large number of respiratory sufferers. Additionally, using CT scans is not encouraged in children for routine assessments due to harmful radiation. Finally, there are a variety of scenarios in which clinicians will use different tools (including modelling outputs) to make decisions [16]. For example, the recent Salford lung study compared metered-dose inhalers and dry-powder inhalers in several thousand patients [17, 18]. Modelling could support similar studies, but performing full patient-specific CPFD analyses based on CT data would pose high cost and time requirements. This simulation cost would be significantly lowered by an X-ray based image-processing tool and a rapid surrogate model to emulate CPFD simulations, the latter is subject of future work. Therefore, an additional image-processing approach that can extract 3D respiratory systems from sparser imaging such as chest X-rays is an unmet need which has been addressed by the present study.

To perform deposition simulations based on patient chest X-rays, we require a novel image-processing approach to reconstruct 3D respiratory geometries from 2D chest X-rays. Reconstruction of 3D skeletal anatomies including femora, vertebra or knee joints from planar and bi-planar X-ray images has been performed using statistical shape and appearance models (SSAMs) [19–24]. Statistical shape models (SSMs) take a set of corresponding geometries (usually represented with landmarks that represent key features), and quantifies how shape features change across a population (based on eigenvalues and eigenvectors of the covariance matrix) [25]. A SSAM performs the same operation, including an additional feature such as gray value at each landmark. In this approach, patient lungs may be extracted from a chest X-ray by combining (i) available information such as lung outline and gray-value distribution, as well as (ii) a SSAM that describes how lung shapes and projected gray-value varies from a database of known (3D) lung shapes [25–27]. Respiratory statistical shape models (SSMs) have previously been used for diagnostic purposes, as Irving et al. developed an approach to extract 3D airways from a single chest X-ray and diagnose tuberculosis in children [28, 29]. A lung shape model was also used to quantify differences in lung morphology in smokers and non-smokers [30]. Recent 2D-3D reconstruction studies have reconstructed full CT scans from a single or limited number of X-ray images using deep learning approaches [31–33], which would require an additional segmentation step to extract the final respiratory geometry. This approach is purely based on the pixel information, and does not have any constraint on the anatomical shape [34], allowing erroneous CT reconstructions to cause noise in the segmented geometry. A simpler approach would be to use a SSAM as a generator for possible shapes and evaluate their fit to the given X-ray, as done by Baka et al. [20].These examples of SSM and SSAM applications in skeletal and respiratory imaging show that SSAMs are a suitable approach for extracting full respiratory systems from planar or bi-planar chest X-rays. However, a key issue in this approach is that only the first bifurcation-level (or 'generation') of the airway tree is visible on a chest X-ray [29], as the low density lung structures are overlapped by denser structures such as the spine, rib cage and heart [35–37]. Similarly, only around 5-10 generations (out of 23) are

visible on high resolution CT scans [8, 38]. Thus, even with high resolution imaging, the deep lung airways are left unaccounted for and must be included for patient-specific deposition predictions.

The limitations of imaging resolution have been overcome in several ways. The remaining levels of the conducting airway tree can be generated (approximately up to generation 16) using statistical models based on information from the visible airways [38–40]. The remaining portion of the airway tree (bifurcations 17 to 23) are the respiratory 'acinar' region, which is a very small (characteristic length is $\mathcal{O}(100\,\mu\text{m})$) and intricate structure that is impossible to image *in vivo* [41–43]. Acinar flows have been studied with CPFD [43–46], but it is highly computationally expensive to directly resolve the full conducting airways and deep lung in one simulation [47–49]. Typically, the deep lung mechanics are treated with reduced-dimension models such as 1D airway networks or 0D lumped parameter models which are coupled to 3D CPFD simulations at the 3D airway outlet [50–58]. To model the full respiratory system, we therefore couple 3D CPFD domains produced from our image-processing framework with a 0D lumped parameter model to validate the entire image-processing and simulation framework against experimental deposition data [59, 60].

Therefore, we aimed to (i) develop a computational framework to reconstruct 3D airways and lungs from 2D image(s); (ii) develop a segmentation tool for reconstructing airways and lungs from clinical CT data; (iii) evaluate the ability of the developed image-processing tools to reproduce morphological properties and deposition of the ground truth airways, and (iv) validate deposition predictions against experimental data. Our developed computational framework is made publicly available online (`https://github.com/jvwilliams23/respiratory2Dto3Dpaper`).

# 2 Methods and materials

## 2.1 Patient image data and ground truth generation

We used 51 clinical CT scans from the LUNA16 challenge [61], which is part of The Cancer Imaging Archive (TCIA) [62]. The data was accessed on the 2nd of March 2020 and is fully anonymized. The scans were acquired for lung cancer nodule detection which does not depend on inhalation state, therefore no information on breath-hold procedures was reported [61]. The scans used were chosen randomly by Tang et al. [14], and they provided radiologist-verified lung lobe segmentations online. We created airway segmentations for the same scans using a semi-automatic region-growing approach, based on work by Nardelli et al. [63]. These segmentations have been independently verified by a radiologist with 9 years experience. We excluded 13 samples from the LUNA16 dataset, as the segmentations only included the trachea and first bifurcation. The remaining 38 samples were used for training the models discussed in the following sections.

To validate our modelling approach, we compared to anonymized experimental data of radiopharmaceutical aerosol deposition from Conway et al. [59, 60], which we refer to throughout as the 'Southampton/Air Liquide dataset' (accessed on 21st of May 2021). They performed experiments with inhaled aerosol in six healthy and six asthmatic subjects. Patients sat in an erect (sitting) position while aerosol was delivered through an AKITA nebuliser. The AKITA nebuliser is breath-actuated to release drug only during inhalation, and a controlled (constant) inhalation profile to optimise deposition [64]. Deposition was imaged with single photon emission computed tomography (SPECT) and combined with low resolution X-ray CT captured in the same machine (SPECT-CT), which relates deposition images to anatomical structure of the lungs [65]. To relate deposition measurements to airway morphology, high resolution CT was acquired for each patient on a different machine on a different day [59, 66]. Full information on scanning equipment and protocol are provided elsewhere [59, 66], and only key points related to our study are discussed here. High resolution CT images were acquired in supine position, which is known to decrease lung volume compared to erect position [67]. Additionally, Conway et al. discussed imaging issues in the upper airway [59]. In half of the images acquired, the epiglottis was completely closed, meaning there was no path for air to pass from the mouth to the lung airways. This may have occurred since the images were taken during exhalation, where the glottis area is at its narrowest [68, 69]. Furthermore, motion artefacts may have blurred the already narrow glottis to appear fully closed. To allow for air and drug particles to pass from the mouth to the lung, we manually cleaned up segmentations in these problematic cases (Supplementary Material Section 1).

## 2.2 Shape reconstruction from a chest X-ray image

To reconstruct 3D lung and airway geometries from an X-ray image, we created a statistical shape and appearance model (SSAM) based on ground truth airway and lung lobe segmentations. The appearance is based on digitally reconstructed radiographs (DRRs) derived from the patient CT data (described below). The DRRs were also used for testing how well the SSAM fits to an X-ray. As there was no available dataset of paired chest CT scans and chest X-ray images for validation of our 2D-3D reconstruction, it is common to use DRRs for testing reconstruction methods [31, 32]. Below we describe the process to create a SSAM from patient CT data and segmentations described above (Section 2.1).

### 2.2.1 Pre-processing and landmarking

To landmark the lungs we used the growing and adaptive meshes algorithm [70] to mesh the lung lobes. This is an automated, unsupervised method which takes points from a surface mesh (STL file) as an input. It begins with a small of landmarks randomly selected from a baseline surface mesh. The number of landmarks are iteratively increased until the accuracy of how well the shape is described converges. This is then adapted to the rest of the dataset, preserving the correspondence between landmarks and anatomical features. We did not consider lobar fissures in the landmarks since they are not visible on the chest X-ray. This gave us 2966 landmarks in total for the lungs.

To landmark the airways, we skeletonised the segmentation [71]. We included the trachea, main and lobar bronchi in the skeleton landmarks, as airways below this will not be visible on the X-ray image. We uniformly sampled 10 points across the length of each branch as landmarks to capture any curvature. At each landmark along the centerline, we also extracted landmarks to describe the surface by sampling points around the circumference of the cross section.

We created DRRs for all (remaining) 38 patients in the LUNA16 dataset which we used to train the SSAM and to assess SSAM performance. This was done by integrating the Hounsfield units for each voxel in the anterior-posterior direction of the scan to create a frontal chest X-ray. We also created additional DRRs in the sagittal plane to test the influence of including an additional projection on the morphologies predicted by the SSAM. The gray-value at the landmark location could then be used in the appearance part of the SSAM. The DRRs represent patient X-rays at the same time point that the CT data used to extract landmarks was taken. The DRRs were also used for assessing our SSAM performance, as we can compare the 3D reconstructions of the lungs and airways from the SSAM to CT at the same instant. By comparing lungs and airways at the same time instant, our results are not influenced by changes in lung or airway shape during breathing.

Similar to Baka et al. [20], we computed a Canny edge map of the chest X-ray to use for guiding our SSAM fitting. To reduce noise in the edge map, we first downsampled the X-ray to be four times coarser and applied global histogram equalization. We then applied a Canny edge map [72], where the Gaussian kernel has a width of two pixels. We found this combination of pre-processing steps to produce an edge map with the lowest amount of noise from overlapping structures such as the rib cage.

### 2.2.2 Statistical shape and appearance modelling

Once a set of landmarks and DRRs were extracted from the LUNA16 dataset (Section 2.2.1), all sets of landmarks were aligned to zero mean and isotropic scaling was applied such that the coordinates had zero mean and unit standard deviation. Landmark gray values were also normalised to zero mean and unit standard deviation per sample. We then create a SSAM to describe and parameterise the covariance of the lung and airway shape-appearance data in the LUNA16 dataset. The SSAM parameters are then iteratively adjusted based on how well the lung and airway shape and appearance agrees with the X-ray image (Fig 1a-b).

The SSAM uses principal component analysis (PCA) to create a model of the variation in the shape and the gray-values of the corresponding chest X-ray image(s). The state vector containing shape and gray-value at each landmark, $i$, is $\boldsymbol{z}_i = (\boldsymbol{x}_i, g_i^1, ..., g_i^{N_{XR}})$, where $\boldsymbol{x}_i$ is the landmark coordinate, $g_i^1$ is the landmark gray-value for the first projection, $N_{XR}$ is the number of X-ray images per patient (each with its own $g$). In this study we use $N_{XR} = 1$ and $N_{XR} = 2$. In both cases, a frontal X-ray is

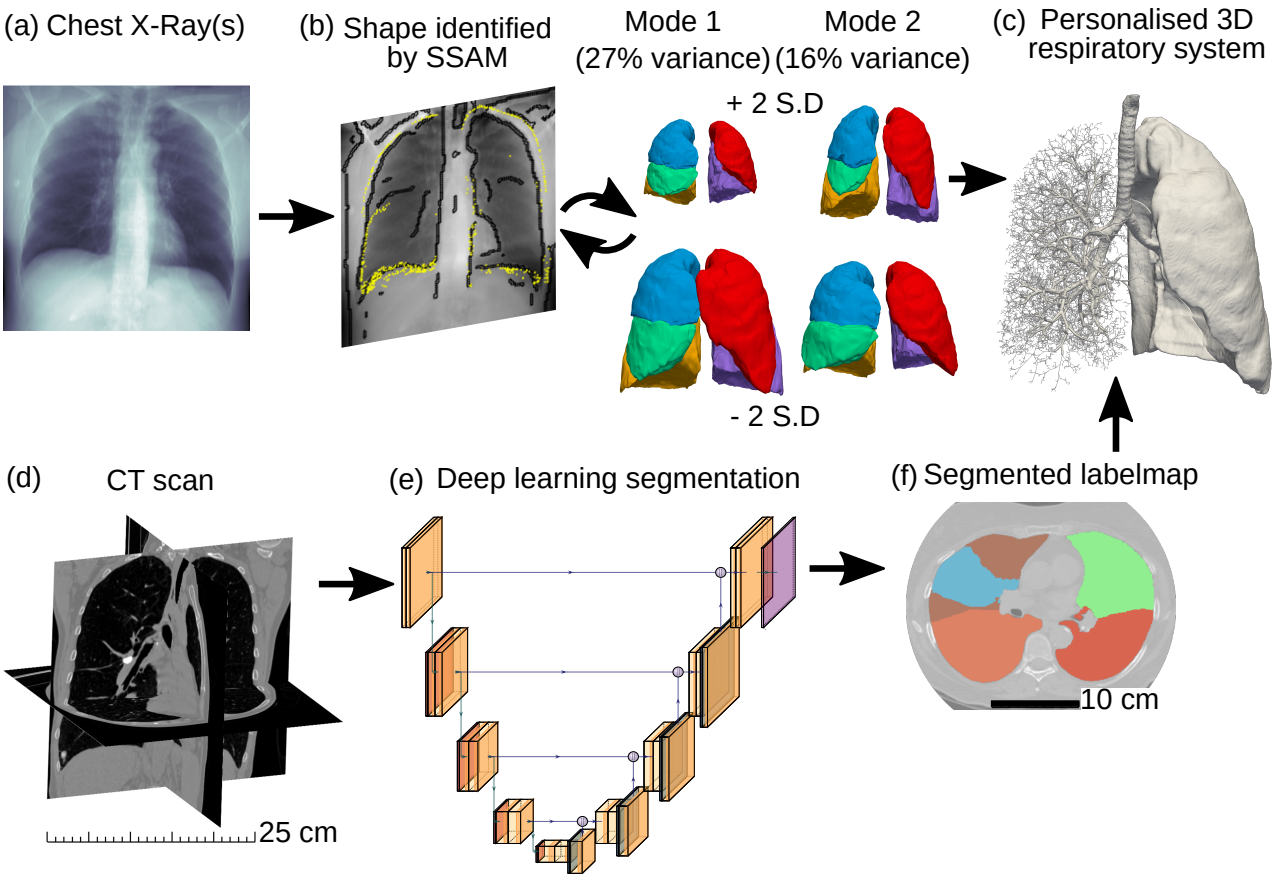


**Fig 1.** Schematic overview of image-processing framework. Chest X-ray image(s) are provided as an input (a), where the shape is identified by our SSAM by iteratively adapting lung outline and gray-value based on a database of known lung shapes (b-c). The SSAM reconstruction can then be used to generate a surface mesh of lungs with a full airway tree (c). Alternatively, a 3D CT scan may be provided as an input (d) to our trained CNN architecture (e). The CNN returns a segmented labelmap for lungs and airways (f), which can also be used to generate the full respiratory domain (c).

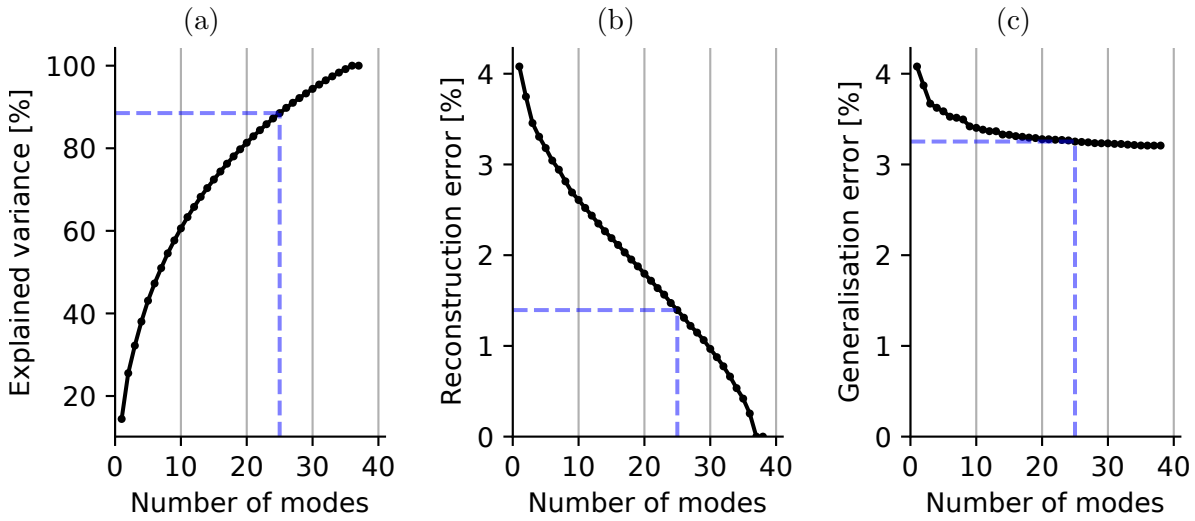


**Fig 2.** Assessment of sensitivity of SSAM metrics to number of modes included in the model. Panels show (a) explained variance based on number of modes, (b) reconstrucion error for shapes in training set, (c) generalisation error for unseen shapes. When the model describes 90% of the population variance (indicated by blue dashed line), the generalisation error in adapting the model to unseen shapes is saturated. Tests were performed using the final 38 samples from the LUNA16 dataset.

used (Fig 1a). In tests with $N_{XR} = 2$, we also provided a lateral view X-ray image. The state vector containing all landmarks is a column vector, $\boldsymbol{z} = (\boldsymbol{z}_1, ..., \boldsymbol{z}_{N_L})$, where $N_L$ is the number of landmarks. Therefore, $\boldsymbol{z}$ has a size of $(3+N_{XR})N_L$. A new state vector can be approximated by the linear function

$$\boldsymbol{z} \approx \bar{\boldsymbol{z}} + \sum_{m=1}^{N_m} \boldsymbol{\Phi}_m \cdot \left( \boldsymbol{b}_m \sqrt{\boldsymbol{\sigma}_m^2} \right), \tag{1}$$

where $\bar{\boldsymbol{z}}$ is the mean of the state vector (averaged over all samples in the training dataset). $\boldsymbol{\Phi}$ is a $(3 + N_{XR})N_L \times N_m$ matrix, giving the main modes of variation of the training set, where $N_m$ is the number of modes. $\boldsymbol{\sigma}_m^2$ is the variance described by mode $m$. $\boldsymbol{b}_m$ is the shape parameter for mode $m$, where Fig 1(b) shows the effect of varying the weighting of $\boldsymbol{b}$ for the first two modes. By using PCA, the complexity of the model can be reduced by including only the number of modes, $N_m$, that are required to capture a specified amount of variance in the modelled shape. In this study, 90% of the population variance is included in the model ($N_m = 25$, Fig 2a). The SSAM was created using our Python library '`pyssam`' for statistical shape and appearance modelling [73].

The accuracy obtained by the SSAM with $N_m$ modes can be assessed in terms of its ability to reconstruct a shape in the training dataset (reconstruction error). More importantly, the SSAM can be assessed by its ability to generalise and reconstruct an unseen shape (not in the training set), which is known as generalisation error. When the shape model contains 100% of the training population variance (Fig 2a), the reconstruction error should tend to zero as all information required to reconstruct the shape is contained within the model (Fig 2b). We assessed reconstruction and generalisation error with leave-one-out testing, which was averaged over the 30 loops of the entire training set. The mean absolute error is presented, which is expressed as a percentage of the lung bounding-box size. We observed that $N_m = 25$ is a suitable choice to reduce the SSAM dimensionality, as the generalisation error did not increase further by including additional modes (Fig 2c).

The shape parameters, $\boldsymbol{b}$, can then be optimised with respect to the fit of the SSAM outline ($\mathcal{L}_{fit}$) and the similarity to the training set ($\mathcal{L}_{prior}$). The appearance model optimises the shape by minimising the gray-value loss ($\mathcal{L}_g$). Gray-value loss is calculated as the difference between the image gray-value at the landmark location and the expected gray-value from Eq 1 [21, 74, 75]. We also used the anatomical shadow loss ($\mathcal{L}_{AS}$) for airways as proposed by [29]. Anatomical shadow compares local gray-values inside a small control area inside and outside of the projected airway surface, which can enhance airway contrast compared to dense overlapping structures such as the spinal column. Therefore, the optimal

isotropic scaling, translation and shape parameters are found by minimising the equation

$$\mathcal{L} = C_{fit}\mathcal{L}_{fit} + C_{prior}\mathcal{L}_{prior} + C_g\mathcal{L}_g + C_{AS}\mathcal{L}_{AS} \tag{2}$$

where $C$ is a coefficient that controls the weight of each fitting term.

The fit of the projected landmarks outline ('silhouette' landmarks) to the lung outline on the X-ray image is found by

$$\mathcal{L}_{fit} = 1 - \frac{1}{N_{SL}} \sum_{i=1}^{N_{SL}} D_i \tag{3}$$

where $N_{SL}$ is the number of silhouette landmarks, $D_i$ is a function based on the distance between each silhouette landmark and the nearest X-ray image outline point. Silhouette landmarks were defined as landmarks where the nearest surface mesh point is shared by two faces, with one face normal pointing towards and another away from the image source. For example, for an anterior-posterior (frontal) chest X-ray image, the projection is normal to the $y$-direction and therefore the two faces nearest to the silhouette landmark will have positive and negative $y$ components in the normal vector. The silhouette landmarks represent the outline of the shape normal to the projection [20].

Baka et al. [20] normalised the outline distance to bounds of [0,1] by

$$D_i = \exp\left( - \frac{\min(||\boldsymbol{x}_{\mathrm{SSM},i} - \boldsymbol{x}_{\mathrm{XR}}||)}{C_{dist}} \right). \tag{4}$$

where $C_{dist} = 5$ was used [20]. $\mathcal{L}_{fit}$ is not evaluated for the airways as the airway outline extracted from a Canny edge map is weak and unreliable due to overlapping denser regions on the X-ray image.

Improbable shapes (relative to our training set) are penalised by the Mahalanobis distance between the created shape and the mean shape [20, 70] calculated by

$$\mathcal{L}_{prior} = \frac{1}{N_L} \sum_{i=1}^{N_L} \sqrt{(\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^T \operatorname{cov}[\boldsymbol{x}, \bar{\boldsymbol{x}}]^{-1} (\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)}, \tag{5}$$

where $N_L$ is the total number of landmarks, $\operatorname{cov}[\boldsymbol{x}, \bar{\boldsymbol{x}}]^{-1}$ is the inverse covariance matrix between $\boldsymbol{x}_i$ and $\overline{\boldsymbol{x}}_i$.

We use the appearance model to minimise the difference between the modelled gray-value ($g_{model}$) and the gray-value at the landmark location ($g_{target}$). This is found by

$$\mathcal{L}_g = \frac{1}{N_L} \sum_{i=1}^{N_L} |g_{model,i} - g_{target,i}|. \tag{6}$$

Finally, the anatomical shadow loss is found for each silhouette landmark [29]. This is done by creating two regions-of-interest (ROIs) at some distance $\pm s_{AS}$ normal to the border of the silhouette landmarks. Each ROI is a circle with radius $r_{AS}$, and the gray-value 'inside' and 'outside' of the airway projection are compared as

$$\mathcal{L}_{AS} = \frac{1}{N_{SL}} \sum_{i=1}^{N_{SL}} \frac{g_{inside,i} - g_{outside,i}}{g_{outside,i}}. \tag{7}$$

We found the optimal parameters to be $s_{AS} = 20$ pixels and $r_{AS} = 14$ pixels.

Shape parameters were optimised using the NGO (Nevergrad optimiser) algorithm [76] in the Nevergrad Python library [77]. This is a competence map method, which adjusts the optimisation algorithm based on the number of optimisation iterations and number of parameters optimised. We chose this method as it can therefore adjust to more modes of variation as our dataset grows in future studies. We found this gradient-free optimisation to show good robustness against local minima. The shape parameters for each mode were initialised to $\boldsymbol{b}_m = 0$, and were bound to $\boldsymbol{b}_m \pm 3$.

SSAM hyperparameters in the loss function (Eq 2) which produced the lowest maximum error in reconstructed lung space volume were found through Gaussian process regression [78, 79], also known as Kriging [80]. We performed 100 optimisation iterations, using a subset of 11 randomly selected samples

from the LUNA16 dataset. This gave the optimal model parameters to be $C_{fit} = 0.795$, $C_g = 0.687$ and $C_{prior} = 4.4 \times 10^{-4}$. The anatomical shadow loss coefficient was later found through a simple grid search as $C_{AS} = 0.2$.

Landmarks obtained from fitting the SSAM to an unseen shape (called 'new landmarks') were converted to a surface by morphing a template mesh and corresponding template landmarks to the 'new' landmarks [81]. New mesh points ($\boldsymbol{p}_{i,\text{new}}$) are created from the new landmarks ($\boldsymbol{x}_{i,\text{new}}$), template mesh points ($\boldsymbol{p}_{i,\text{template}}$) and template landmarks ($\boldsymbol{x}_{j,\text{template}}$) by

$$\boldsymbol{p}_{i,\text{new}} = \boldsymbol{p}_{i,\text{template}} + \sum_{j}^{N_L} k(\boldsymbol{p}_{i,\text{template}}, \boldsymbol{x}_{j,\text{template}}) w_j \tag{8}$$

where $k$ is a kernel used for morphing the mesh and $w_j$ represents the weights controlling how much each landmark point is morphed. The weights are computed by solving (8) for $w_j$ and setting $\boldsymbol{p}_{i,\text{new}} = \boldsymbol{x}_{j,\text{new}}$. We used a surface morphing algorithm [81] with a Gaussian kernel, as we found it to produce a smoother deformation when morphing. The kernel is $k(a, b) = \exp(||a - b||/2\,\sigma^2)$, where $\sigma = 0.3$ [82].

### 2.3 Segmentation from volumetric CT

To segment lung lobes and airways from volumetric CT scans (Fig 1d), we used convolutional neural networks [15] such as the U-Net shown in Fig 1e [83, 84]. We implemented the U-Net in PyTorch [85]. The U-Net architecture was detailed by Ronneberger et al. [83]. For lung segmentation we used a pre-trained U-Net [12]. For segmenting the airways we trained a U-Net using segmentations from the LUNA16 challenge (Fig 1e). Prior to training, all images and labelmaps were cropped to the extent of the airway labelmap to conserve memory and allow us to segment the entire 3D image, instead of slice-by-slice [12], or sliding-box approaches [86]. With the aim of lowering memory consumption, we also trained an ENet ('efficient neural network') [87]. The ENet has approximately 15 times fewer parameters than the U-Net architecture [88], which creates lower memory consumption during training and faster processing time for 3D images [87, 89].

The loss function used for training all CNNs was [14, 90], given as

$$\mathcal{L} = \mathcal{L}_{DICE} + \mathcal{L}_{focal}, \tag{9}$$

where $\mathcal{L}_{DICE}$ is the DICE loss and $\mathcal{L}_{focal}$ is the focal loss. The DICE loss maximises the total intersection between a modelled labelmap and the ground-truth labelmap during training, and is found by

$$\mathcal{L}_{DICE} = \frac{1}{CN} \sum_{c}^{C} \sum_{i}^{N} \left( 1 - \frac{p_{ic} g_{ic}}{p_{ic} g_{ic} + (1 - p_{ic}) g_{ic} + p_{ic}(1 - g_{ic})} \right), \tag{10}$$

where $p_{ic}$ is the probability of a voxel $i$ belonging to class $c$ as predicted by the CNN. $g_{ic}$ is the ground truth value for the same voxel (which will be 1 or 0). This is averaged over all voxels in the image $N$ and the total number of classes $C$.

The focal loss weights the loss more towards 'harder to learn' voxels with a focusing parameter, $\gamma$. Focal loss was developed for images with high class imbalances, such as airway segmentations, where we found the number of voxels representing the airways to be on average 63 times fewer than the number of background voxels. The focal loss was found by

$$\mathcal{L}_{focal} = -\frac{1}{CN} \sum_{c}^{C} \sum_{i}^{N} \alpha g_{ic} (1 - p_{ic})^{\gamma} \log(p_{ic}), \tag{11}$$

where we set the focusing parameter $\gamma = 5$ based on a parametric study of training with $\gamma$ set to 1, 3 and 5 (Supplementary Material Figure S4). We used $\alpha = 1$ as recommended by [14].

To train the U-Net and ENet, we used the Adam optimiser with a batch size equal to one to minimise memory consumption. Data was augmented with rotations of $\pm 15°$. We created 50 new augmented datasets per epoch. Training was stopped once the validation loss had not decreased further in the most recent 10 epochs. To train the U-Net and ENet models, we used a compute node with 756

GB of RAM on Heriot-Watt University's high performance computing cluster ROCKS. Training took 3 days for the U-Net (42 epochs) and 5 hours for the ENet (45 epochs). The ENet DICE coefficient was lower than the U-Net (0.89 compared to 0.93, Figure S5a). The ENet had a significantly improved inference time compared to the U-Net (ENet gave 5.7 times speedup, Figure S5b).

### 2.4 Distal airway generation

To create the rest of the conducting airways not visible in CT or from our airway SSAM, we implemented a volume-filling airway generation algorithm [38, 39]. The algorithm generates a full conducting airway tree based on information from the upper airways such as branch length and diameter. The lung space volume was discretised into seed points with a uniform spacing found by $\Delta = (V_{lung}/N_T)^{1/3}$ where $N_T = 30,000$ is the number of terminal nodes in the conducting airway tree.

The algorithm for generating conducting airways can be found in detail elsewhere [38]. We will briefly cover it here. Seed points were first clustered based on their nearest terminal branch in the airway tree. The centroid of each cluster was found and a splitting-plane was defined based on the centroid and the node of the parent branch. This splitting-plane was used to divide each seed point cluster into two more clusters. A new branch was generated extending towards each seed point set by 40% of the distance to the centroid. When the seed point set had only one node, or the branch length was less than 2 mm, the branch is classed as a terminal bifurcation and the nearest seed point was deleted. This was repeated until no seed points remained in the tree. Once the tree was grown, a diameter was assigned to each branch segment as described in Supplementary Material Section 5. In our CPFD simulations, the image-based airways are used until generation 3 and airway tree generated from the volume-filling algorithm [38] is used to generate a surface from generations 3 to 6, similar to [91].

### 2.5 Deposition simulation configuration

To prepare the airway segmentations of the LUNA16 dataset for deposition simulations, we added a generic mouth-throat geometry obtained from a healthy adult patient [92] as done in [7] in cases where the upper airways were not imaged (chest CT, as is most common). This was done by scaling the surface mesh of the healthy patient's upper airways such that the diameter at the interface between healthy patient upper airways and the patient's trachea reconstructed was the same [7]. As jet-like flow structure produced in the upper airways has a key influence on the swirling flow observed in the trachea [93], it is important to include at least a generic representation of the upper airways in respiratory CPFD models [7, 93]. Images in the Southampton/Air Liquide dataset contained the lungs and the mouth-throat region (unlike the LUNA16 dataset). Therefore, we used these images directly and obtained segmentations of the mouth and throat by region-growing. We did not use the CNN for this as the mouth-throat region was not present in the training data. Also, this region can easily be segmented with region-growing [11]. The mouth-throat segmentation was then joined to the trachea directly.

The deposition simulations in this study were performed using OpenFOAM v6 [94] using our custom solver deepLungMPPICFoam [95]. Briefly summarised, we solved mass and momentum conservation equations for transport of an incompressible fluid (air). As described in [7], our computational fluid dynamics mesh had a spacing of $\Delta = 500\,\mu m$, which gave excellent agreement with flow through 3D printed airways [92]. The fluid turbulence modelling is discussed in detail in [96]. Particles were tracked in a Lagrangian approach by solving Newton's equations of motion accounting for drag and gravity, as we showed in our previous study that particle-particle interactions are not influential [7]. Particles were deleted when they reached an outlet, and were given a 'sticking' condition when they hit the wall boundary. The number of particles tracked was 200,000 with diameter, $d_p = 4\,\mu m$ as used in our previous study [7]. In simulations with airways from the LUNA16 dataset, particles were released from the inlet over a period of 0.1 s with initial velocity as 10 m/s to represent a metered-dose inhaler [97]. More information on simulation configuration for comparing to the Southampton/Air Liquide dataset is given later in this section.

To drive flow in the simulation, we model pressure at the outlet of the 3D domain using a lumped-parameter model based on a resistance-compliance circuit [51]. In this model, the pleural pressure required to drive flow during inhalation, $p_d$, is found by

$$p_d(t) = R_{global}Q(t) + \frac{V(t)}{C_{global}} - p_{atm}, \tag{12}$$

where $R_{global}$ is the global resistance of the circuit, $C_{global}$ is the global capacitance, $Q(t)$ is the flowrate as a function of time, $V(t)$ is the volume of air inhaled as a function of time, and $p_{atm}$ is the atmospheric pressure. As the flow is incompressible, there is no effect of atmospheric pressure and therefore it is dropped ($p_{atm} = 0$). In simulations of inhalation in airways from the LUNA16 dataset, we model tidal inhalation where the volume of air inhaled to the lung is [51]

$$V(t) = -\frac{1}{2}\left[V_T \cos\left(2\pi\frac{t}{T_B}\right) - V_T\right] \tag{13}$$

where $V_T$ is the tidal volume and $T_B$ is the time for one breath. Oakes et al. estimated respiratory parameters such as $R_{global}$, $C_{global}$, $V_T$ using empirical relationships which are often based on parameters such as age, weight, height, gender [51]. We used the adult parameters from [51], $R_{global} = 7 \times 10^{-3}\,\mathrm{cm\,H2O\,s/mL}$ and $C_{global} = 59\,\mathrm{mL/cm\,H2O}$. Pressure at the outlet is then calculated by solving the following equation at each outlet (subscript $_{(i)}$) at each time-step of our simulation:

$$p_{(i)} = R_{(i)}\phi_{(i)} + \frac{V_{(i)}}{C_{(i)}} + p_d. \tag{14}$$

where $\phi_{(i)}$ is the flowrate (also called 'flux') at the outlet computed from the fluid solver. Volume at the outlet is then found by integrating the flux in time. Local compliance and resistance at each outlet are estimated based on

$$C_{(i)} = \frac{A_{(i)}\alpha_{(L)}}{A_{(L)}}C_{global} \quad R_{(i)} = \frac{A_{(L)}}{A_{(i)}\alpha_{(L)}}R_{global} \tag{15}$$

where the $A_{(L)}$ is the sum of areas of all outlets in each lung, $L$. $\alpha_{(L)}$ is the fraction of (static) volume of a lung to the combined volume of both lungs, which assumes that the volume change during inhalation is a factor of the lung's volume only (not accounting for localised disease).

By defining Neumann boundary conditions for velocity at the inlets and outlets, the simulation can be unstable. Particularly, when the flow at the outlet is reversed (due to turbulence or local changes in geometry curvature) the flux at the face has a negative sign (creating positive pressure) and can cause numerical divergence [98, 99]. To prevent backflow instabilities, we (i) set maximum pressure and minimum flux at each outlet to zero, (ii) use limiters when calculating gradients at outlet faces to prevent small negative fluxes caused by interpolation errors. We also (iii) set the maximum Courant number to be 0.1 when the flow is developing, and Courant number is set to 0.2 when the flow is developed (quantified by $0.2 < |\sin(2\pi t/T_B)|$).

Simulations of the Southampton/Air Liquide dataset modelled inhalation from a breath-actuated AKITA nebuliser [59]. This provided a constant inhalation flowrate of 18 L/min. Therefore, we assigned $V(t) = V_T\, t/T_{inhale}$, where $T_{inhale}$ is the time for one inhalation, $T_{inhale} = T_B/2$. In the experimental study, inhalation duration was varied between 2 and 3.3 s (shallow and deep breathing). However, Fleming et al. [60] observed no significant difference in regional deposition when changing from shallow to deep inhalation. As our model cannot account for particle motion in the deep lung, we cannot model particles re-entering the central airways from the deep lung during exhalation. Therefore, we used an inhalation duration of 2 s for our simulations to reduce simulation clock-time, since the longer inhalation and breath-hold would only influence particle motion in the deep lung (increasing time for deposition due to diffusion or settling). As the inhalation flowrate is steady and we could not model exhalation, we only modelled one inhalation. This was also done by de Backer et al. [100] to compare with SPECT-CT data, although they used a steady solver which cannot capture transient flow features.

In the Southampton/Air Liquide experimental study [59, 60], the particle diameters were $d_p = 3.1$ and 6.05 µm. To simplify our analysis, we used monodisperse particles as the experimental size

distribution was reported as being narrow [59]. Simulated particles were released from a 5 mm disk in the AKITA mouthpiece, which represented the diameter of the tubing that delivers the particles from the nebuliser. Simulated particles were released with initial velocity of 0 m/s, since the particle timescale was small it would quickly adapt to the local fluid velocity in the mouthpiece ($\tau_p = 28\,\mu s$ for $d_p = 3.1\,\mu m$ and $\tau_p = 108\,\mu s$ for $d_p = 6.05\,\mu m$).

Simulations of flow in airways from the LUNA16 dataset were performed using ROCKS with 28 CPUs each (Intel Xeon Silver), taking between one and three days. Simulations of flow in patient airways from the Southampton/Air Liquide dataset were performed on Oracle cloud with two bare metal compute nodes each with 36 CPUs (BM.optimised3.36, Intel Xeon Gold), taking between one and three days depending on the airway geometry.

### 2.6 Model evaluation

To evaluate our model, we first aimed to verify our 2D-3D reconstruction algorithm and CNN segmentation by comparing reconstructed lung and airway morphological properties using the LUNA16 and Southampton/Air Liquide datasets (Fig 3). The morphological analysis was done using the U-Net for the airways in the LUNA16 dataset, and the ENet for airways in the Southampton/Air Liquide dataset (due to larger memory requirements). For both datasets, the CNN segmentation of the lobes was done using a pretrained 2D U-Net [12].

We then compared CFD deposition results from the 2D-3D reconstruction and the CNN segmentation against CFD simulations based on ground truth segmentations from the LUNA16 dataset (Section 2.1). Finally, we aimed to validate our drug deposition modelling by comparing deposition predictions in airways segmented from CT images with our CNN to *in vivo* SPECT deposition data described in Section 2.1 [59, 60]. As our SSAM did not contain training data from the mouth-throat region, we did not test the SSAM against the Southampton/Air Liquide deposition measurements.

The Southampton/Air Liquide experimental data contained measurements from deposition in six healthy and six asthmatic patients. Two experiments were performed for each patient, testing varying combinations of aerosol size, breathing pattern (shallow/deep) and inhaled gas (air compared to He/O2) giving 24 total measurements. We omitted cases with He/O2 inhaled gas (4/24 experiments). Images in three asthmatic patients had blockages in the upper airways due to tongue positioning, which meant we could not segment the entire upper airways and these patients were omitted (4 of remaining 20 experiments). Additionally, as our model does not account for particle transport in the deep lung or exhalation we omitted the 'shallow' inhalation in cases where shallow and deep inhalations were the independent variable compared (4 of remaining 16 cases). Therefore, we performed 12 simulations to evaluate our model against the available experimental data [59, 60].

For all CFD analyses, we evaluated deposition fraction computed as the number of particles depositing in a region (relative to the total number of inhaler particles). Deposition predictions were assessed based on regional drug deposition fraction (right and left lung) for both the LUNA16 and Southampton/Air Liquide datasets.

For CFD analyses of the LUNA16 dataset we also computed a local drug deposition concentration metric known as deposition enhancement factor (DEF) [101, 102]. We calculate this using the number of particles deposited within a fixed distance (1 mm, area $A_{conc} = \pi\,(1\,\text{mm})^2$) of the central point of each wall face [103]. This is made relative to the global deposition by

$$\text{DEF} = \frac{N_{p,deposit}(A_{conc})/A_{conc}}{N_{p,deposit}(A_{tot})/A_{tot}}, \tag{16}$$

where $N_{p,deposit}(A)$ is the number of deposited particles in a surface area $A$. $A_{tot}$ is the total airway surface area. We use this instead of the sum of the deposition efficiencies as the defined areas may overlap, and this therefore prevents particles being counted multiple times towards the global average. To minimise differences in total number of deposited particles or total surface area interfering with quantitative comparison between DEF in the ground truth and SSAM or CNN, we used the ground truth denominator from Eq 16 in DEF calculations in the SSAM and CNN [104].

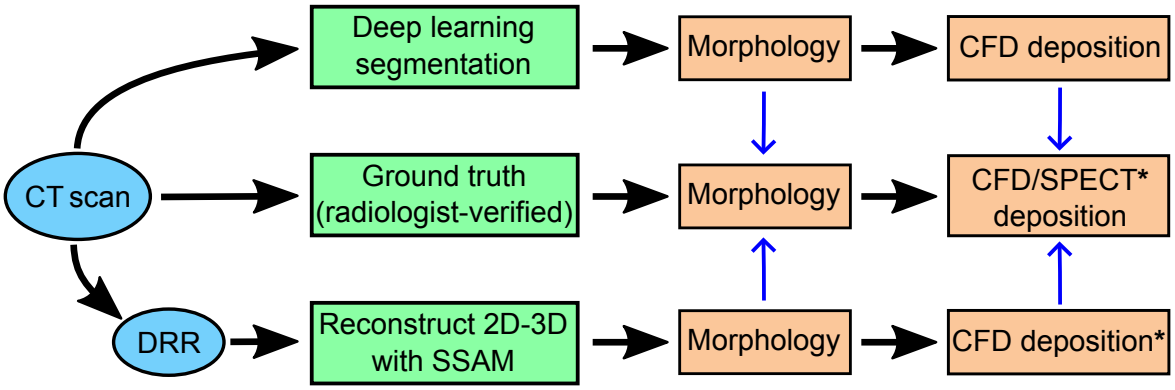


**Fig 3.** Flowchart overview of the model evaluation procedure. Model inputs are given in blue ellipses, image-processing methodology are in green boxes, outputs are in orange boxes. Blue arrows show model quantities validated against the ground truth. For the Southampton/Air Liquide dataset, ground truth deposition is obtained from SPECT and SSAM deposition is not computed (boxes with *).

# 3 Results

## 3.1 Morphological assessment on LUNA16 dataset

We found our SSAM to yield a relative lung space volume error of 9.9% median and a 95th percentile of 34.3% on the LUNA16 dataset (Fig 4a and Table 1). The maximum error for the right lung was 33.64% and 41.66% for the left lung. The concordance correlation coefficient between SSAM lung volume and ground truth lung volume was 0.893 with two DRRs used for SSAM fitting (Fig 4b), and 0.884 for one DRR in the SSAM (Figure S6a). The total volume error (averaged over left and right lungs) had a median value of 9.8% and a 95th percentile of 26.1%. Larger error in the left lung may be due to the presence of the heart interfering with the edge-map, which would cause a poor fit at the inner side of the lung. In contrast, the absolute relative error for the U-Net lung segmentation gave a median value of 1.35% and a 95th percentile of 3% (Fig 4). The U-Net is well-suited to segmenting lungs from CT, due to the high-contrast and well-resolved boundary between dark (air-filled) lung parenchyma and surrounding soft and hard-tissue. The SSAM lung reconstruction yielded a larger maximum error due to the poor contrast between lung and surrounding tissues on some radiographs.

The median lung volume error for the U-Net and SSAM were increased on the Southampton/Air Liquide dataset (Table 1). The correlation coefficient obtained by the SSAM was 0.73, which is a reduction from the 0.88 on the LUNA16 dataset. This may be improved by including a more diverse dataset in SSAM training, as was done for the lung U-Net by Hofmanninger et al. [12]. The inclusion of more diverse training data may explain how the U-Net was able to achieve similar correlation coefficients on both datasets.

In Fig 5 we compared the output SSAM landmarks against the input X-ray image and lung outline map in two outlier cases and the best case (in terms of lung space volume error, Fig 4). In both cases, we observed that the outer edge of the left lung has incorrectly fitted to a spurious edge introduced to the lung edge map. In the largest error case (Fig 5a), the lower corner of the SSAM landmark point cloud has fitted to the outline of the patient's torso. In the case with the second largest error (Fig 5b), the left lung outer edge and right lung lower edge have mistakenly fitted to an edge in the soft tissue. In the lowest error case (Fig 5c), there is a spurious edge for the outline of the patient's torso that slightly interferes with the outer edge of the patient's right lung. However, this is located very near the lung itself as the patient has a narrow torso, which therefore appears to cause a minor influence on the SSAM fitting.

Diameter predictions did appear to be sensitive to outliers, as the maximum diameter error was above 30% in the main bronchi (Fig 6). We compare airway diameter error with a single DRR (anterior-posterior plane projection) and two DRRs (anterior-posterior and sagittal plane projections) on the LUNA16 dataset. The maximum trachea diameter error was 20.4% for one projection and 21.9% for two projections, which shows the trachea is less sensitive to outliers than the main bronchi due to higher visibility on an X-ray image. The absolute trachea diameter error had a median value

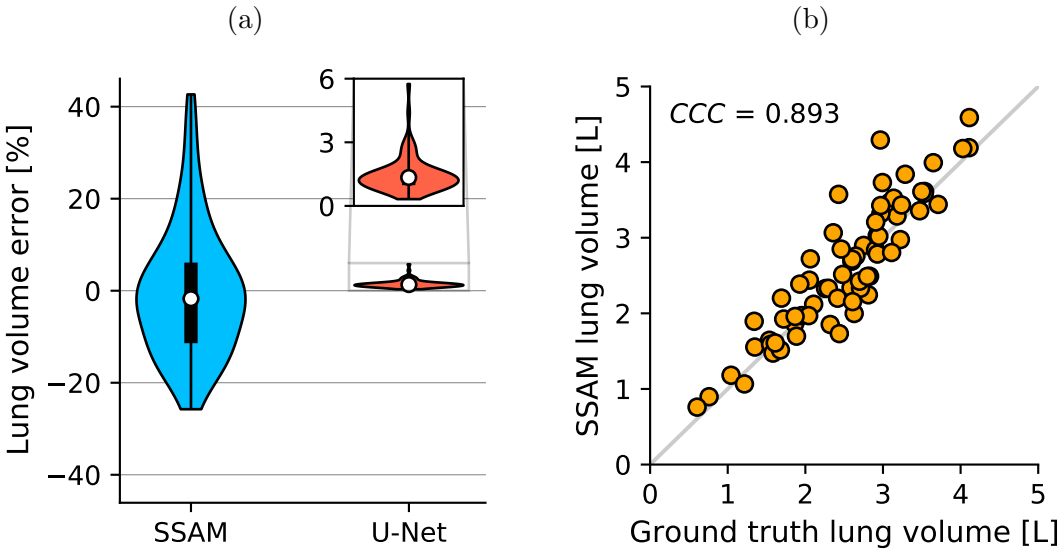


**Fig 4.** Evaluating SSAM error in lung space volume compared to ground truth lungs from the LUNA16 dataset. (a) Relative error from SSAM and U-Net reconstructions of patients from LUNA16 dataset. Inset shows U-Net lung space volume error, where the U-Net results were produced using a pretrained U-Net [12]. (b) Identity plot comparing absolute lung space volumes from SSAM and ground truth. $CCC$ is the concordance correlation coefficient.

**Table 1.** Summary of reported errors for various tests. "UoS/AL" refers to the Southampton/Air Liquide dataset [59,60]. Lung U-Net results were produced from a pretrained U-Net [12]. SSAM results shown are with two projections ($N_{XR} = 2$).

| Metric | Dataset | Model | Median | 95th percentile | $CCC$ |
|---|---|---|---|---|---|
| Lung volume | LUNA16 | SSAM | 9.9% | 34.3% | 0.89 |
| Lung volume | LUNA16 | U-Net [12] | 1.35% | 3% | 0.998 |
| Lung volume | UoS/AL | SSAM | 11.2% | 26.2% | 0.73 |
| Lung volume | UoS/AL | U-Net [12] | 6.5% | 8.7% | 0.97 |
| Airway diameter | LUNA16 | SSAM | 7.8% | 23.4% | 0.71 |
| Airway diameter | LUNA16 | U-Net | 5.5% | 12.2% | 0.96 |
| Airway diameter | UoS/AL | SSAM | 11.3% | 33.3% | 0.53 |
| Airway diameter | UoS/AL | ENet | 6.1% | 13.6% | 0.92 |
| Lung deposition | LUNA16 | SSAM | 8.2% | 15.9% | 0.64 |
| Lung deposition | LUNA16 | U-Net | 5.3% | 11.6% | 0.85 |

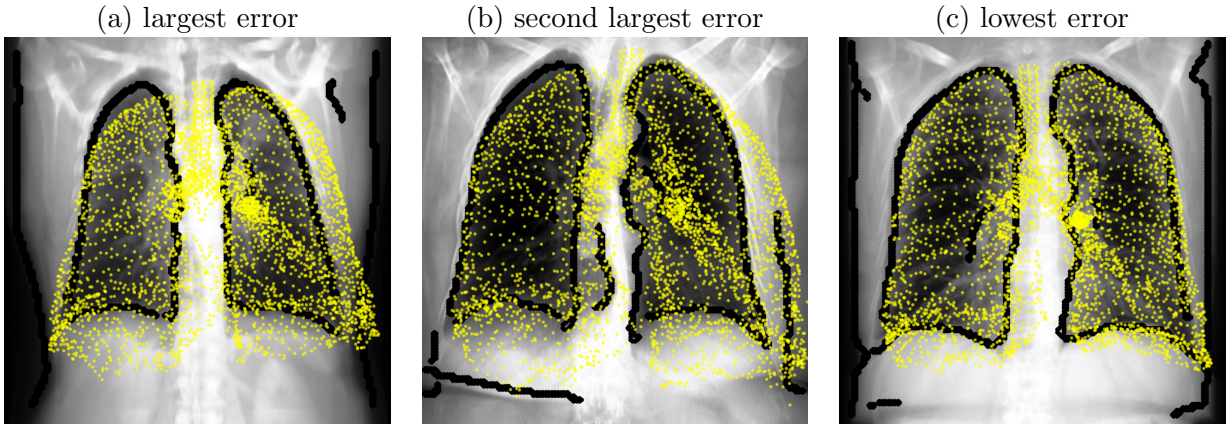


**Fig 5.** Example of fitted SSAM compared to projection and lung outline for two largest lung space volume errors in the LUNA16 dataset. Yellow markers are output landmarks from the SSAM, and black markers represent the X-ray edges detected by the Canny edge map described in Section 2.2.1. Panels show (a) the largest error, (b) second largest error and (c) the lowest error. In (a) the right lung error was 31% and left lung error was 42.6%. In (b) the right lung error was 15.5% and left lung error was 38.3%. In (c) the right lung error was 1.8% and left lung error was 0.06%.

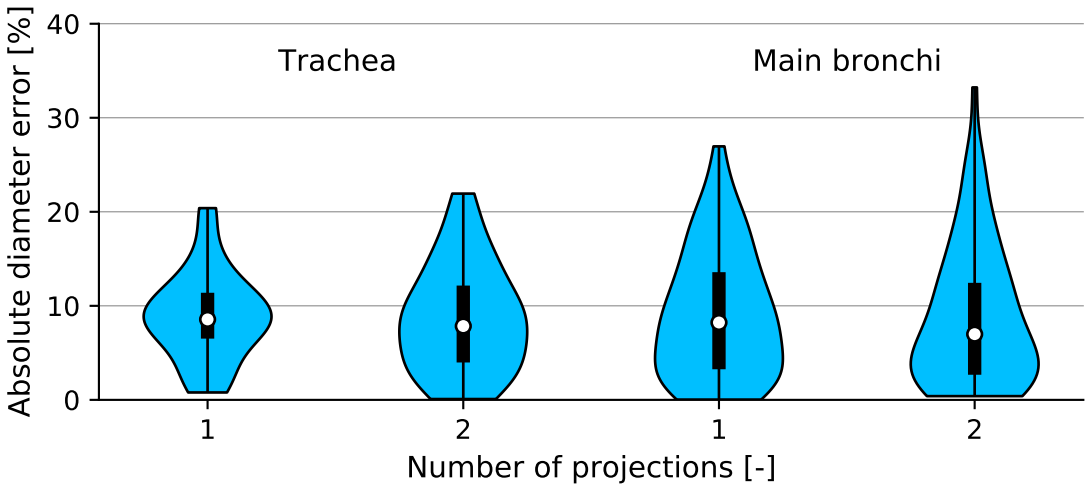


**Fig 6.** Comparison of airway diameter errors predicted by our SSAM compared to the ground truth segmentations from the LUNA16 dataset. We show the influence of including an additional projection on diameter error in the trachea and main bronchi.

of 8.5% and 7.8% for one and two projections, respectively. The main bronchi maximum error was 26.9% and 33.2% for one and two projections, and the upper quartile error was 13.6% and 12.4%. The concordance correlation coefficient was slightly higher for two projections than one (0.71 compared to 0.657, Figure S7). There was no statistically significant difference between the absolute error when comparing SSAM results with one and two projections (Wilcoxon signed-rank test $p = 0.6$ and $p = 0.69$ for the trachea and main bronchi, respectively). Similar to the lung volume error discussed above, the correlation coefficient of SSAM airway diameter predictions in the Southampton/Air Liquide dataset was lower than the LUNA16 dataset (0.71 compared to 0.53, Table 1). The CNN-based segmentations yielded similar performance on both datasets, as the median error was 5.5% and 6.5% on the LUNA16 (U-Net) and Southampton/Air Liquide (ENet) datasets, respectively. Correlation coefficients of airway diameter on both datasets were above 0.9 (Table 1), despite using an ENet on the Southampton/Air Liquide dataset which has fewer trainable parameters.

To evaluate the effect of imaging error propagation into distal airway diameters, we evaluated the diameter generated from the distal airway generation algorithm from the ground truth compared to the SSAM and CNN (Fig 7). Here, the ground truth refers to results from the airway generation algorithm

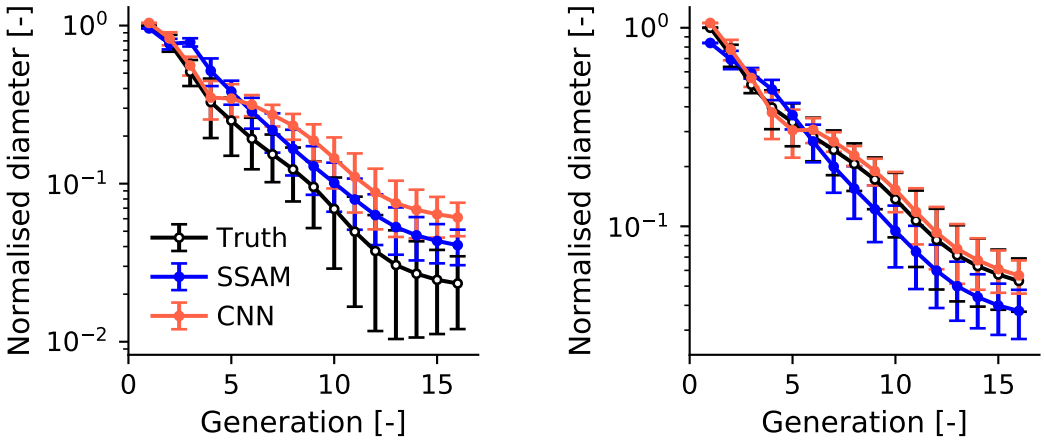


**Fig 7.** Generated airway diameter for our image-processing modalities (SSAM and U-Net CNN) on two cases from the LUNA16 dataset. Markers show the normalised mean diameter per generation. Error bars show $\pm 1 S.D.$ where $S.D.$ is the standard deviation.

using the central airway tree from the ground truth segmentations. CNN performance varied due to the dependence on the diameter in the central airways. The median absolute error in mean airway diameter per generation was 24.2% (similar to the SSAM with parent model, Figure S8). Additionally, the maximum absolute error was 77.6%, compared to 76% for the SSAM. The CNN airways show improved predictions of the diameter standard deviation (Fig 7), as the maximum error in standard deviation was 125.5% for the SSAM and 69.6% for the CNN.

### 3.2 Deposition analysis

In this section, we first compare deposition predictions from CPFD simulations in the LUNA16 dataset. We compare predicted deposition in ground truth airways with deposition predictions in airways obtained from SSAM and CNN segmentations. Following this, we then compare our simulated deposition from airways segmented by the CNN with experimental results from the Southampton/Air Liquide dataset.

Errors in deposition fraction from simulations of flow in airways from the LUNA16 dataset are shown in Fig 8. Comparing simulations in the CNN with the ground truth, the median absolute error was 5.3%, with a 95% confidence interval of $-3.3 \pm 12.1\%$ (Fig 8a). The Pearson correlation coefficient between ground truth and CNN deposition results was 0.88. Deposition in the SSAM airway had a similar median absolute error as the CNN deposition (8%, Fig 8b). SSAM deposition error was more widely distributed, as the 95% confidence interval was $-3.5 \pm 20.5\%$. The Pearson correlation coefficient between deposition in the SSAM with the ground truth was 0.6. The error range for deposition in the SSAM was similar to diameter error (Fig 6) which had a confidence interval of $-2.1 \pm 21.7\%$.

To evaluate the effect of SSAM reconstruction on local deposition hotspots, we computed the deposition enhancement factor (DEF). The deposition hotspots are highly sensitive to local flow and geometry [7, 105], which itself is susceptible to geometric differences by image acquisition and image processing [106]. We can see that even for the case with the lowest deposition error (Fig 9), the deposition hotspots in the SSAM do not match the ground truth. Particularly, in the ground truth and CNN segmented airways (Fig 9a,c), the drug is noticably more dispersed than in the SSAM (Fig 9b). Less dispersed hotspots in the SSAM is likely due to additional turbulence generated in the CT-based airways which is not present in the SSAM, since it contains a reduced number of principal components. A lower number of principal components essentially simplifies the geometry and removes intricate details that may onset turbulence such as cartiligenous rings in the trachea [107, 108].

Deposition hotspots in the CNN-based airways appear qualitatively similar to the ground truth (Fig 9a,c). The surface area with DEF $> 1$ at the first bifurcation in the ground truth is 17% larger than the CNN. However, the intensity of the hotspots is significantly different since the average DEF at the first bifurcation was 2.49 for the ground truth (maximum value 400), compared to 0.93 for the

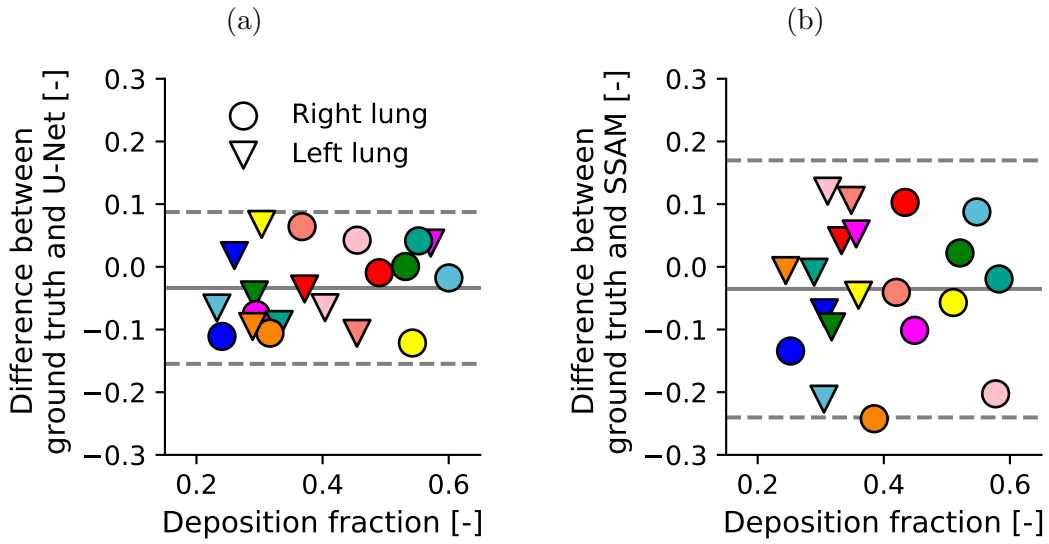


**Fig 8.** Bland-Altman plot comparing drug deposition in each lung predicted in ground truth airways, compared to airways reconstructed by (a) the U-Net, and (b) the SSAM. Number of patients tested was 10 samples from the LUNA16 dataset. Each marker represents one lung, and each colour represents a different patient.

CNN (maximum value 225). At the first bifurcation, the mean DEF in our SSAM was 1.06 (maximum value 225). The surface area with DEF $> 1$ at the first bifurcation was 2.16x larger in the ground truth than the SSAM. Interestingly, the mean DEF at the first bifurcation in the SSAM and U-Net are similar (1.06 compared to 0.93 for SSAM and U-Net, respectively). However, there is a significant difference in the surface area covered by the deposition hotspots here (Fig 9).

Finally, we performed simulations of nebulised aerosol inhalation in healthy and asthmatic patient cases of the Southampton/Air Liquide dataset [59, 60]. The maximum deposition absolute error averaged over both lungs was 12.1%, which belonged to the case with most obstructed upper airway images shown in Figure S1a (case H02 with small particles). The median, 75th percentile and maximum absolute error was 4.9%, 7.9% 13.1%, respectively. The concordance correlation coefficient between simulated and experimental lung deposition measurements was 0.43. The mouth-throat deposition had a concordance correlation coefficient of 0.525 (Fig 10b). The maximum mouth-throat deposition error was also in case H02 with small particles (27.2%). The linear best-fit line shows that the cases with high error in the mouth-throat region causes the distribution to depart from the 45 ° line (Fig 10a,b). We observed that mouth-throat deposition was best in the two cases where the glottis did not require manual cleanup (grey and yellow markers in Fig 10b). The error in these cases was 1.3% and 3.1%. The median and 75th percentile mouth-throat deposition errors were 5.7% and 11%, respectively. These results show that our models agree well with *in vivo* deposition data, particularly in the absence of imaging artefacts (Fig 10).

In an attempt to understand how our simulations would perform in a perfect case with no imaging difficulties in the upper airways, we partially corrected the simulation and experimental data by normalising the deposition fraction by $1 - DF_{mouth}$, where $DF_{mouth}$ is the deposition fraction in the mouth (Fig 10c). This makes deposition relative to number of particles entering the trachea, rather than the total number of particles entering the mouth. We apply this only to cases where the experimental lung deposition fraction was 5% larger than the simulated lung deposition (only outlier cases above the 45° line), as these were cases where the mouth-throat deposition was severly over-estimated (for example, green markers in Fig 10). Once these problematic cases were corrected, we see an improvement in the correlation between experimental and simulated data, as the concordance correlation coefficient increases from 0.432 to 0.810 (Fig 10a,c). This can be seen quantitatively by the fitted linear regression line nearly matching the 45° line. The resultant $r^2$ value is 3.3 times larger, at a value of $r^2 = 0.699$. The median, 75th percentile and maximum errors all decreased to 2.12%, 4.14% and 8.6%, respectively. This idealised test shows that the physical model is working

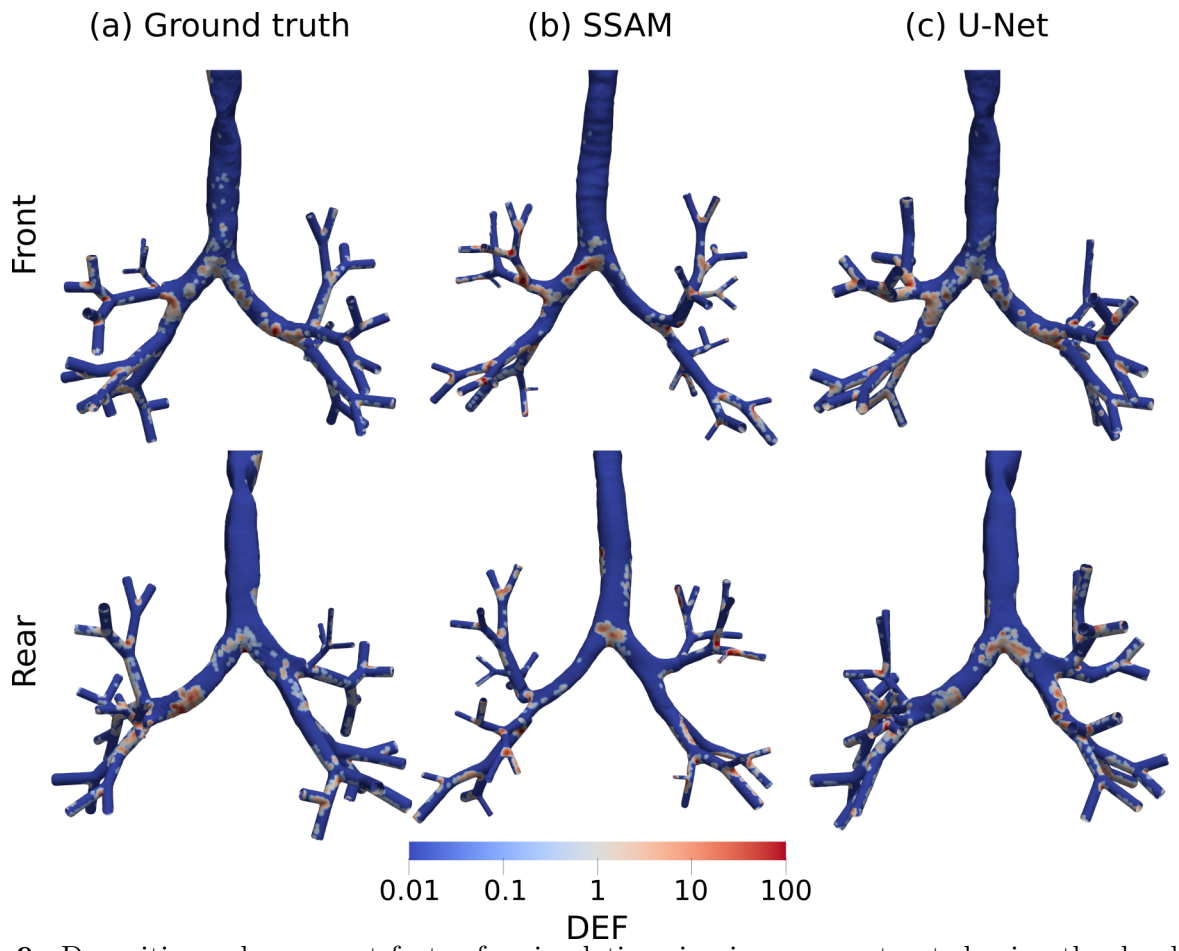


**Fig 9.** Deposition enhancement factor for simulations in airways constructed using the developed SSAM and U-Net, compared to simulations in a ground truth segmentation. This represents one sample from the LUNA16 dataset. Columns show (left) ground truth, (middle) SSAM, and (right) U-Net segmented airways. Rows show (top) front and (bottom) rear view of the airways.

well, but manually cleaning the segmentations around the glottis can introduce uncertainty that lowers correlation with experimental data (Fig 10a,c). Additionally, this idealised test points to an avenue for future development of image-processing tools for this region.

# 4 Discussion

In this study, we aimed to develop an image-processing framework to automatically generate personalised computational respiratory systems from volumetric CT and planar X-ray image(s). To that end, we developed a statistical shape and appearance model to parameterise the variation in 3D lung and airway shapes across a population. An algorithm was proposed to reconstruct 3D lung and airway geometries, by iteratively adjusting shape parameters based on a given X-ray image. We also trained a convolutional neural network (CNN) based on the U-Net and ENet architectures [83, 87], which allow us to automatically segment airways from volumetric CT scans. We evaluated the ability of our proposed framework to reproduce morphological properties and deposition results that were obtained from the ground truth airway/lung geometries. The CNN-based airways showed good agreement with the ground truth on all metrics. Our SSAM showed good agreement with the ground truth, although it was sensitive to outliers in predictions of morphological properties due to the low amount of information available in a planar X-ray image. Finally, we aimed to validate our proposed framework against experimental *in vivo* deposition measurements [59, 60]. After cleaning the upper airway segmentations to account for glottis motion artefacts, we obtained good agreement with the experimental data.

Despite imaging uncertainties in the upper airway geometry, we obtained good agreement with experimental deposition measurements (median deposition error 4.9%, Fig 10). Available *in vivo* experimental studies of drug deposition are limited due to the difficult and expensive nature of the experiments. Therefore, most modelling studies with validation compare to *in vitro* experiments, such as flow or deposition measurements in airway casts [7, 109–111]. Alternatively, some CPFD validation studies have used varying forms of experimental *in vivo* data [50, 100, 112, 113]. Tian et al. [112] compared their geometrically-simplified model to 2D scintigraphy data at varying particle sizes. De Backer et al. [100] compared patient-specific models with SPECT-CT data, but only for tracer particles ($d_p \approx 1\,\mu m$). Additionally, their outlet boundary conditions uses a fixed flowrate based on the regional lung space volume change from CT images at full inhalation and exhalation. This similar approach was also used for defining 3D domain outlet boundary conditions in a recent validation study [114]. In contrast, the outlet pressure in our study is approximated with a model based on lung space volume fraction, resistance and compliance (Section 2.5). Our approach has the benefit that it requires less radiation (only one CT or X-ray image required) to model regional ventilation to each lung. Kuprat et al. [114] determined particle deposition in the deep lung using a multiple-path particle deposition (called 'MPPD') model, which showed good agreement with experimental measurements of the steady-state dosage retained in the patients lungs after five breaths [115]. Implementing an OpenFOAM-MPPD model coupling for predictions of particle deposition in the deep lung is a potential next step to allow us to validate deposition in the deep lung. Other *in vivo* validation studies have compared to 3D deposition measurements in rats [50, 113], although the Reynolds number and Stokes number will change in these cases since the airway diameter is smaller than humans. Very recently Grill et al. [116] have performed a validation study on the same dataset as us [59]. Their model allows predictions of deposition at lobar and full conducting airway tree level, compared to ours which focuses on central airways and left/right lungs. However, we have also validated on the additional asthmatic patient dataset [60]. Moreover, our main contribution was focused on developing an image-processing approach to generate respiratory geometries for these simulations. Therefore, by validating our full patient-specific modelling and imaging approach against 3D *in vivo* data at varying particle sizes, our work represents a key development in advancing towards incorporating deposition models into clinical care.

Morphological and deposition results produced by the developed SSAM had a wide distribution, as the SSAM was sensitive to outliers. We observed that outliers occurred in cases where the outline of the lung on the X-ray image was noisy or low quality (Fig 5). However, the SSAM largely provided good agreement with the ground truth, as median absolute error was under 10% for the lung space

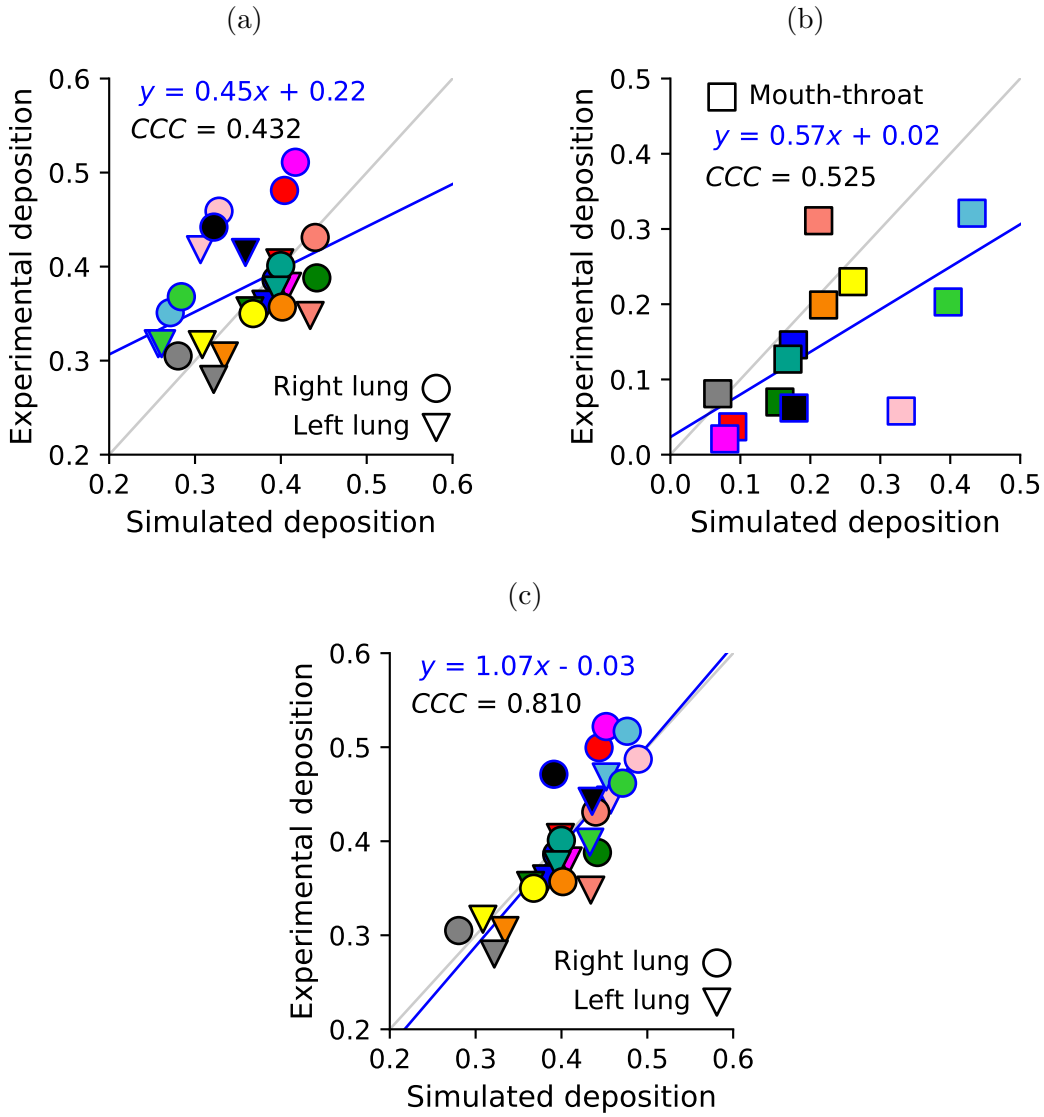


**Fig 10.** Comparison of simulated deposition against experimental measurements from the Southampton/Air Liquide dataset [59, 60]. Airways for simulations were obtained using the ENet CNN (due to larger memory requirements). Panels show (a) right and left lung deposition, (b) mouth-throat deposition, and (c) the partially corrected right and left lung deposition. Cases which are adjusted by correcting over-predicted mouth-throat deposition are shown with a blue border on the marker. *CCC* is the concordance correlation coefficient. The linear regression lines shown have (a) $r^2 = 0.213$, (b) $r^2 = 0.36$, (c) $r^2 = 0.699$. The cases requiring no manual cleanup of the glottis have grey and yellow markers.

volume (Fig 4), airway diameter (Fig 6), and lung deposition (Fig 8b). We also observed that the maximum error for airway diameter (32%) was less than the maximum error for regional deposition (24%), which shows that morphological errors in the SSAM do not propagate into significantly worse regional deposition results despite the nonlinear influence of diameter error on local velocity, Reynolds number and Stokes number on deposition.

When assessing local deposition hotspots, we observed significantly different deposition patterns in the SSAM compared to the ground truth (Fig 9a,b). Therefore, for exact predictions of deposition hotspots, the SSAM is not a suitable approach and a CNN segmentation from CT should be used. This occurs since hotspots are highly influenced by near-wall flow structures. This shows the complex geometrical features of the airways that create secondary flows responsible for deposition cannot be reconstructed from the limited information available in a planar X-ray image. Our SSAM shows similar performance to a previous study [21], which aimed to reconstruct femora from 2D radiographs, as the average errors were below 10%, but the SSAM had difficulty predicting local morphological properties. The quality of our reconstruction from 2D images may be improved by using a neural network (transformer) architecture to find the optimal shape parameters ($\boldsymbol{b}$) from the X-ray image, as this approach was shown to reconstruct a full volumetric CT from a series of X-ray images [31].

By analysing the lung space volume error outliers, we observed that the input lung edge map appeared to create some difficulties due to spuriously introduced edges (Fig 5). Li et al. [117] proposed an approach to extract lung-field outlines from chest X-rays using SSAMs. Using a SSAM instead of a general edge detection algorithm constrains the lung outline with the SSAM modes, and ensures no non-lung edges are highlighted to potentially interfere with the fitting. Alternatively, fitting of the SSAM could be improved by replacing the global loss function parameters (model coefficients, $C$ in Eq 2) with patient-specific values. This could be obtained by determining the optimal parameters for each patient in the training database, through the same Gaussian process regression described in Section 2.2.2. The input chest X-ray may then be passed through a convolutional neural network with four output neurons (one for each $C$ in Eq 2). These minor alterations may improve fitting the loss function given in Eq 2 and yield a more robust SSAM algorithm.

Error between experimental and simulated deposition in CNN-segmented airways in the LUNA16 dataset gave a 95% confidence interval of $-3.3 \pm 12.1\%$ (Fig 8a). This is likely due to errors in distal airway diameter (Fig 7) that influenced deposition in the distal airways and particles reaching the deep lung. The distal airway diameter may also have influenced CNN-segmented airway deposition through changing the outlet diameter and the computed outlet boundary condition through Eq 15. The ability of the CNN to segment distal airways was limited by the low resolution of LUNA16 datasets (mean slice thickness 1.6 mm), which is much lower than the high resolution CT used in the Southampton/Air Liquide dataset (mean slice thickness 0.5 mm) [59, 60]. We are also limited by the quality of ground truth obtained by our semi-automatic algorithm which was verified by a radiologist to have no leakage, but suffered from a limited number of branches segmented. This was due to the low image resolution, which can create significant difficulty in obtaining a full airway tree without leakage [118, 119]. The CNN training could be improved to mitigate this by masking voxels belonging to trachea and main bronchi, which forces the model to focus on learning to segment small airways [11].

To understand required improvements for future iterations of our modelling, the clinical setting the framework will be used in should be considered. In cases where a clinician is dealing with a patient in a scheduled checkup appointment, the simulation time of days achieved in this study is acceptable [16]. In time-critical emergency situations such as a patient undergoing an exacerbation, using CPFD-based modelling as a biomarker to help treatment planning is infeasible due to the turnaround time [16], and a less time-consuming model should be developed. To deliver real-time prediction of clinical quantities of interest, geometric parameters (such as SSM weights) and flow variables can be used to fit machine learning models to simulation data to provide predictions of deposition within minutes [120–122]. Morales et al. [123] showed that wall-shear stress could be predicted from patient-specific vascular meshes with minimal pre-processing using geometric deep learning, which has been developed for analysing non-Euclidean geometries such as graphs or meshes [124]. This approach could be applied to the respiratory system to evaluate properties such as deposition hotspots (Fig 9) based on a mesh obtained with our segmentation algorithm, without the additional step of fitting to a SSM [120]. Our

developed imaging and simulation framework could be used to generate training data for machine learning approaches, as well as to benefit from the speed improvements created by machine learning for CFD. Combining these modelling approaches is crucial to feasibly incorporate patient-specific models in clinical practice.

### 4.1 Limitations

A limitation of this study was the difficulty in capturing the upper airway anatomy (Supplementary material Section 1). The glottis can dilate and narrow during inhalation and exhalation, respectively [68, 69], which created difficulties in capturing CT images of the glottis [59]. Changes in the shape and area of the airways here can cause increased deposition in the glottis [125], but also downstream due to the flow separation at the expansion which creates an unsteady wake whose structure will likely vary based on patient anatomy [126–128]. Therefore, our attempt at cleaning up the glottis cannot perfectly represent instantaneous patient-specific flow patterns observed *in vivo*. Despite this, we observed good agreement with experimental lung and mouth-throat deposition measurements (Fig 10). An alternative approach to ours is a recent study on the same dataset by Grill et al. [116] who replaced the 3D mouth-throat geometry with an analytical model to predict deposition in the mouth-throat and resultant dosage entering the trachea. This also showed very good agreement with the experimental data. Future efforts should be focussed towards combining our developments with dynamic upper airway models [125, 128], to enable modelling the complex biomechanical fluid-structure interaction of respiration. However, to perform fluid-structure interaction simulations of these cases still requires a reliable baseline airway model, which requires manual cleanup of outlier cases shown in Fig 10. An improved approach to manually cleaning the glottis segmentations may be to determine the missing airway structure based on statistical shape models, as done for segmenting obstructed bronchi by Irving et al. [28]. Therefore, a template mesh could be morphed to approximate the glottis shape based on the shape of the nearby upstream and downstream airways (the throat and trachea). Combining this improved segmentation cleanup approach with dynamic mesh simulations in the upper airways is likely to produce a robust approach for modelling upper airway dynamics.

An additional limitation is the use of a healthy patient mouth-throat geometry [92] as a representative generic geometry for the upper airways in the LUNA16 dataset. Due to unavailable imaging data above the trachea in most clinical CT datasets, it was not possible to segment this from CT with our CNN, or include it as part of our SSAM to reconstruct from X-ray images. This of course limits the understanding of how well our imaging approach can capture the upper airways and flow patterns specific to a patient. The effect of this would be negligible on the CNN segmentation, as the upper airways are easily segmented with a region-growing approach [11]. It is unclear how well the SSAM could reconstruct the mouth-throat from a patient X-ray image, as there is no existing SSM that shows the main modes of variance in this part of the airways. However, literature agrees that the glottis cross-sectional area is the most significant upper airway morphological parameter influencing deposition [93, 129]. Moreover, realistic mouth-throat geometries were shown to give similar regional deposition predictions as a simplified mouth-throat geometry with an elliptical cross-section when $d_p \leq 10\,\mu\text{m}$ [129]. This suggests that it would not be essential for a SSAM to capture the exact shape of the mouth-throat, but an accurate prediction of the glottal cross sectional area is required. With a good estimation of glottal area, the SSAM could produce deposition that agrees with the deposition as shown for our generic healthy patient mouth-throat (Fig 8).

Our deep lung model was a 0D approximation of deep lung mechanics [50, 51]. Using this 0D model as an outlet boundary condition for the 3D domain (3D-0D model) has been shown to accurately predict lobar deposition in healthy and emphysematous rats [50]. In this 0D model, particles reaching the outlets are deleted and classified as depositing in the deep lung. Oakes et al. [130] developed a 1D model for fluid and particle transport in distal airways (3D-1D model), which tracks particle concentration in the distal airways. This 1D formulation showed minor improvements in agreement with experimental lobar deposition measurements over the 0D formulation [130]. A key benefit of the 1D formulation is that particles can re-enter the 3D domain during exhalation as it is known how many particles are still floating in the distal lung [130]. An inability to model particles re-entering the domain during exhalation is a limitation of our model, as we cannot compare the fraction of drug exhaled

with experimental measurements [59, 60]. In the majority of experiments, the exhaled fraction was below 10% [59, 60], meaning this has only a minor influence on the validity of our model for inhalation studies. Incorporating a 1D model for particle and flow transport in distal airways would enable further comparison to SPECT-CT data, as we could assess if our model captures how deep particles penetrate into the lung, rather than only regional deposition given here (Fig 10).

Our 0D deep lung model used global resistance and compliance metrics which were based on empirical correlations from experimental measurements [51, 131]. The resistance and compliance per outlet was then based on the outlet area and the lung's volume fraction (relative to combined volume). Oakes et al. [51] used this procedure as a initial condition, and then determined the true resistance (accounting for flow resistance in the 3D domain) by iteratively computing the pressure drop between the trachea and each outlet and adjusting the resistance until convergence. This would require many simulations for each patient to determine the exact distribution of flow and particles delivered to each individual outlet in the domain. As we are only interested in regional deposition (right and left lung) at this stage to validate our imaging and model predictions, this was not necessary. Therefore, we have not accounted for 3D domain flow resistance as it has no effect on the distribution of particles between right and left lung, which is dominated by the lobe volume fraction, $\alpha_{(L)}$, and airway morphology. In future studies where exact local deposition is required, the iterative simulation approach of Oakes et al. [51] may be improved by combining full 3D-0D simulations with a low-fidelity surrogate model to decrease the number of time-consuming simulations [78].

A final limitation is the use of DRRs to train and validate the SSAM. As the clinical end-user would use real chest X-rays as an input, here we discuss the potential issues of this and their mitigations. DRRs are commonly used in the absence of paired X-ray and CT images in 2D-3D reconstruction studies [31–33]. We observed the primary difference between DRRs and real chest X-rays is the reduced detail in DRRs due to the coarser resolution of the CT used to derive the DRR (Figure S2, Supplementary Material Section 2). This additional detail adds noise to the pixel intensity profiles (Figure S3), which could easily be removed with a Gaussian filter. Alternatively, one could use robust PCA [132] to limit the influence of noisy data on SSAM fitting. Another complication of real chest X-ray data may connected devices that could occlude the underlying lung structure [133], which could also be removed with robust PCA as shown in face recognition literature [134].

# 5 Conclusion

We have developed a rapid image-processing and mathematical modelling pipeline to enable patient-specific predictions of drug deposition with data as sparse as a single chest X-ray image. This approach segmented lungs and airways from volumetric CT data with high accuracy using convolutional neural networks [12, 83, 87]. We also developed an approach to extract lung and airway geometries from chest X-rays using a statistical shape and appearance model that iteratively adjusted shape parameters based on the outline and gray-value distribution of an unseen X-ray image. This approach was shown to reconstruct patient respiratory morphologies in good agreement with the ground truth data, with the exception of a few outliers. Deposition in airways from both automated approaches agreed well with deposition in ground truth airways obtained by semi-automatic segmentation. Finally, the imaging and modelling framework was compared to experimental *in vivo* measurements [59, 60]. The quality of images of the upper airway (mouth-throat) was mixed and required cleanup [59], which introduced uncertainty. Despite this uncertainty in the mouth-throat geometry, we achieved good agreement with experimental regional deposition. To enable future integration of physical models into healthcare settings and e-health frameworks, the ability to generate patient-specific respiratory systems and deposition predictions automatically from sparse data is crucial. Predicted deposition information could be used as part of the clinical development of new inhaled therapeutics, and reduce the need for expensive and irradiating *in vivo* deposition imaging. Our developed and experimentally validated framework is an essential step towards clinical implementation of patient-specific modelling, which will allow for automated and reliable predictions of patient therapeutic response to an aerosol drug.

## Supporting information

**S1 File Supporting information.** Supplementary methodological details and results. We detail our manual cleanup of the upper airways in the Southampton/Air Liquide dataset. We then compare qualitative differences between a DRR and real chest X-ray to understand potential challenges working with real chest X-ray images. We provide results on the two CNN architectures tested for airway segmentation, including DICE coefficient and inference time. We provide additional results on SSAM reconstruction accuracy (reported as dimensional values). We then provide results that justify the choice of distal airway diameter model for the SSAM-reconstructed airways.

## Acknowledgments

The authors thank Daniel Bustamante for his work in performing LUNA16 airway segmentations. The authors acknowledge the National Cancer Institute and the Foundation for the National Institutes of Health, and their critical role in the creation of the free publicly available LIDC/IDRI Database used in this study.

## References

1. The Global Asthma Network. The Global Asthma Report; 2018.
2. van Boven JF, Cushen B, Sulaiman I, Greene G, MacHale E, Mokoka MC, et al. Personalising adherence-enhancing interventions using a smart inhaler in patients with COPD: an exploratory cost-effectiveness analysis. NPJ primary care respiratory medicine. 2018;28(1):1–3.
3. Honkoop P, Usmani O, Bonini M. The Current and Future Role of Technology in Respiratory Care. Pulmonary Therapy. 2022; p. 1–13.
4. Lim K, Harris D, Sail P, Parry M. Robust Method to Predict Inspiratory Flowrate from the Acoustic Signature of Swirl-Based DPIs using Deep Learning. In: Journal of Aerosol Medicine and Pulmonary Drug Delivery. vol. 32. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2019. p. A27–A27.
5. Nowak N, Kakade PP, Annapragada AV. Computational fluid dynamics simulation of airflow and aerosol deposition in human lungs. Annals of biomedical engineering. 2003;31(4):374–390.
6. Koullapis P, Kassinos SC, Bivolarova MP, Melikov AK. Particle deposition in a realistic geometry of the human conducting airways: Effects of inlet velocity profile, inhalation flowrate and electrostatic charge. Journal of biomechanics. 2016;49(11):2201–2212.
7. Williams J, Kolehmainen J, Cunningham S, Ozel A, Wolfram U. Effect of patient inhalation profile and airway structure on drug deposition in image-based models with particle-particle interactions. International Journal of Pharmaceutics. 2022;612:121321.
8. Kleinstreuer C, Zhang Z. Airflow and particle transport in the human respiratory system. Annual review of fluid mechanics. 2010;42:301–334.
9. Sonka M, Park W, Hoffman EA. Rule-based detection of intrathoracic airway trees. IEEE transactions on medical imaging. 1996;15(3):314–326.
10. Tschirren J, Yavarna T, Reinhardt J. Airway segmentation framework for clinical environments. In: Proceedings of 2nd international workshop on pulmonary image analysis; 2009. p. 227–38.
11. Garcia-Uceda A, Selvan R, Saghir Z, Tiddens H, de Bruijne M. Automatic airway segmentation from Computed Tomography using robust and efficient 3-D convolutional neural networks. Scientific Reports. 2021;.

12. Hofmanninger J, Prayer F, Pan J, Röhrich S, Prosch H, Langs G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European Radiology Experimental. 2020;4(1):1–13.

13. Yun J, Park J, Yu D, Yi J, Lee M, Park HJ, et al. Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net. Medical image analysis. 2019;51:13–20.

14. Tang H, Zhang C, Xie X. Automatic Pulmonary Lobe Segmentation Using Deep Learning. arXiv preprint arXiv:190309879. 2019;.

15. LeCun Y, Bengio Y, et al. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks. 1995;3361(10):1995.

16. Sin DD, Hollander Z, DeMarco ML, McManus BM, Ng RT. Biomarker development for chronic obstructive pulmonary disease. From discovery to clinical implementation. American journal of respiratory and critical care medicine. 2015;192(10):1162–1170.

17. New JP, Bakerly ND, Leather D, Woodcock A. Obtaining real-world evidence: the Salford Lung Study. Thorax. 2014;69(12):1152–1154.

18. Woodcock A, Janson C, Rees J, Frith L, Löfdahl M, Moore A, et al. Effects of switching from a metered dose inhaler to a dry powder inhaler on climate emissions and asthma control: post-hoc analysis. Thorax. 2022;77(12):1187–1192.

19. Albrecht T, Lüthi M, Gerig T, Vetter T. Posterior shape models. Medical image analysis. 2013;17(8):959–973.

20. Baka N, Kaptein BL, de Bruijne M, van Walsum T, Giphart J, Niessen WJ, et al. 2D–3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. Medical image analysis. 2011;15(6):840–850.

21. Väänänen SP, Grassi L, Flivik G, Jurvelin JS, Isaksson H. Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. Medical image analysis. 2015;24(1):125–134.

22. Zheng G, Nolte LP, Ferguson SJ. Scaled, patient-specific 3D vertebral model reconstruction based on 2D lateral fluoroscopy. International journal of computer assisted radiology and surgery. 2011;6(3):351–366.

23. Youn K, Park MS, Lee J. Iterative approach for 3D reconstruction of the femur from uncalibrated 2D radiographic images. Medical engineering & physics. 2017;50:89–95.

24. Dworzak J, Lamecker H, von Berg J, Klinder T, Lorenz C, Kainmüller D, et al. 3D reconstruction of the human rib cage from 2D projection images using a statistical shape model. International journal of computer assisted radiology and surgery. 2010;5(2):111–124.

25. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models-their training and application. Computer vision and image understanding. 1995;61(1):38–59.

26. Cootes TF, Edwards GJ, Taylor CJ. Active appearance models. IEEE Transactions on pattern analysis and machine intelligence. 2001;23(6):681–685.

27. Heimann T, Meinzer HP. Statistical shape models for 3D medical image segmentation: a review. Medical image analysis. 2009;13(4):543–563.

28. Irving B, Goussard P, Gie R, Todd-Pokropek A, Taylor P. Segmentation of obstructed airway branches in CT using airway topology and statistical shape analysis. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2011. p. 447–451.

29. Irving B, Douglas T, Taylor P. 2D X-ray airway tree segmentation by 3D deformable model projection and registration. In: Proceedings of the Fifth International Workshop on Pulmonary Image Analysis; 2013. p. 43–52.

30. Osanlouy M, Clark AR, Kumar H, King C, Wilsher ML, Milne DG, et al. Lung and fissure shape is associated with age in healthy never-smoking adults aged 20–90 years. Scientific Reports. 2020;10(1):1–13.

31. Shen L, Zhao W, Xing L. Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning. Nature biomedical engineering. 2019;3(11):880–888.

32. Ying X, Guo H, Ma K, Wu J, Weng Z, Zheng Y. X2CT-GAN: reconstructing CT from biplanar X-rays with generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. p. 10619–10628.

33. Tan Z, Li J, Tao H, Li S, Hu Y. XctNet: Reconstruction network of volumetric images from a single X-ray image. Computerized Medical Imaging and Graphics. 2022;98:102067.

34. Bohlender S, Oksuz I, Mukhopadhyay A. A survey on shape-constraint deep learning for medical image segmentation. IEEE Reviews in Biomedical Engineering. 2021;.

35. Gozes O, Greenspan H. Lung Structures Enhancement in Chest Radiographs via CT based FCNN Training. In: Image Analysis for Moving Organ, Breast, and Thoracic Images. Springer; 2018. p. 147–158.

36. Yang W, Chen Y, Liu Y, Zhong L, Qin G, Lu Z, et al. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Medical image analysis. 2017;35:421–433.

37. Liang J, Tang YX, Tang YB, Xiao J, Summers RM. Bone suppression on chest radiographs with adversarial learning. In: Medical Imaging 2020: Computer-Aided Diagnosis. vol. 11314. SPIE; 2020. p. 45–50.

38. Bordas R, Lefevre C, Veeckmans B, Pitt-Francis J, Fetita C, Brightling CE, et al. Development and analysis of patient-based complete conducting airways models. PloS one. 2015;10(12):e0144105.

39. Tawhai MH, Hunter P, Tschirren J, Reinhardt J, McLennan G, Hoffman EA. CT-based geometry analysis and finite element models of the human and ovine bronchial tree. Journal of applied physiology. 2004;97(6):2310–2321.

40. Montesantos S, Katz I, Pichelin M, Caillibotte G. The creation and statistical evaluation of a deterministic model of the human bronchial tree from HRCT images. PLOS one. 2016;11(12):e0168026.

41. Haefeli-Bleuer B, Weibel ER. Morphometry of the human pulmonary acinus. The Anatomical Record. 1988;220(4):401–414.

42. Rausch SM, Haberthür D, Stampanoni M, Schittny J, Wall W. Local strain distribution in real three-dimensional alveolar geometries. Annals of biomedical engineering. 2011;39(11):2835–2843.

43. Dong J, Yang Y, Zhu Y. Recent advances in the understanding of alveolar flow. Biomicrofluidics. 2022;16(2):021502.

44. Sznitman J. Respiratory microflows in the pulmonary acinus. Journal of biomechanics. 2013;46(2):284–298.

45. Khajeh-Hosseini-Dalasm N, Longest PW. Deposition of particles in the alveolar airways: inhalation and breath-hold with pharmaceutical aerosols. Journal of aerosol science. 2015;79:15–30.

46. Hofemeier P, Sznitman J. The role of anisotropic expansion for pulmonary acinar aerosol deposition. Journal of biomechanics. 2016;49(14):3543–3548.

47. Koullapis P, Hofemeier P, Sznitman J, Kassinos SC. An efficient computational fluid-particle dynamics method to predict deposition in a simplified approximation of the deep lung. European Journal of Pharmaceutical Sciences. 2018;113:132–144.

48. Koullapis P, Ollson B, Kassinos SC, Sznitman J. Multiscale in silico Lung Modeling Strategies for Aerosol Inhalation Therapy and Drug Delivery. Current Opinion in Biomedical Engineering. 2019;.

49. Koullapis P, Stylianou F, Sznitman J, Olsson B, Kassinos S. Towards whole-lung simulations of aerosol deposition: A model of the deep lung. Journal of Aerosol Science. 2020;144:105541.

50. Oakes JM, Marsden AL, Grandmont C, Darquenne C, Vignon-Clementel IE. Distribution of aerosolized particles in healthy and emphysematous rat lungs: Comparison between experimental and numerical studies. Journal of biomechanics. 2015;48(6):1147–1157.

51. Oakes JM, Roth SC, Shadden SC. Airflow simulations in infant, child, and adult pulmonary conducting airways. Annals of biomedical engineering. 2018;46(3):498–512.

52. Roth CJ, Yoshihara L, Ismail M, Wall WA. Computational modelling of the respiratory system: discussion of coupled modelling approaches and two recent extensions. Computer Methods in Applied Mechanics and Engineering. 2017;314:473–493.

53. Comerford A, Förster C, Wall WA. Structured tree impedance outflow boundary conditions for 3D lung simulations. Journal of biomechanical engineering. 2010;132(8):081002.

54. Ismail M, Gravemeier V, Comerford A, Wall W. A stable approach for coupling multidimensional cardiovascular and pulmonary networks based on a novel pressure-flow rate or pressure-only Neumann boundary condition formulation. International Journal for Numerical Methods in Biomedical Engineering. 2014;30(4):447–469.

55. Pozin N, Montesantos S, Katz I, Pichelin M, Vignon-Clementel I, Grandmont C. A tree-parenchyma coupled model for lung ventilation simulation. International journal for numerical methods in biomedical engineering. 2017;33(11):e2873.

56. Kuprat A, Jalali M, Jan T, Corley R, Asgharian B, Price O, et al. Efficient bi-directional coupling of 3D computational fluid-particle dynamics and 1D Multiple Path Particle Dosimetry lung models for multiscale modeling of aerosol dosimetry. Journal of Aerosol Science. 2020; p. 105647.

57. Poorbahrami K, Oakes JM. Regional flow and deposition variability in adult female lungs: A numerical simulation pilot study. Clinical Biomechanics. 2019;66:40–49.

58. Feng Y, Zhao J, Hayati H, Sperry T, Yi H. Tutorial: Understanding the transport, deposition, and translocation of particles in human respiratory systems using Computational Fluid-Particle Dynamics and Physiologically Based Toxicokinetic models. Journal of Aerosol Science. 2021;151:105672.

59. Conway J, Fleming J, Majoral C, Katz I, Perchet D, Peebles C, et al. Controlled, parametric, individualized, 2-D and 3-D imaging measurements of aerosol deposition in the respiratory tract of healthy human subjects for model validation. Journal of aerosol science. 2012;52:1–17.

60. Fleming J, Conway J, Majoral C, Katz I, Caillibotte G, Pichelin M, et al. Controlled, parametric, individualized, 2-D and 3-D imaging measurements of aerosol deposition in the respiratory tract of asthmatic human subjects for model validation. Journal of aerosol medicine and pulmonary drug delivery. 2015;28(6):432–451.

61. Setio AAA, Traverso A, De Bel T, Berens MS, Van Den Bogaard C, Cerello P, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Medical image analysis. 2017;42:1–13.

62. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging. 2013;26(6):1045–1057.

63. Nardelli P, Khan KA, Corvò A, Moore N, Murphy MJ, Twomey M, et al. Optimizing parameters of an open-source airway segmentation algorithm using different CT images. Biomedical engineering online. 2015;14(1):62.

64. Munro S, Vos W, Surujbally R, Mignot B. Using a combination of In-Silico Lung Deposition Modelling and Secondary Endpoint Data taken from a Phase 3 Clinical Trial in patients with Severe Asthma, to provide evidence that Lung Deposition may be Enhanced when using a Smart Nebuliser instead of a Conventional Nebuliser. In: Journal of Aerosol Medicine and Pulmonary Drug Delivery. vol. 33. Mary Ann Liebert; 2020. p. A22–A22.

65. Fleming J, Conway J, Majoral C, Tossici-Bolt L, Katz I, Caillibotte G, et al. The use of combined single photon emission computed tomography and X-ray computed tomography to assess the fate of inhaled aerosol. Journal of aerosol medicine and pulmonary drug delivery. 2011;24(1):49–60.

66. Montesantos S, Katz I, Fleming J, Majoral C, Pichelin M, Dubau C, et al. Airway morphology from high resolution computed tomography in healthy subjects and patients with moderate persistent asthma. The anatomical record. 2013;296(6):852–866.

67. Ibanez J, Raurich J. Normal values of functional residual capacity in the sitting and supine positions. Intensive care medicine. 1982;8(4):173–177.

68. Brancatisano T, Collett P, Engel L. Respiratory movements of the vocal cords. Journal of Applied Physiology. 1983;54(5):1269–1276.

69. Scheinherr A, Bailly L, Boiron O, Lagier A, Legou T, Pichelin M, et al. Realistic glottal motion and airflow rate during human breathing. Medical engineering & physics. 2015;37(9):829–839.

70. Ferrarini L, Olofsen H, Palm WM, Van Buchem MA, Reiber JH, Admiraal-Behloul F. GAMEs: growing and adaptive meshes for fully automatic shape modeling and analysis. Medical image analysis. 2007;11(3):302–314.

71. Lee TC, Kashyap RL, Chu CN. Building skeleton models via 3-D medial surface axis thinning algorithms. CVGIP: Graphical Models and Image Processing. 1994;56(6):462–478.

72. Canny J. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence. 1986;(6):679–698.

73. Williams J, Ozel A, Wolfram U. pyssam – a Python library for statistical modelling of biomedical shape and appearance. arXiv preprint arXiv:230104416. 2023;.

74. Sarkalkan N, Weinans H, Zadpoor AA. Statistical shape and appearance models of bones. Bone. 2014;60:129–140.

75. Castro-Mateos I, Pozo JM, Cootes TF, Wilkinson JM, Eastell R, Frangi AF. Statistical shape and appearance models in osteoporosis. Current osteoporosis reports. 2014;12(2):163–173.

76. Meunier L, Rakotoarison H, Wong PK, Roziere B, Rapin J, Teytaud O, et al. Black-Box Optimization Revisited: Improving Algorithm Selection Wizards through Massive Benchmarking. IEEE Transactions on Evolutionary Computation. 2021;.

77. Rapin J, Teytaud O. Nevergrad - A gradient-free optimization platform; 2018. `https://GitHub.com/FacebookResearch/Nevergrad`.

78. Perdikaris P, Karniadakis GE. Model inversion via multi-fidelity Bayesian optimization: a new paradigm for parameter estimation in haemodynamics, and beyond. Journal of The Royal Society Interface. 2016;13(118):20151107.

79. Angelikopoulos P, Papadimitriou C, Koumoutsakos P. Bayesian uncertainty quantification and propagation in molecular dynamics simulations: a high performance computing framework. The Journal of chemical physics. 2012;137(14):144103.

80. Krige DG. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Southern African Institute of Mining and Metallurgy. 1951;52(6):119–139.

81. Grassi L, Hraiech N, Schileo E, Ansaloni M, Rochette M, Viceconti M. Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur. Medical engineering & physics. 2011;33(1):112–120.

82. Carr JC, Fright WR, Beatson RK. Surface interpolation with radial basis functions for medical imaging. IEEE transactions on medical imaging. 1997;16(1):96–107.

83. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer; 2015. p. 234–241.

84. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention. Springer; 2016. p. 424–432.

85. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, Garnett R, editors. Advances in Neural Information Processing Systems 32. Curran Associates, Inc.; 2019. p. 8024–8035.

86. Juarez AGU, Selvan R, Saghir Z, de Bruijne M. A joint 3D UNet-Graph Neural Network-based method for Airway Segmentation from chest CTs. In: International Workshop on Machine Learning in Medical Imaging. Springer; 2019. p. 583–591.

87. Paszke A, Chaurasia A, Kim S, Culurciello E. Enet: A deep neural network architecture for real-time semantic segmentation. arXiv preprint arXiv:160602147. 2016;.

88. Comelli A, Dahiya N, Stefano A, Benfante V, Gentile G, Agnese V, et al. Deep learning approach for the segmentation of aneurysmal ascending aorta. Biomedical Engineering Letters. 2021;11(1):15–24.

89. Lieman-Sifry J, Le M, Lau F, Sall S, Golden D. FastVentricle: cardiac segmentation with ENet. In: International Conference on Functional Imaging and Modeling of the Heart. Springer; 2017. p. 127–138.

90. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2980–2988.

91. Nousias S, Zacharaki EI, Moustakas K. AVATREE: An open-source computational modelling framework modelling Anatomically Valid Airway TREE conformations. PloS one. 2020;15(4):e0230259.

92. Banko A, Coletti F, Schiavazzi D, Elkins C, Eaton J. Three-dimensional inspiratory flow in the upper and central human airways. Experiments in Fluids. 2015;56(6):117.

93. Feng Y, Zhao J, Kleinstreuer C, Wang Q, Wang J, Wu DH, et al. An in silico inter-subject variability study of extra-thoracic morphology effects on inhaled particle transport and deposition. Journal of Aerosol Science. 2018;123:185–207.

94. Weller HG, Tabor G, Jasak H, Fureby C. A tensorial approach to computational continuum mechanics using object-oriented techniques. Computers in physics. 1998;12(6):620–631.

95. Williams J, Ozel A, Wolfram U. jvwilliams23/deepLungFoam: v0.1.0; 2022. Available from: `https://doi.org/10.5281/zenodo.7448214`.

96. Williams J, Wolfram U, Ozel A. Neural stochastic differential equations for particle dispersion in large-eddy simulations of homogeneous isotropic turbulence. Physics of Fluids. 2022;34(11):113315.

97. Liu X, Doub WH, Guo C. Evaluation of metered dose inhaler spray velocities using Phase Doppler Anemometry (PDA). International journal of pharmaceutics. 2012;423(2):235–239.

98. Esmaily Moghadam M, Bazilevs Y, Hsia TY, Vignon-Clementel IE, Marsden AL. A comparison of outlet boundary treatments for prevention of backflow divergence with relevance to blood flow simulations. Computational Mechanics. 2011;48(3):277–291.

99. Esmaily Moghadam M, Vignon-Clementel IE, Figliola R, Marsden AL, of Congenital Hearts Alliance (MOCHA) Investigators M, et al. A modular numerical method for implicit 0D/3D coupling in cardiovascular finite element simulations. Journal of Computational Physics. 2013;244:63–79.

100. De Backer JW, Vos WG, Vinchurkar SC, Claes R, Drollmann A, Wulfrank D, et al. Validation of computational fluid dynamics in CT-based airway models with SPECT/CT. Radiology. 2010;257(3):854–862.

101. Balashazy I, Hofmann W, Heistracher T. Computation of local enhancement factors for the quantification of particle deposition patterns in airway bifurcations. Journal of Aerosol Science. 1999;30(2):185–203.

102. Longest PW, Vinchurkar S, Martonen T. Transport and deposition of respiratory aerosols in models of childhood asthma. Journal of Aerosol Science. 2006;37(10):1234–1257.

103. Dong J, Tian L, Ahmadi G. Numerical assessment of respiratory airway exposure risks to diesel exhaust particles. Experimental and Computational Multiphase Flow. 2019;1(1):51–59.

104. Dong J, Shang Y, Tian L, Inthavong K, Qiu D, Tu J. Ultrafine particle deposition in a realistic human airway at multiple inhalation scenarios. International journal for numerical methods in biomedical engineering. 2019; p. e3215.

105. Farghadan A, Poorbahrami K, Jalal S, Oakes JM, Coletti F, Arzani A. Particle transport and deposition correlation with near-wall flow characteristic under inspiratory airflow in lung airways. Computers in biology and medicine. 2020;120:103703.

106. MacDonald C, Hellmuth R, Priba L, Murphy E, Gandy S, Matthew S, et al. Experimental Assessment of Two Non-Contrast MRI Sequences Used for Computational Fluid Dynamics: Investigation of Consistency Between Techniques. Cardiovascular Engineering and Technology. 2020;11(4):416–430.

107. Zhang Y, Finlay WH. Measurement of the effect of cartilaginous rings on particle deposition in a proximal lung bifurcation model. Aerosol Science and Technology. 2005;39(5):394–399.

108. Russo J, Robinson R, Oldham MJ. Effects of cartilage rings on airflow and particle deposition in the trachea and main bronchi. Medical engineering & physics. 2008;30(5):581–589.

109. Koullapis P, Kassinos SC, Muela J, Perez-Segarra C, Rigola J, Lehmkuhl O, et al. Regional aerosol deposition in the human airways: The SimInhale benchmark case and a critical assessment of in silico methods. European Journal of Pharmaceutical Sciences. 2018;113:77–94.

110. Holbrook LT, Longest PW. Validating CFD predictions of highly localized aerosol deposition in airway models: In vitro data and effects of surface properties. Journal of Aerosol Science. 2013;59:6–21.

111. Longest PW, Tian G, Walenga RL, Hindle M. Comparing MDI and DPI aerosol deposition using in vitro experiments and a new stochastic individual path (SIP) model of the conducting airways. Pharmaceutical research. 2012;29(6):1670–1688.

112. Tian G, Hindle M, Lee S, Longest PW. Validating CFD predictions of pharmaceutical aerosol deposition with in vivo data. Pharmaceutical research. 2015;32(10):3170–3187.

113. Oakes JM, Scadeng M, Breen EC, Prisk GK, Darquenne C. Regional distribution of aerosol deposition in rat lungs using magnetic resonance imaging. Annals of biomedical engineering. 2013;41(5):967–978.

114. Kuprat A, Price O, Asgharian B, Singh R, Colby S, Yugulis K, et al. Automated bidirectional coupling of multiscale models of aerosol dosimetry: Validation with subject-specific deposition data. Journal of Aerosol Science. 2023;174:106233.

115. Darquenne C, Lamm WJ, Fine JM, Corley RA, Glenny RW. Total and regional deposition of inhaled aerosols in supine healthy subjects and subjects with mild-to-moderate COPD. Journal of aerosol science. 2016;99:27–39.

116. Grill MJ, Biehler J, Wichmann KR, Rudlstorfer D, Rixner M, Brei M, et al. In silico high-resolution whole lung model to predict the locally delivered dose of inhaled drugs. arXiv preprint arXiv:230704757. 2023;.

117. Li X, Luo S, Hu Q, Li J, Wang D, Chiong F. Automatic lung field segmentation in X-ray radiographs using statistical shape and appearance models. Journal of Medical Imaging and Health Informatics. 2016;6(2):338–348.

118. Tschirren J, Hoffman EA, McLennan G, Sonka M. Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. IEEE transactions on medical imaging. 2005;24(12):1529–1539.

119. Lo P, Van Ginneken B, Reinhardt JM, Yavarna T, De Jong PA, Irving B, et al. Extraction of airways from CT (EXACT'09). IEEE Transactions on Medical Imaging. 2012;31(11):2093–2107.

120. Hoeijmakers M, Waechter-Stehle I, Weese J, Van de Vosse F. Combining statistical shape modeling, CFD, and meta-modeling to approximate the patient-specific pressure-drop across the aortic valve in real-time. International journal for numerical methods in biomedical engineering. 2020;36(10):e3387.

121. Hoeijmakers M, Huberts W, Rutten M, van de Vosse F. The impact of shape uncertainty on aortic-valve pressure-drop computations. International Journal for Numerical Methods in Biomedical Engineering. 2021;37(10):e3518.

122. Su B, Zhang JM, Zou H, Ghista D, Le TT, Chin C. Generating wall shear stress for coronary artery in real-time using neural networks: Feasibility and initial results based on idealized models. Computers in biology and medicine. 2020;126:104038.

123. Morales Ferez X, Mill J, Juhl KA, Acebes C, Iriart X, Legghe B, et al. Deep learning framework for real-time estimation of in-silico thrombotic risk indices in the left atrial appendage. Frontiers in Physiology. 2021;12:694945.

124. Bronstein MM, Bruna J, LeCun Y, Szlam A, Vandergheynst P. Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine. 2017;34(4):18–42.

125. Zhao J, Feng Y, Fromen CA. Glottis motion effects on the particle transport and deposition in a subject-specific mouth-to-trachea model: A CFPD study. Computers in Biology and Medicine. 2020;116:103532.

126. Luo X, Pedley T. A numerical simulation of unsteady flow in a two-dimensional collapsible channel. Journal of Fluid Mechanics. 1996;314:191–225.

127. Cisonni J, Van Hirtum A, Luo XY, Pelorson X. Experimental validation of quasi-one-dimensional and two-dimensional steady glottal flow models. Medical & biological engineering & computing. 2010;48(9):903–910.

128. Bhardwaj S, Koullapis P, Kassinos SC, Sznitman J. Fate of inhaled aerosols under the influence of glottal motion in a realistic in silico human tracheobronchial tree model. European Journal of Pharmaceutical Sciences. 2022;173:106172.

129. Xi J, Yuan JE, Yang M, Si X, Zhou Y, Cheng YS. Parametric study on mouth–throat geometrical factors on deposition of orally inhaled aerosols. Journal of Aerosol Science. 2016;99:94–106.

130. Oakes JM, Shadden SC, Grandmont C, Vignon-Clementel IE. Aerosol transport throughout inspiration and expiration in the pulmonary airways. International journal for numerical methods in biomedical engineering. 2017;33(9):e2847.

131. D'angelo E, Calderini E, Torri G, Robatto F, Bono D, Milic-Emili J. Respiratory mechanics in anesthetized paralyzed humans: effects of flow, volume, and time. Journal of Applied Physiology. 1989;67(6):2556–2564.

132. De la Torre F, Black MJ. Robust principal component analysis for computer vision. In: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001. vol. 1. IEEE; 2001. p. 362–369.

133. Gambato M, Scotti N, Borsari G, Zambon Bertoja J, Gabrieli JD, De Cassai A, et al. Chest X-ray Interpretation: Detecting Devices and Device-Related Complications. Diagnostics. 2023;13(4):599.

134. Luan X, Fang B, Liu L, Yang W, Qian J. Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. Pattern Recognition. 2014;47(2):495–508.