

**Functional Lung Imaging in Humans Using Positron Emission
Tomography**

by

Dominick Layfield

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003

© Dominick Layfield, 2003. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper
and electronic copies of this thesis document in whole or in part.

Author

.....
Department of Mechanical Engineering
June 5th, 2003

Certified by

.....
José-Gabriel Venegas
Associate Professor, Harvard Medical School
Thesis Supervisor

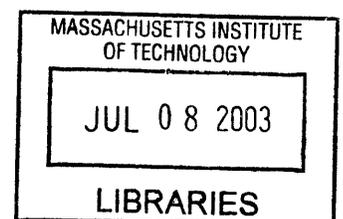
Certified by ..

.....
Derek Rowell
Professor, Department of Mechanical Engineering
Chairman, Thesis Committee

Accepted by

.....
Ain A. Sonin
Chairman, Committee for Graduate Students

BARKER



Functional Lung Imaging in Humans Using Positron Emission

Tomography

by

Dominick Layfield

Submitted to the Department of Mechanical Engineering
on June 5th, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis deals with a method of functional lung imaging using Positron Emission Tomography (PET). In this technique, a radioactive tracer, nitrogen-13, is dissolved in saline solution, and injected into a peripheral vein. By analysis of the tracer kinetics through the lung, measured using PET, a three-dimensional image of perfusion and ventilation can be generated.

In the first part of this thesis, a new tracer-preparation system, suitable for use in human subjects, is described. The system is remotely operated, highly automated, and incorporates numerous redundant safeguards to protect the patient.

The second part of the thesis details a formal approach to the analysis of the experimental data. A model of the tracer in the right heart and lungs is developed, and used to estimate physiological parameters for large to medium-sized regions of diseased lung. As regions of interest are made smaller, the amount of imaging noise in PET data increases. Consequently parameter estimates become less reliable as finer resolution is used. In order to retain as much spatial information as possible, a new approach is explored, in which voxels with similar kinetics are grouped together, and parameters are estimated for the whole group; in this way, spatial resolution is conserved at the expense of parametric discretization. The viability of the approach is demonstrated by high-resolution analysis of ventilation dysfunction in asthmatic subjects.

Thesis Supervisor: José-Gabriel Venegas

Title: Associate Professor

Acknowledgments

First and foremost, I would like to thank my advisor, José Venegas, for his guidance, support, and patience.

I view this thesis as a collaborative effort. Without the hard work of everyone in our laboratory at MGH, my research could not have been performed. In particular, I should like to thank Guido Musch for pushing the hospital to approve our human studies. I have also benefited greatly from collaboration with Marcos Vidal Melo and Scott Harris. I am grateful to Steve Weise and Sandy Barrow, who manned the PET camera and reconstructed our data, and to Jack Correia, Bill Buceliewicz and the staff of the hospital cyclotron.

I should like to thank my thesis committee, Derek Rowell and Roger Kamm, for their forbearance while my thesis painfully trickled out of me, always behind schedule.

Finally, I would like to thank Peter Madden, whose hospitality I have exploited shamefully; Henrietta, whose love carried me through my first four years at MIT; and Jill, who motivated me to finish, and has waited patiently for me to do so.

Contents

1	Introduction	11
1.1	Overview	11
1.1.1	About This Thesis	11
1.2	Functional Lung Imaging	12
1.3	Positron Emission Tomography	12
1.3.1	Emission Scans	13
1.3.2	Transmission Scans	14
1.3.3	Resolution	14
1.3.4	PET Camera	16
1.3.5	Reconstruction Algorithms	16
1.4	Nitrogen-13	16
1.4.1	Infusion-Washout Technique	17
1.4.2	Diseased Lung States	20
1.5	Imaging Noise	20
1.5.1	Quantification	21
1.5.2	Signal to Noise	26
1.6	Originality and Authorship	26
2	Tracer Preparation System: Overview	28
2.1	Introduction	28
2.1.1	Purpose of Device	28
2.1.2	History	28
2.1.3	New System	30

<i>CONTENTS</i>	6
2.2 Principle of Operation	30
2.2.1 Introduction	30
2.2.2 Overview of operation	31
2.2.3 Architecture	33
2.2.4 Execution Sequence	33
2.2.5 Design and Function of the Absorbing Chamber	38
2.2.6 Encapsulation and Transfer of the Nitrogen-13 Bubble	38
2.2.7 Dissolution of Bubble in Saline	41
2.2.8 Percussing Mechanism	42
2.2.9 Sampling	42
2.2.10 Infusion	44
3 Tracer Preparation System: Detail	45
3.1 Software Architecture	45
3.1.1 Platform	45
3.1.2 General Architecture	45
3.1.3 Injector control algorithm	46
3.1.4 Screenshots	47
3.2 Stopcock Actuators	47
3.3 Choice of Flush Gas	52
3.4 Bubble composition	54
3.4.1 Nitrogen-13 volume	54
3.4.2 Analysis of Bubble Composition	54
3.5 Solubility of gases.	55
3.5.0.1 Henry's Law	55
3.5.0.2 Partition Coefficient	55
3.5.0.3 Variation of solubility with temperature	56
3.5.0.4 Variation of solubility with dissolved salts	56
3.5.0.5 Empirical data at physiological conditions	56
3.6 Degassing	60
3.6.0.6 Mechanism	60

<i>CONTENTS</i>	7
3.7 Sodium Hydroxide	62
3.7.0.7 Change in Volume	62
3.7.1 Dissolution of gas in absorber	62
3.7.1.1 Analysis	62
3.7.1.2 Optimization of absorbing process	63
3.7.1.3 Modification of $V_g : V_l$ ratio	64
3.7.1.4 Retention of dissolved gases	64
3.8 Design of Absorbing Chamber	67
3.8.1 Compliant Chamber	67
3.8.2 Pressure-Volume Characteristic	69
3.9 Carbon Monoxide catalyst	70
3.10 Cyclotron Preparation	71
3.10.1 Duration of bombardment	71
3.10.1.1 Analysis	72
3.10.2 Target Priming	72
3.11 Infusion	73
3.12 Passive Syringe	74
3.12.1 Bubble vol/surface area/pressure calculation.	76
3.13 Performance History	77
4 Lung Model	79
4.1 Model Formulation	79
4.1.1 Bolus Infusion	79
4.1.2 Heart Model	80
4.1.3 Convective Delay	81
4.1.4 Aerated Lung	81
4.1.5 ROI-level description	82
4.1.6 Summary of parameters	84
4.1.7 Camera behavior	84
4.1.8 Combined description of heart	85
4.1.9 Combined description of aerated lung during apnea	85

<i>CONTENTS</i>	8
4.2 Shunt	85
4.3 Theoretical Results	87
4.3.1 Integrating effect of well-inflated lung	87
4.3.2 Apnea peak	87
4.3.3 End-apnea level	89
4.3.4 Apnea integral	90
4.4 Evaluation of Model Output	90
4.5 Washout	91
4.5.1 Washout Integral	92
4.6 Equivalent Ventilation	93
4.6.1 Oxygenation from \dot{V}/\dot{Q} ratios	93
4.6.2 Calculation of ‘equivalent’ ventilation	94
4.7 Parameter Identification	94
4.7.1 Identification of Cardiac Parameters	95
4.7.1.1 Double Infusion Model	97
4.7.2 Identification of Lung Parameters	99
4.7.2.1 Perfusion	99
4.7.2.2 Ventilation	101
4.8 Summary	104
5 Small ROI Parameter Estimation	105
5.1 Introduction	105
5.2 Image Noise	105
5.2.1 The ‘Pop-Up’ Effect	107
5.3 Parameter Estimation	110
5.3.1 Non-spatial grouping	110
5.3.1.1 Work of Kimura et al	111
5.4 Principal Component Analysis	111
5.4.0.2 Mathematics	111
5.4.1 Application to lung imaging	113
5.4.2 Principal components from synthetic data	115

5.4.2.1	Concept	116
5.4.2.2	PCA of synthetic data	116
5.4.3	Weighted PCA	116
5.5	Grouping	118
5.5.0.1	Orthogonal partitioning	118
5.5.0.2	Partitioning weighted by eigenvalue	120
5.5.0.3	Underpopulated groups	121
5.6	Parameter Estimation	121
5.6.1	Washout model	122
5.6.2	4-parameter fit	122
5.6.3	Initial parameter estimates	123
5.6.4	3-parameter fit	123
5.6.5	Direct amplitude estimation	123
5.6.6	'Magic Point'	124
5.6.7	Single-Compartment Fit	124
5.7	Example	126
5.7.1	Single-Compartment Groups	131
5.8	Grouping Criteria	133
5.8.1	Revised example	135
5.9	Validation	136
5.9.1	Reconstruction of global kinetics.	136
5.9.2	Quality of model fit	136
5.9.3	Analysis of artificial data	138
5.9.3.1	Generation of test data	138
5.9.3.2	Quantifying Performance	138
5.9.3.3	Optimal Number of Groups	139
5.9.3.4	Results	139
5.9.4	Perfusion Prediction	139
5.10	Summary	142

<i>CONTENTS</i>	10
6 Ventilation Disruption in Asthmatics	143
6.1 Subject Selection	143
6.2 Experimental Protocol	143
6.3 Tracer Kinetics	144
6.3.1 Constancy of Ventilation	144
6.4 Perfusion	147
6.4.1 Method	147
6.4.2 Results	147
6.5 Ventilation	150
6.5.1 Method	150
6.5.2 Results	150
6.6 Summary	152
7 Conclusion	159
A Important Experiments	161
A.1 Carbon Monoxide Content	161
A.2 Optimal NaOH concentration	162
A.3 Tracer loss to NaOH	164
A.4 Mass Spectrography of Bubble	164
A.4.1 Test Setup	165
A.4.2 Results	166
A.4.2.1 Interpretation	172
A.4.2.2 Validity	175
A.5 Target Priming	175
A.5.1 Test setup	176
A.5.1.1 Target priming apparatus	176
A.5.1.2 Measuring radioactive yield	176
A.5.2 Results	179
A.5.3 Existence of Second Species?	179
A.5.4 Data from Earlier Tests	184
Bibliography	186

Chapter 1

Introduction

1.1 Overview

Historically, a major impediment to our understanding of lung disease has been the difficulty in visualizing both the anatomy and function of the lung.

Existing techniques to image lung function all require that a tracer or contrast medium of some kind be introduced into the body, typically one tracer to measure ventilation, and another to measure perfusion. Although these methods are varied, in general they suffer from the following defects: they produce low-resolution (and often two-dimensional) images; they are poorly quantitative; and they require separate scans to measure ventilation and perfusion.

The only method currently available to simultaneously image ventilation and perfusion uses nitrogen-13 measured with Positron Emission Tomography.

Variations on this method were first used experimentally at the end of the 1960's [15], and the technique saw limited clinical use [7]. But despite the emergence of PET, which offered much better resolution and quantification than planar scintigraphy, the technique fell into disuse (last data published 1992 [4]), principally because the tracer solution was manually prepared and caused unacceptable radioexposure to the operator.

In recent years, this method has been resurrected through the work of Venegas et al, who developed apparatus for remote preparation of the tracer solution, and conducted several animal studies with this system [17, 19, 6]. However, this tracer-preparation apparatus had a number of deficiencies that made it unsuitable for use with humans.

1.1.1 About This Thesis

The goal of this PhD thesis is to develop this technique to study lung disease in humans.

This thesis is in two distinct parts. The first part [chapters 2, 3] describes the development of an apparatus for batch production and delivery of nitrogen-13 tracer solution, which is suitable for use in human studies.

The second part relates to analysis of the data obtained in such studies. A model of the the kinetics of the tracer is developed [chapter 4], and used to analyze global kinetics. Regional kinetics, however, exhibit very large amounts of imaging noise, preventing the use of conventional parameter estimation techniques to create high-resolution parametric images. In chapter 5 I describe a method of obtaining reliable parametric estimates while retaining spatial information. Finally, in chapter 6 I demonstrate the application of this technique to the analysis of experimental data from asthmatic volunteers.

1.2 Functional Lung Imaging

The principal functional parameters in the lung are perfusion, \dot{Q} , the amount of blood flowing to the lung, and ventilation, \dot{V} , the amount of air flowing to the lung. Thus the term, ‘Functional Lung Imaging’, describes methods to image \dot{V} , \dot{Q} , or the ratio, \dot{V}/\dot{Q} .

Because breathing is tidal, ventilation is usually quantified as an average over a breath-cycle. Additionally, because the airway tree has significant volume, tidal breathing convects some of this *dead-space* volume into the alveoli. By convention, \dot{V} only describes the mean rate of ‘new’ air into the alveoli.

These parameters are of fundamental importance because they determine how well the lung can oxygenate blood. In order for the lung to function optimally, these parameters should be matched: if blood flows to part of the lung which is not ventilated, that blood will not be oxygenated. If part of the lung is not perfused, then it is inefficient to circulate air to that region, as it will not participate in gas exchange. It should be obvious, however, that the former dysfunction is more serious than the latter.

The lung has a number of mechanisms to ensure that \dot{V} and \dot{Q} are matched. In regions of low ventilation, the partial pressure of oxygen in the alveolar airspace (P_{AO_2}) is decreased, which causes constriction of pulmonary arterioles to limit perfusion: a phenomenon known as *hypoxic vasoconstriction*. Similarly, in *hypocapnic pneumoconstriction* small airways constrict in response to the diminished carbon dioxide content that occurs when perfusion is restricted. Several other, slower-acting matching mechanisms are believed to exist.

In lung disease, primary derangements in perfusion (e.g. pulmonary embolism), or ventilation (e.g. asthma) are typically not life-threatening unless they are of so great a magnitude that the remaining functioning lung is unable to compensate. However, secondary derangements in the \dot{V}/\dot{Q} matching mechanisms can have dire consequences, as these result in poorly oxygenated blood.

1.3 Positron Emission Tomography

Positron Emission Tomography (PET) is a method for non-invasively imaging the spatial distribution of a radioactive tracer introduced into the body by a variety of methods.

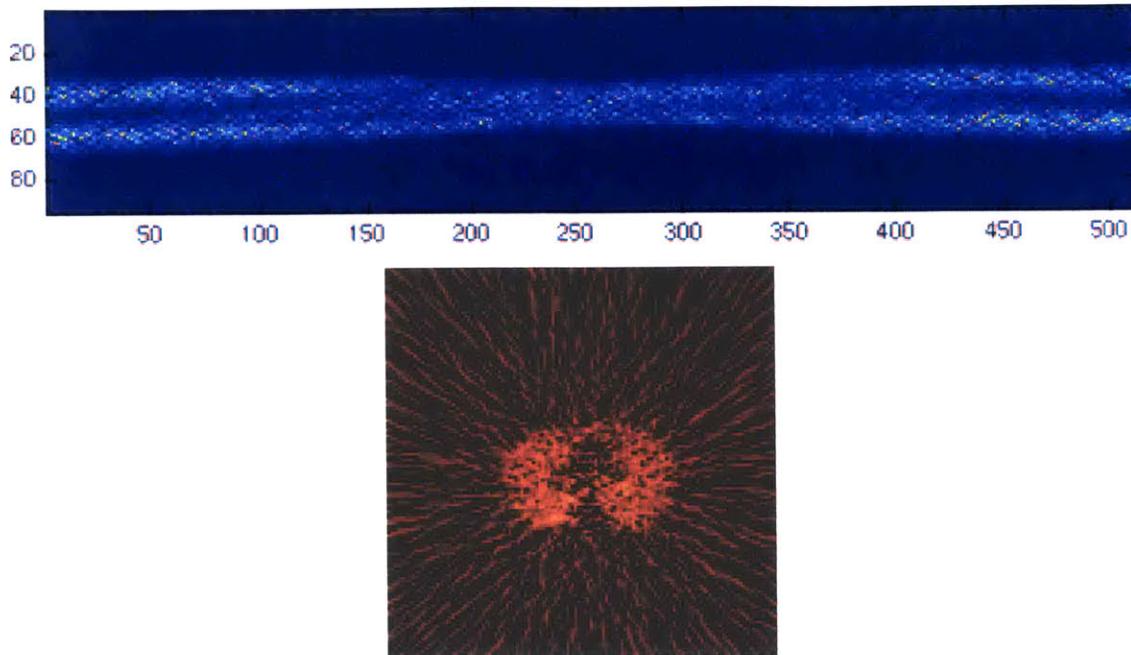


Figure 1.1: A sinogram (top), and the corresponding reconstructed image (below). The sinogram is a list of the number of coincidences between each detector pair. In this PET scanner, there are 512 detectors (numbered along the base of the sinogram) in each slice, but coincidences are only detected with the 96 detectors (numbered along the side) located opposite the first. This sinogram is taken from a 2.5 second apnea frame, and due to the short duration, the reconstructed image is fairly noisy.

1.3.1 Emission Scans

The tracer may be either an elemental form, or compound of, a radioactive isotope that decays by positron (anti-electron) emission. When such a decay occurs, the positron annihilates with a nearby electron and a pair of gamma photons are produced with equal and opposite momentum. These high-energy (511 keV) photons propagate away from the annihilation site, and are detected by a cylindrical ring of detectors surrounding the subject. The coincident arrivals of these photon pairs are recorded as a *sinogram*: a list of the number of coincidence events observed for each pair of detectors.

The sinogram can be used to reconstruct a three-dimensional image of the tracer distribution within the imaging field, using algorithms similar to those used in x-ray Computed Tomography (CT) imaging. A variety of reconstruction algorithms exist; the method used for the images shown in this thesis belongs to the class of back-projection algorithms.

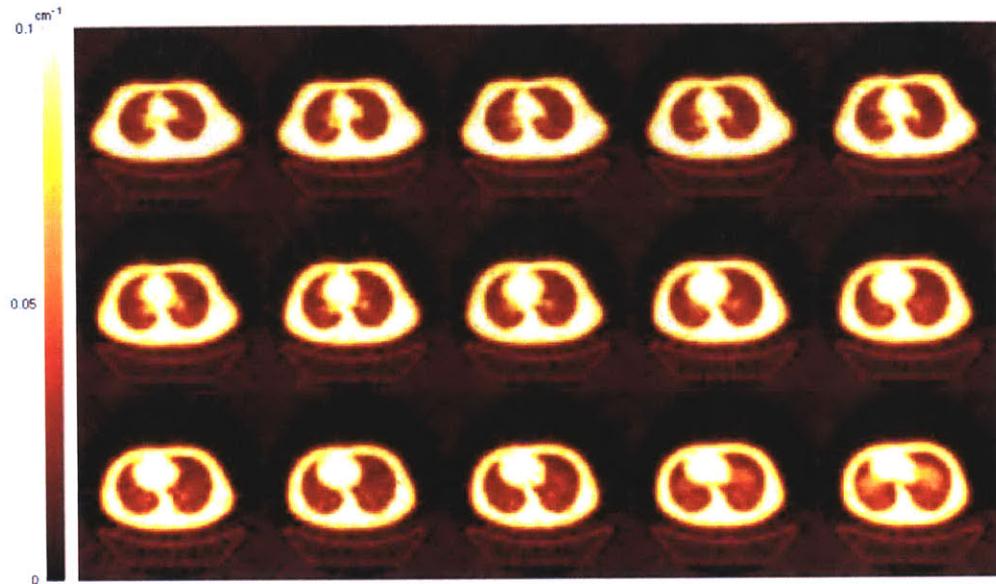


Figure 1.2: Transmission scan, in units of linear attenuation. Fifteen transverse (coronal) slices, separated by 6.5mm are shown, running in a cranial to caudal (top left to bottom right) direction. The subject, a human male, is lying supine. The lungs are clearly visible as two low density regions, separated by the mediastinum in which the outline of the heart can be seen. In the last slice, the dome of the diaphragm, just entering the slice, appears as an area of slightly higher density. Unlike CT images, which are normally shown looking foot-to-head, this image is head-to-foot, so that image left is patient left.

1.3.2 Transmission Scans

To accurately quantify tracer concentration, we need to compensate for the attenuation due to tissue density. This is achieved by making a transmission scan which accompanies the emission scan described above. In a transmission scan, a positron source is continuously rotated around the perimeter of the imaging field, while the patient is within the scanner. The resulting sinogram is subtracted from a sinogram generated from a transmission scan of an empty field, is used to reconstruct an image of tissue opacity.

These images, analogous to low-resolution CT scans, are normally used in the reconstruction of emission data, but are also useful on their own, for example to delineate the lung region, or assess tissue density within the lung.

1.3.3 Resolution

The resolution of PET images has three principal determinants. Firstly, there are fundamental limitations of the technique that no improvement in camera technology can overcome. Secondly, there

are limitations associated with the geometry of the camera itself. Thirdly, statistical imaging noise is a major problem, and we are forced to strike a compromise between resolution and noise. These determinants are discussed below.

The indirect nature of the emission is a fundamental limitation of the technique: the event that is being detected is really the positron-electron annihilation, not the decay event (positron emission) itself. The mean free path (MFP) of the positron depends inversely on the electron density of the local environment. In air, the MFP is around 1.6 meters; in water, it is around 1.5mm [11]¹. Another absolute limitation on the resolution of the method is due to the fact that dense tissue not only absorbs photons, but also deflects them (Compton scatter), so that resolution is compromised by the amount of surrounding material (in our case, a thick chest wall). Yet another fundamental limitation is that the gamma photons are not emitted in perfectly equal and opposite directions: their velocities depend on the momenta of the electron and positron that annihilate. These effects all reduce the validity of the underlying tomographic assumption, that the location of the decay event lies along the strip of space joining the pair of crystals that detected the coincidence.

The geometry of the scanner also sets absolute limits on the achievable resolution: we can never see detail finer than the size of the intersection areas of strips between detectors. These intersection areas grow larger with distance away from the ring center, so that the maximum achievable resolution deteriorates with radius. We can improve the theoretical resolution by increasing the number (or equivalently reducing the size) of detectors in the ring, but this produces diminishing returns as we approach the limitations described above.

PET imaging is vulnerable to statistical effects of the emission and detection process. The stochastic nature of radioactive decay means that signal-to-noise improves with the total number of events detected. Thus increasing tracer concentration, imaging duration, and *region-of-interest* (ROI) volume improves the reliability of our estimates. A consequence of this volumetric dependence is that we are always faced with a compromise between noise and resolution in reconstructed images. We can reconstruct at high-resolution, and see very large amounts of noise, or reconstruct at lower resolution and improve the accuracy of our estimates. Statistical imaging noise is exacerbated by attenuation of photons by surrounding tissue: in patients with a thick chest wall, the reduced number of photons that reach detectors mean that we see increased noise (and hence the practically-achievable resolution is degraded).

It is thus hard to characterize the resolution of a PET image, which depends fundamentally on the local tissue density, surrounding tissue density, and location within the imaging field. In practice, due to noise-resolution trade-offs, it also depends on local tracer concentration and on tracer concentration in surrounding tissue.

Rough Estimates

Despite the above discussion, it is useful to provide some sort of quantification to provide a rough idea of the length scales are involved.

¹Note that the MFP is not equal (and not trivially-related) to the FWHM resolution.

The data presented in this these were all acquired using a Scandtronix PC4096 camera. Based on the geometry of this camera, theoretical analysis predicts that a point source located in the center of the imaging field should produce a response with FWHM of 6mm.²

Within the lung, however, this resolution is not achieved, due principally to the long MFP (as tissue density is low), scattering by the chest wall, and by resolution-noise compromises forced by the short imaging durations required to follow tracer kinetics.

1.3.4 PET Camera

As mentioned above, the same PET camera, a Scandtronix PC4096, was used to acquire all of the data presented in this thesis.

This camera has 8 rings, spaced 13mm apart. Each ring contains 512 detectors with effective width of 6.2mm. Coincidences are detected within each ring, and between adjacent rings, so that a total of $8+7=15$ axial sinograms are generated.

1.3.5 Reconstruction Algorithms

All images and derived data presented here were reconstructed using filtered back-projection (FBP) with identical settings used for each reconstruction. This algorithm has the advantage of being well-understood and has statistical properties that are easy to model. However, FBP reconstruction introduces some undesirable artifacts, generating, for example, *negative* estimates of the tracer concentration of some voxels. Newer and better algorithms exist which iteratively converge on (penalized) maximum-likelihood estimates, and which are capable of modeling a greater range of the PET acquisition phenomena. I have experimented with some of these algorithms, but have been unable to fully validate the results in time to replace the data in this thesis.

1.4 Nitrogen-13

Nitrogen-13 (^{13}N) is a radioactive isotope of nitrogen. It has a short half-life (9.965 minutes) and decays by positron emission ($E=1.19$ MeV) to carbon-13. Since it has identical chemical properties to the stable isotope (^{14}N), it is biologically inert unless compounded.

The molecular form ($^{13}\text{N} - ^{13}\text{N}$ or $^{13}\text{N} - ^{14}\text{N}$) is weakly soluble in water, with a partition coefficient of ~ 0.014 at body temperature (about 40 times less soluble than carbon dioxide, and about half as soluble as oxygen³). Because of this low solubility, dissolved nitrogen-13 in the bloodstream redistributes almost entirely into the airspaces of the lung.

These properties make molecular nitrogen-13 an ideal agent for PET imaging. The overwhelming majority introduced into the body is exhaled rapidly, minimizing exposure of the patient to radiation.

²This number has been validated by empirical tests, which gave a FWHM of 8mm, using a pin source of approximately 2mm width.

³Obviously, the effective solubility of oxygen in whole blood is much higher due hemoglobin affinity.

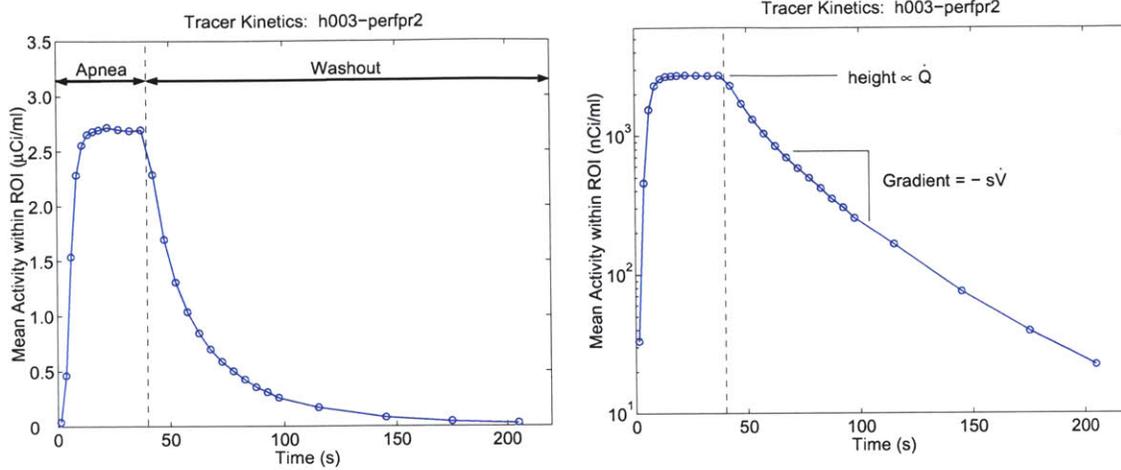


Figure 1.3: Whole-lung tracer kinetics from a healthy subject, shown on linear (L) and logarithmic (R) scales. During apnea tracer arrives in the lung and reaches a stable concentration proportional to perfusion (\dot{Q}). The washout is essentially uniform, and appears as a straight line on a logarithmic scale, with negative gradient equal to specific ventilation ($s\dot{V}$).

Furthermore, the short half-life means that any residual tracer dissolved in peripheral tissue decays rapidly, and ceases to exist within an hour or two.

Nitrogen-13 can be used in a number of ways as a tracer for pulmonary imaging. It can be inhaled directly, and used to measure ventilation (from wash-in and wash-out kinetics) and lung inflation (from steady-state data). Alternatively, it can be dissolved in saline solution and either infused continuously during normal breathing (to assess VQ ratio), or in a bolus, as described below.

1.4.1 Infusion-Washout Technique

In the infusion-washout technique, the radioisotope is dissolved in saline solution, and infused as a bolus into a peripheral vein. The tracer is convected by the venous circulation through the right heart until it reaches the lungs. Due to its poor solubility, the tracer diffuses out of solution and redistributes in the alveolar airspaces.

At the point of infusion, the patient is instructed to hold his/her breath. Without any tidal ventilation, the tracer concentration develops in the lung in proportion to local perfusion.⁴ The subject then resumes normal breathing, and the tracer is washed out of the lung at a rate equal to the local ventilation. By acquisition of a series of sequential scans, we can measure the kinetics of the tracer, and thus deduce, in a single pass, regional perfusion (\dot{Q}) and ventilation (\dot{V}).

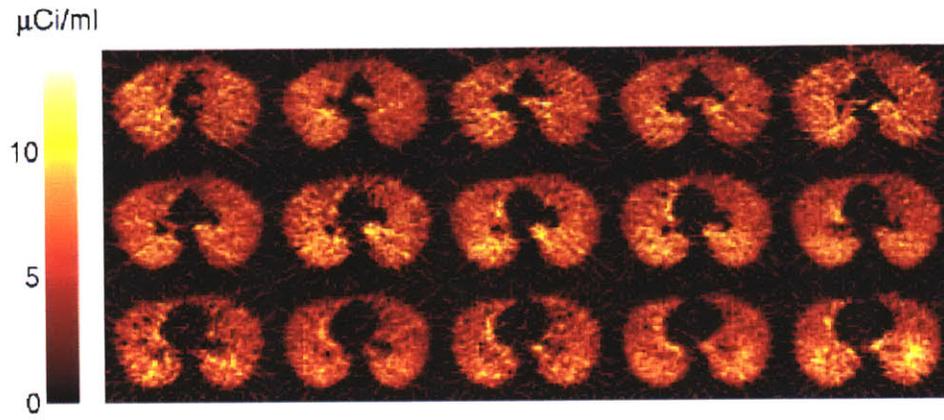


Figure 1.4: Perfusion image for normal human subject. This image is generated by taking the mean of the apnea plateau frames. The lung is not masked. This subject was slim, a large amount of activity was infused (18mCi), and the plateau was fairly long (30s): as a consequence, this example represents close to a best case of image quality obtainable using this camera and reconstruction method.

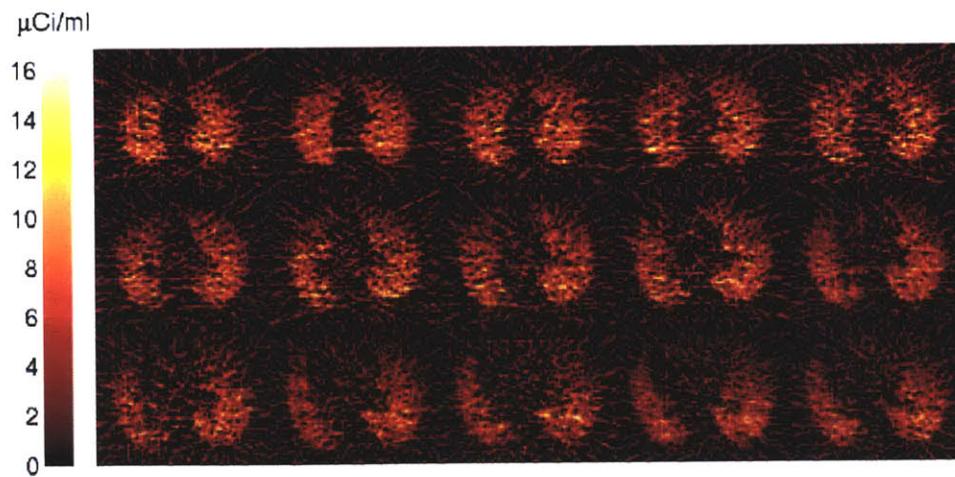


Figure 1.5: Perfusion image from another normal human subject. In this case, the subject was relatively obese, and attenuation by the chest-wall was substantial. Moreover, the plateau duration was short (10s) due to slow delivery of tracer to the lung and the subject's limited capacity to breath-hold. Again, the image is left unmasked to emphasize the amount of noise in the image.

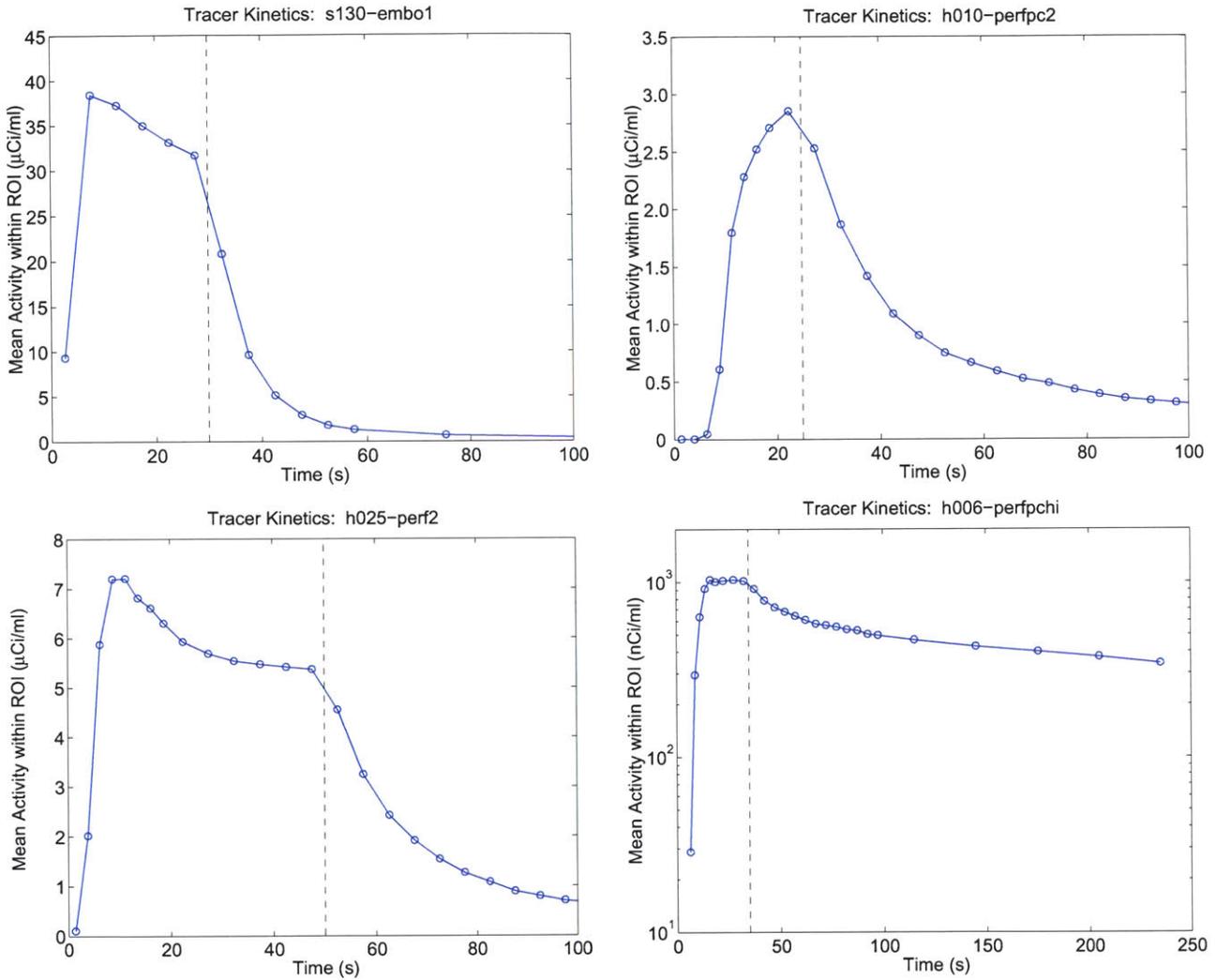


Figure 1.6: Examples of abnormal kinetics: (top left) noticeable tracer resorption in a region of very high perfusion; (top right) subject unable to breath-hold long enough for tracer equilibration; (bottom left) lung with significant shunt fraction; (bottom right) tracer retention in a bronchoconstricted lung.

1.4.2 Diseased Lung States

As described above, the core assumption of this method is that during apnea, the ^{13}N tracer accumulates in proportion to local perfusion, and that local specific ventilation is inversely proportional to the time constant of the washout.

In a healthy lung, these assumptions are reasonable. However, in our studies of diseased lung states, it has become clear that a more sophisticated analysis is required. For example:

- A bronchoconstricted human may have difficulty breath-holding long enough for complete equilibration. Tracer may still be arriving in the lung at the resumption of normal breathing.
- In lung that is collapsed or flooded, the tracer does not reach a stable plateau during apnea. The tracer does not evolve into the gas phase and is not retained in the lung, but instead is said to *shunt* – to pass straight through the lung.
- The washout of asthmatic and injured lung exhibits multi-compartmental behavior. In this case, it is impossible to calculate a single specific ventilation, $s\dot{V}$, as even small regions exhibit both fast and slow time-constants.
- In regions of gas-trapping, the tracer concentration decreases during the washout phase due to resorption of tracer back into the blood, making complete trapping difficult to distinguish from hypoventilation.
- Similarly, in regions of hyperperfusion, re-absorption of tracer during apnea can be confused with shunt or premature ventilation.

Some examples of altered kinetics from diseased lungs are shown in figure 1.6.

1.5 Imaging Noise

The noise in a PET image arises from many sources: firstly, ‘quantum’ noise intrinsic to positron emission and annihilation (variation in the rate of spontaneous decay, distance traveled by a positron before annihilation); secondly, photon attenuation and scattering (gamma photons may be completely absorbed or deflected by dense tissue); thirdly, camera-related sources (dead-time in photomultiplier tubes, finite temporal resolution of coincidence detection circuitry etc., geometry artifacts); finally artifacts are introduced by imperfect reconstruction algorithms.

Without wanting to be drawn into a complete discussion of this complex issue, some observations are useful which will enable us to make a simple approximation.

i. Noise increases with resolution. As ROI’s are made smaller and smaller, we see more and more noise in the kinetics. This trade-off is also evident during reconstruction. Achievable resolution

⁴This is an approximation, the details of which are discussed in chapter 4.

depends on the position in the imaging field (resolution is fundamentally better in the center of the imaging field where more coincidence strips intersect), on local tissue density (if a positron is emitted in an area of low density, like an airway lumen, it will travel further before annihilation), and on surrounding tissue density (which attenuates and scatters the emitted gamma photons). However, if we ignore all of these latter factors, effectively assuming a ‘perfect’ PET camera, we can state approximately that noise (measured by standard deviation) is inversely proportional to the square root of volume (ROI or voxel).

ii. Noise increases with total activity. Two important measures need to be discriminated carefully: increasing tracer concentration improves the signal to noise *ratio*, but the *absolute* amount of noise increases with tracer concentration.⁵ In general, however, we are concerned with the amount of noise in a specific region, where the ‘signal’ part of the ‘signal-to-noise’ is determined by *local* tracer content, while the ‘noise’ part is determined not only by the local photon emission, but also by emission along *all* the strips in the sinogram that intersect the ROI: i.e. noise contributions come from a large fraction of the imaging field. To accurately assess the noise in a region, we should properly transform to sinogram space (forward-project) and sum tracer content along contributing strips (direct intersections and scatter contributions). However, such a approach is too complex to be useful here. Instead we make the coarse approximation that noise is proportional to the square root of the total amount of activity within the imaging field.

iii. Noise decreases with imaging duration. As radioactive decay is Poisson-distributed, the variance in total recorded counts is proportional to the duration of acquisition, and hence the standard deviation of recorded counts is proportional to the square root of the recording duration. Thus noise in our estimate of specific activity (proportional to count *rate*) is be inversely proportional to the root of imaging duration.

Other factors are either approximately uniform throughout an imaging sequence (e.g. attenuation by dense tissue), slight (e.g. imperfections in dead-time correction) or too complex to be useful (e.g. regional variation in noise, mentioned above). With these caveats, we state

$$\sigma(\hat{a}) \propto \sqrt{\frac{A_{field}}{\tau V_{ROI}}} \quad (1.1)$$

That is, the standard deviation, σ , of our estimate of specific activity, \hat{a} , is proportional to the square root of total activity, A_{field} , over the product of imaging duration, τ , and region-of-interest volume, V_{ROI} . Note that among the factors affecting noise, the largest factor not explicitly described by this expression is the attenuation of photons by surrounding tissue. This effect is substantial, but, crucially, is constant for all the frames of an image.

1.5.1 Quantification

The frames from the apnea plateau of an image are symmetric in the sense that the ‘signal’ part of each frame is identical: only the noise changes from frame to frame. Because of this symmetry, and

⁵Radioactive decay has a Poisson distribution, where the variance is equal to the mean, so signal to noise, measured by mean/stdev, improves with increasing activity.

because the noise is unbiased as well as uncorrelated, the mean of any combination of frames has the same ‘signal’, but varying amounts of noise. We can use this property to determine the constant of proportionality in equation 1.1.

Our approximate model of noise is as follows:

$$\hat{a} = a + N\left(0, \frac{kA_{field}}{\tau V_{ROI}}\right) \quad (1.2)$$

That is, the estimated specific activity in a region, \hat{a} , is equal to the actual specific activity, a , plus normally-distributed noise with zero mean and variance as described in equation 1.1 (with k as the constant of proportionality).

In our approximate model, we assume that the noise is independent of local activity, and thus does not vary significantly across the lung. Since the noise term is constant, we can write a simple expression for the observed variance in estimated specific activity sampled at different locations in the lung:

$$\sigma^2(\hat{a}) = \sigma^2(a) + \frac{kA_{field}}{\tau V_{ROI}} \quad (1.3)$$

We can now take our plateau data, which consists of combinations of frames of different duration (τ), and consider each voxel within the lung as a region of interest. If we plot the variance in estimated specific activity in these voxels, $\sigma^2(\hat{a})$, against (A_{field}/τ) , then the data should lie on a straight line, with gradient of (k/V_{voxel}) , and with y-intercept of $\sigma^2(a)$.

In practice, it is more useful to restate this relationship by normalizing by the mean activity within the mask, which removes the dependence of the y-intercept on the amount of tracer infused.

$$\begin{aligned} \left(\frac{\sigma^2(\hat{a})}{\hat{a}}\right) &= \left(\frac{k}{V_{voxel}}\right) \left(\frac{A_{field}}{\hat{a}\tau}\right) + \left(\frac{\sigma^2(a)}{\hat{a}}\right) \\ y &= mx + c \end{aligned} \quad (1.4)$$

This relationship is plotted for a sample image in figure 1.7.

This graph provides us with much useful information. Firstly, the gradient gives us the constant of proportionality in our noise model, so that we can quantify the expected noise in a voxel or larger region of interest, for any frame of a image. This gradient is directly related to the fraction of emitted photons that are recorded by the PET camera, and is thus determined predominantly by the amount of attenuation that occurs in the chest wall. If the chest wall is very thick (for example, in an obese patient), the gradient is large.

Figure 1.8 shows a single slice of transmission scans taken from extreme examples: firstly from the patient with the smallest gradient (~ 2.2 nCi/ml²s), and secondly from the patient in which we observed the largest gradient (~ 16 nCi/ml²s). In the latter subject, attenuation dramatically increases the amount of noise in the image: to achieve the same signal-to-noise ratio as in the first subject, we would have to infuse seven times as much tracer. It is not surprising, then, that we get better quality data from thinner subjects.

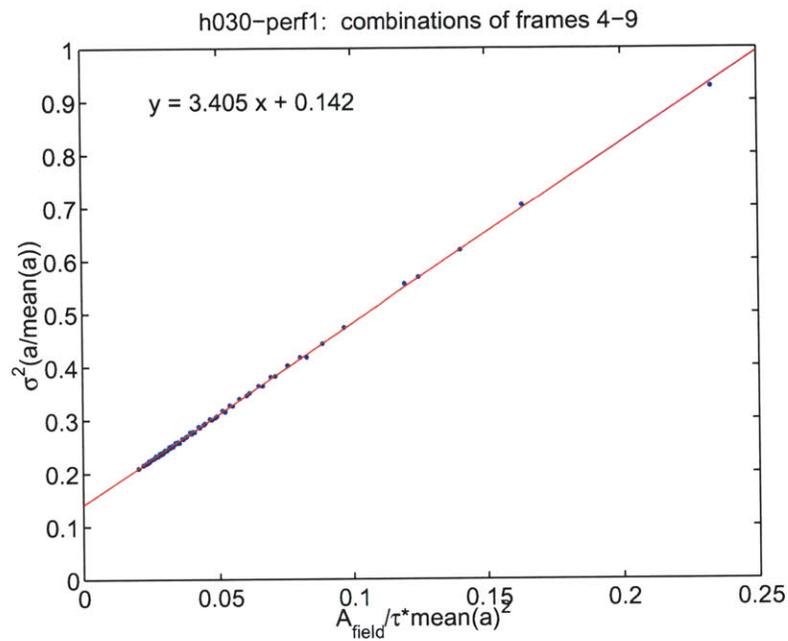


Figure 1.7: Analysis of image noise. By taking combinations of equivalent frames (from apnea plateau) and plotting normalized variance against the total activity in the imaging field, divided by frame duration*mean squared, we generate a straight line. The gradient of this line describes how image noise varies with the amount of total counts in a region of interest and is determined by invariant camera parameters and the thickness of the chest wall of the subject. The y-intercept of this line describes the variance in perfusion that we would observe if we could image for an infinite duration (or equivalently, infuse an infinite amount of tracer).

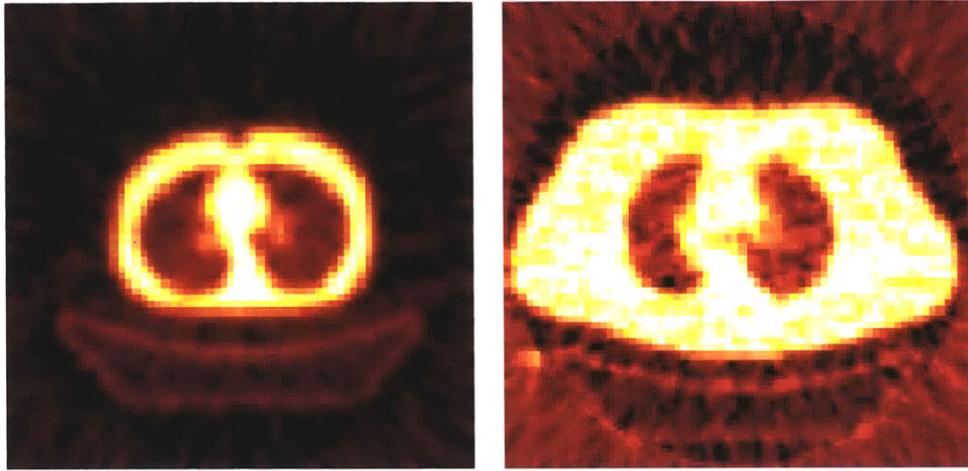


Figure 1.8: Transmission scans taken from a slim subject (left), and obese subject (right). The chest wall thickness affects the number of photon pairs detected by the camera, and thus we see much more noise in images taken from subjects with thick chest walls. In the example above, imaging noise (measured by standard deviation) is about two and a half times higher in the right-hand subject. To compensate, we would need to inject about seven times as much tracer.

Significance of y-intercept

As well as allowing us to quantify noise in the image, this analysis also has the important effect of quantifying the ‘true’ heterogeneity in perfusion (as measured by the normalized variance). This quantity is the y-intercept, c , of the linear regression, because:

$$\frac{\sigma^2(a)}{\bar{a}} = \frac{\sigma^2(\dot{Q}_A)}{\dot{Q}_A^2} \quad (1.5)$$

The assumptions behind this approximation are discussed in section 4.3.1.

The y-intercept is the ‘true’ heterogeneity in the sense that this is the quantity we would measure if we could image for an infinite amount of time, or infuse an infinite amount of tracer. I.e. it is still dependent on the limit of resolution of the camera. Since pulmonary perfusion is generally believed to be fractal in nature, we would see higher heterogeneity in a higher resolution scanner.

Voxel volume

In the above analysis, the gradient of the linear regression is k/V_{voxel} . We can test our model further with a new image in which the resolution is halved by merging adjacent voxels, and seeing how this effects the results. Since the constant k describes attenuation and camera sensitivity, the only change should be in V_{voxel} , which is now four times its previous value (since we are combining pairs of voxel in x and y directions). Thus the gradient should decrease by a factor of four. It does not.

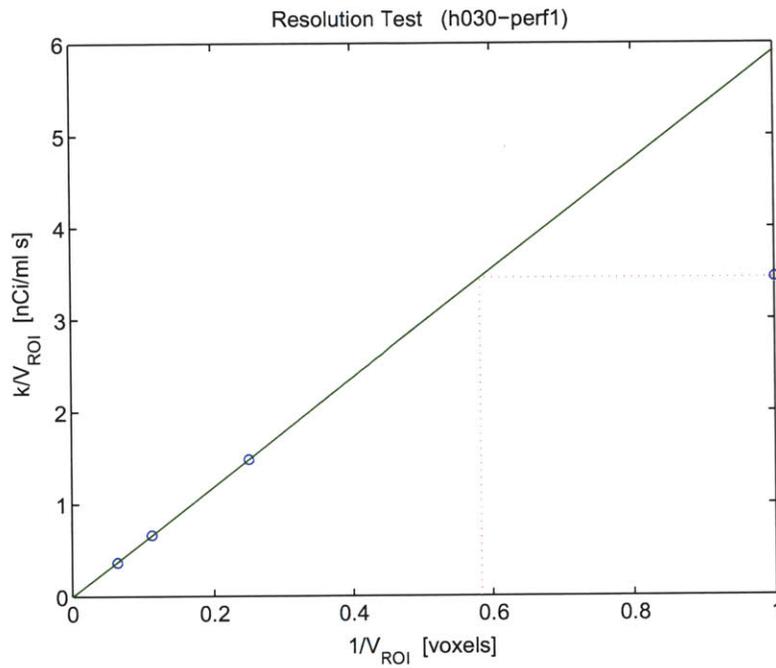


Figure 1.9: Effect of changing resolution (by merging adjacent voxels). The noise gradient is plotted for imaging element volumes of 1, 4, 9, and 16 voxels. The gradient for single voxels is lower than expected because the image was reconstructed with voxel size (4mm) below the effective resolution of the camera. The measured noise on individual voxels equals the expected noise for an element of $1/0.58 = 1.7$ times greater volume, corresponding to a imaging element of 5.2mm side length.

The reason is that the voxel size is actually smaller than the resolution of the camera. When V_{voxel} is below the resolution of the camera, we are effectively interpolating the image; enlarging voxel size only begins to affect the statistical noise once the voxel is larger than the effective resolution. If we degrade the image further, combining nine voxels rather than four, the gradient does decrease by the expected factor of 9/4.

We can estimate the effective resolution of the camera, based on the volume change that would give the observed change in gradient. An example is shown in figure 1.9. This calculation consistently suggests that the effective resolution of the camera is around 5.1 to 5.2mm. [mean=5.146, stdev=0.061, n=66.] This calculation is not just an academic exercise: this information is needed to calculate the expected noise in a large ROI.

Note that we can degrade the resolution of an image in more subtle ways, for example, by filtering the image with low-pass filters of different cut-off frequencies. (A related approach was employed by Venegas & Galletti [16] to explore the fractal nature of pulmonary perfusion: here we are interested only in imaging resolution.)

1.5.2 Signal to Noise

Using the approximate noise model defined above, we expect to see a signal to noise ratio (as defined by mean/stdev) as follows:

$$\text{SNR} = \frac{a}{\sqrt{\frac{kA_{\text{field}}}{\tau V_{\text{ROI}}}}} \quad (1.6)$$

The mean SNR can thus be written:

$$\overline{\text{SNR}} = \sqrt{\frac{\bar{a}\tau V_{\text{ROI}}}{kV_{\text{lung}}}} = \sqrt{\frac{\bar{a}\tau}{m}} \quad (1.7)$$

Where parameter $m = k/V_{\text{ROI}}$ is the gradient of the linear regression. For our studies, a typical value of m would be around 5 nCi/ml²s. Mean concentration in the apnea plateau is typically $\sim 3,000$ nCi/ml. Under these circumstances, a voxel from a 5-second apnea frame would have SNR around unity, while a whole-lung ROI from the same frame would have SNR around 100. A voxel from a 30-second frame in the washout, when, say, 99% of the tracer has washed out would have SNR around 0.3, and the whole lung around 30.

Note that these are mean SNR's. It is worth reiterating the point that while the 'signal' depends on local concentration, noise is mostly determined by total tracer content (and hence mean). Thus areas which have relatively low activity have lower SNR; areas where the activity is higher than average have better SNR.

1.6 Originality and Authorship

I attest that I am the sole author of this thesis, and that the work described herein is wholly my own, but with the following caveats.

Although I single-handedly designed and built the tracer-preparation system described in Part I, almost every stage in the design and implementation followed lengthy discussions with my advisor, Jose Venegas.

Next, the small ROI analysis technique described in chapter 5 was originally described by Kimura et al [8, 9]. However, the technique is herein applied to a different physiologic system, and is refined, extended, and investigated more deeply than in the publications of this author. Thus I feel comfortable representing this as a work of originality.

Lastly, the experimental data from human asthmatics analyzed in chapter 6, and images from other studies, scattered throughout this thesis, are the product of a collaborative effort of our laboratory at Massachusetts General Hospital.

Chapter 2

Tracer Preparation System: Overview

2.1 Introduction

2.1.1 Purpose of Device

The function of the apparatus described here is to prepare a solution of a gaseous radioactive tracer, and to administer the solution to a patient. The tracer used is Nitrogen-13, a short-lived positron-emitting radioisotope of nitrogen, which is dissolved in physiological saline and injected intravenously.

The distribution and kinetics of the tracer within the body are measured, non-invasively, by Positron Emission Tomography (PET). The resulting data are used to study lung physiology.

2.1.2 History

One of the principal barriers to clinical adoption of PET for lung imaging has been the difficulty in preparing the nitrogen-13 tracer solution.

The original apparatus for preparing the tracer (Hammersmith hospital, London, UK) was manually-operated, but this method exposed the operator to very high radiation doses, and was discontinued. [5]

What was required was a system which could be remotely operated, so that the operator could prepare the tracer solution without dangerous exposure.

Such a system was developed by Dr. Venegas, and was successfully used in animal studies at MGH for some years. However, the system was not suitable for clinical use since it was not sterile, and did not incorporate safety features necessary for human studies.

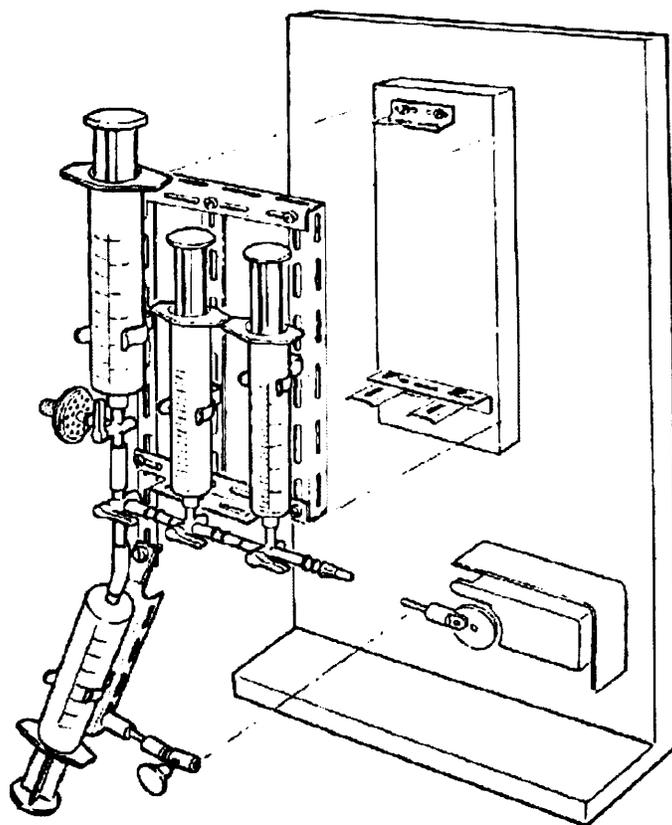


Figure 2.1: Drawing of original tracer preparation system, as used by Rhodes et al. The network of syringes and stopcocks was operated by hand, exposing the operator to high doses of radiation.

2.1.3 New System

We have developed a prototype system for preparation of the nitrogen-13 tracer solution. Our new system is intended for use in pilot studies in humans. No other such systems, commercial or experimental, exist.

The new system overcomes several deficiencies of its experimental predecessor:

- To ensure sterility, all parts of the apparatus that contact the infusate are standard sterile medical components.
- The system incorporates multiple safeguards, both in hardware and in software, which detect and prevent contamination of the infusate with sodium hydroxide, and to detect and prevent the injection of bubbles.
- It is highly automated, and requires minimal operator intervention.

2.2 Principle of Operation

2.2.1 Introduction

Nitrogen-13 is produced in trace quantities by bombardment of carbon dioxide in a cyclotron.

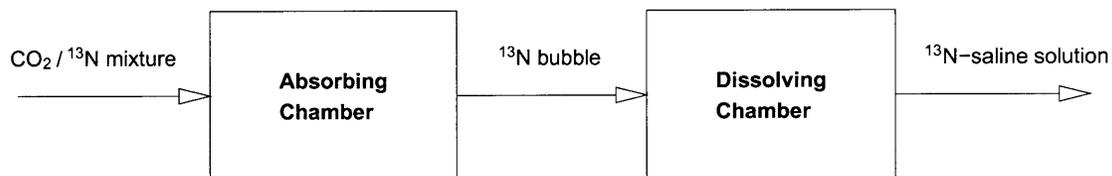


Figure 2.2: Preparation process.

1. The gas received by the apparatus is thus a mixture of carbon dioxide and nitrogen-13. The first stage in the preparation process is to absorb the carbon dioxide. This is achieved by bubbling the mixture into a chamber filled with sodium hydroxide solution.
2. The small volume of nitrogen-13 that has collected at the top of the absorbing chamber is transferred to another chamber, filled with degassed saline, where it is forced into solution under pressure.

The apparatus is operated from a portable computer, located at some distance from the main unit. (Figure 2.3.)

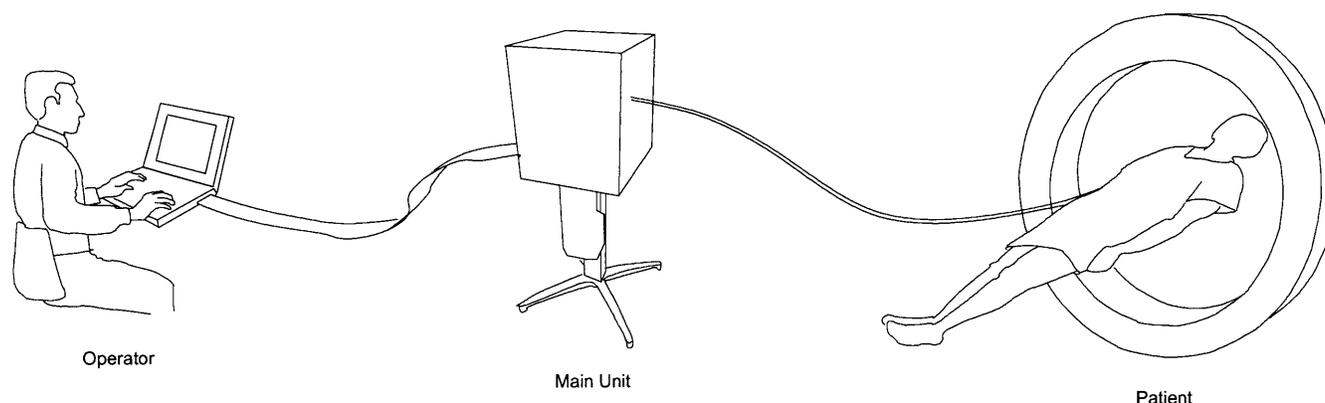


Figure 2.3: System concept.

The main unit encloses a system of syringes and chambers connected by a variety of solenoid valves and automatically-driven stopcocks. A contrast media injector (Medrad Mark IV, modified for computer control) drives the flow through the apparatus, mixes the solution, and gives the infusion to the patient.

2.2.2 Overview of operation

The radioisotope is created in a cyclotron by deuteron bombardment of gaseous carbon dioxide. The bombardment takes around twenty minutes [see section 3.10.1]. The resulting mix of gases remains almost entirely carbon dioxide, but contains minute amounts of nitrogen-13 gas [see section 3.4.1]. This mixture is pumped to the tracer preparation apparatus, which is typically located in the same room as the PET camera.

Gas is received from the cyclotron, and admitted to the absorbing chamber, which is filled with sodium hydroxide solution. The sodium hydroxide reacts with the carbon dioxide, leaving a small bubble of radioactive gas which accumulates at the top of the chamber.

Once all the gas has been received from the cyclotron, the bubble is then transferred to the sterile part of the apparatus. This consists of a system of disposable stopcocks, syringes, and interconnecting tubing. Each of the stopcocks is connected to a servomotor, enabling them to be switched remotely. [See section 3.2.] The flow of fluid through the system is driven by a powered contrast-media injector, which is also controlled remotely.

The bubble of radioactive gas is drawn out of the absorbing chamber by the powered injector. As the bubble exits, sodium hydroxide solution is drawn up out of the chamber after the bubble. A liquid detector is mounted on the line out of the absorbing chamber, and when the meniscus on the trailing edge of the bubble reaches this detector, the first stopcock is switched to allow saline solution to

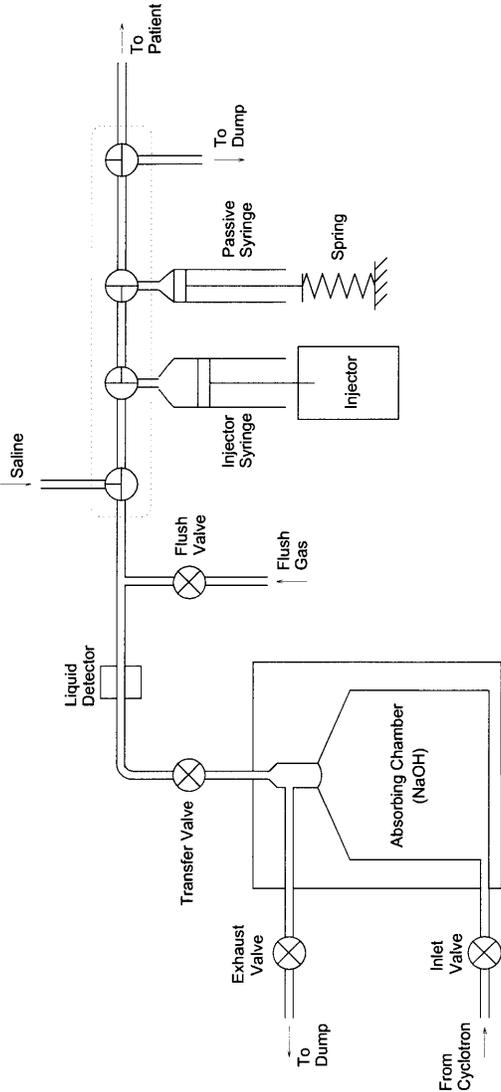


Figure 2.4: Simplified schematic of tracer preparation system

fills in behind the bubble. This allows the bubble to be almost completely transferred to the injector syringe, which is filled with degassed saline solution. [See section 2.2.6.]

The mixture of gas and liquid in the injector syringe is expelled to a second syringe. This second syringe, called the 'passive' syringe, is spring-loaded, so that as the plunger is displaced, the system becomes pressurized. The gas is dissolved by repeated ejection and withdrawal of the mixture, back and forth between the two syringes.

After the mixing is complete, the stopcocks are switched to connect the injector syringe to the patient line. A sample of the tracer is taken and tested for pH (as a final test for NaOH contamination) and its radioactivity measured (to quantify dosing).

The injector is then used to infuse the tracer solution to the patient.

2.2.3 Architecture

The apparatus can be conveniently divided into those parts that do not contact the sterile solution, and those parts which do.

The parts which contact the saline directly must be sterile, and are assembled from disposable medical components. Thus the hydrophobic filter, stopcock manifold, injector syringe, passive syringe, and their interconnections are all off-the-shelf, disposable parts which are replaced before each patient study. These components are mounted on the front panel of the main unit, so that they can be changed quickly.

The remaining parts of the system do not contact the saline, and can thus be constructed from reusable components. The absorbing chamber, the solenoid valves and other components located in the rear compartment (see below) to it are permanently installed.

The system is enclosed in a cabinet which is connected to a high flow-rate vacuum. This maintains a steady flow into the cabinet through its small openings, so that any leaks of radioactivity within the system will be contained and removed. A photograph of the enclosure is shown in figure 2.8.

The cabinet is divided into three compartments. The rear compartment, accessible via a rear door, contains the absorbing chamber and the dump tank, and is watertight so that a catastrophic failure of the absorbing chamber will not result in sodium hydroxide escape. A variety of other permanent flow-measuring and regulating parts are also located in the rear (catalyst, flow meter, pressure transducers, solenoid valves etc.) A central compartment houses all of the electronics of the apparatus, and is protected from contact with any liquid that may leak from a failing component or connection. The front of the cabinet forms a door which encloses the front panel, allowing easy access to the sterile components, which need to be replaced frequently.

2.2.4 Execution Sequence

- Before gas is received from the cyclotron, the system is readied for production: the tubing from the absorbing chamber is flushed with gas, and the remainder of the apparatus is flushed and filled with degassed saline.

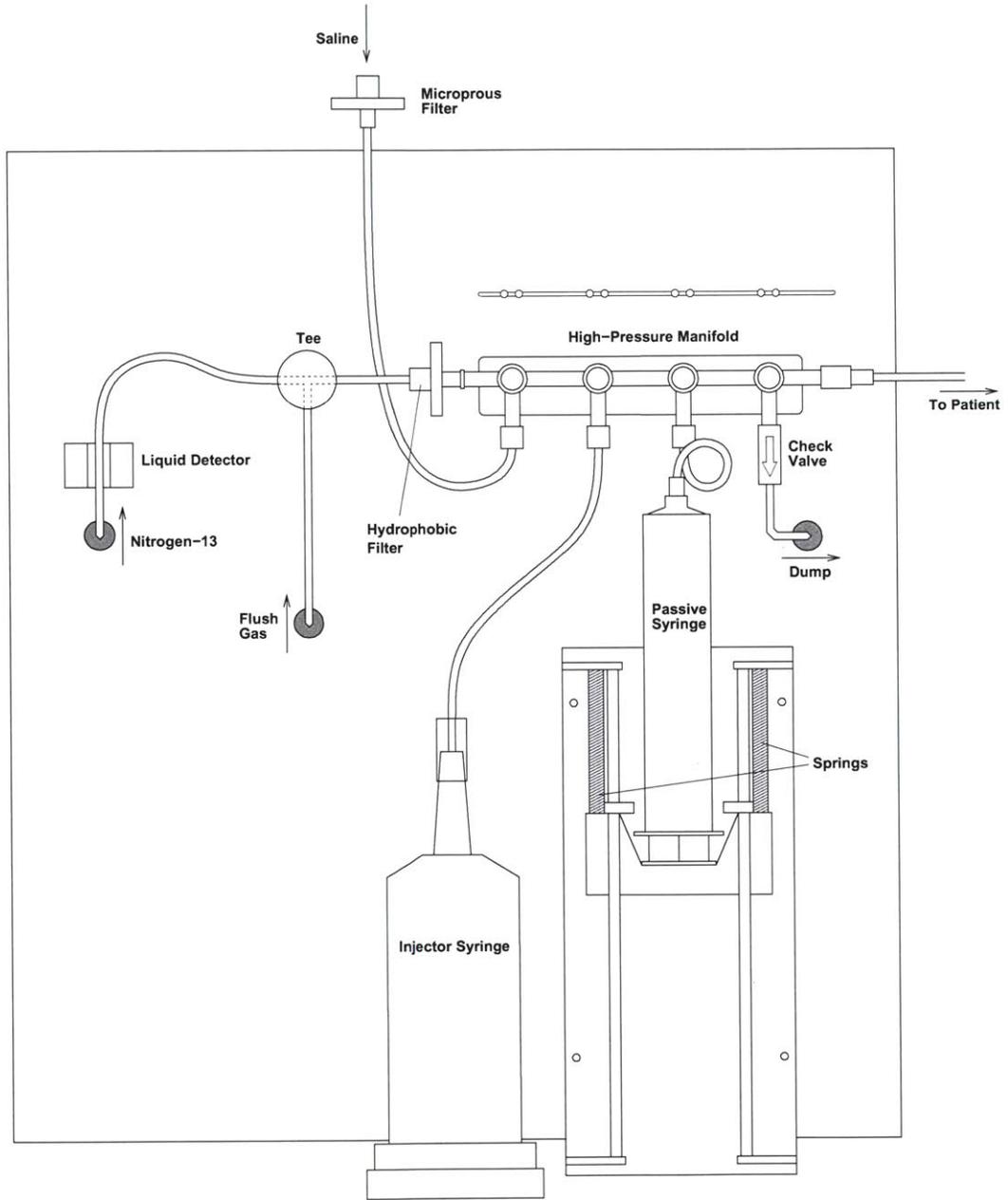


Figure 2.5: Diagram of front panel.

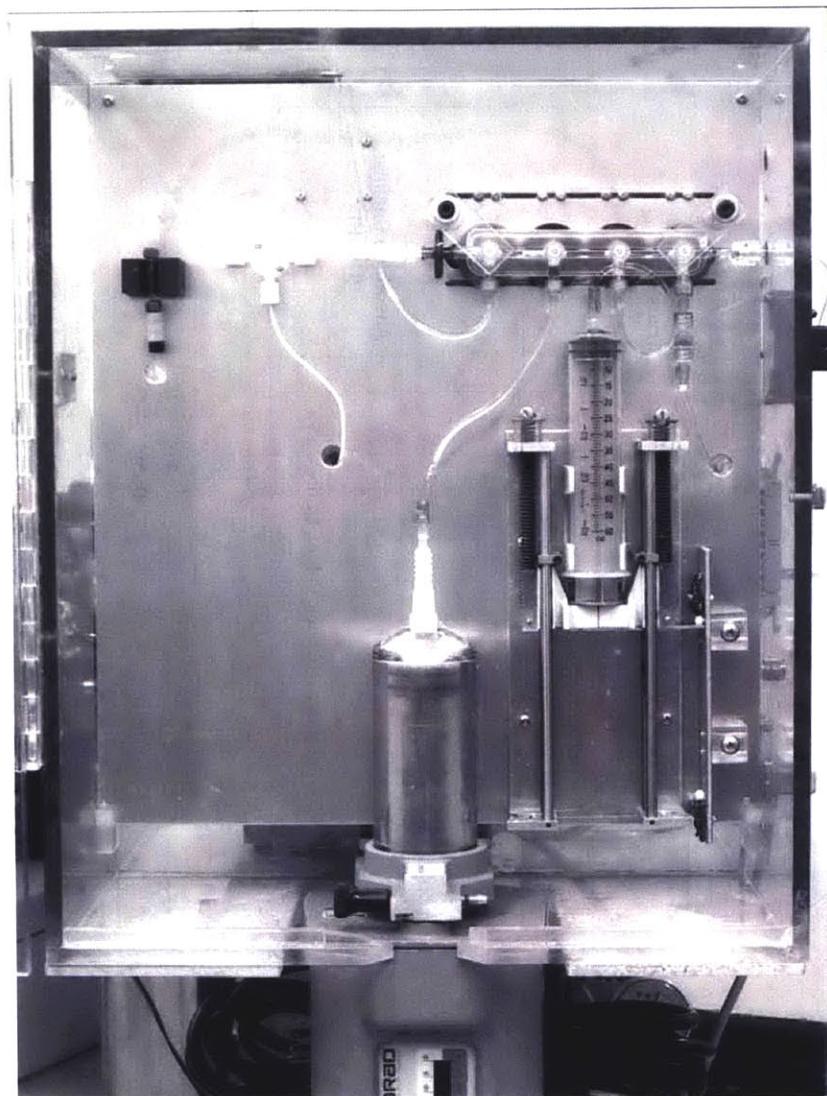


Figure 2.6: Photograph of front panel. Visible are (upper left) the liquid detector, (upper right) stopcock manifold, injector syringe (bottom center), passive syringe (low right).

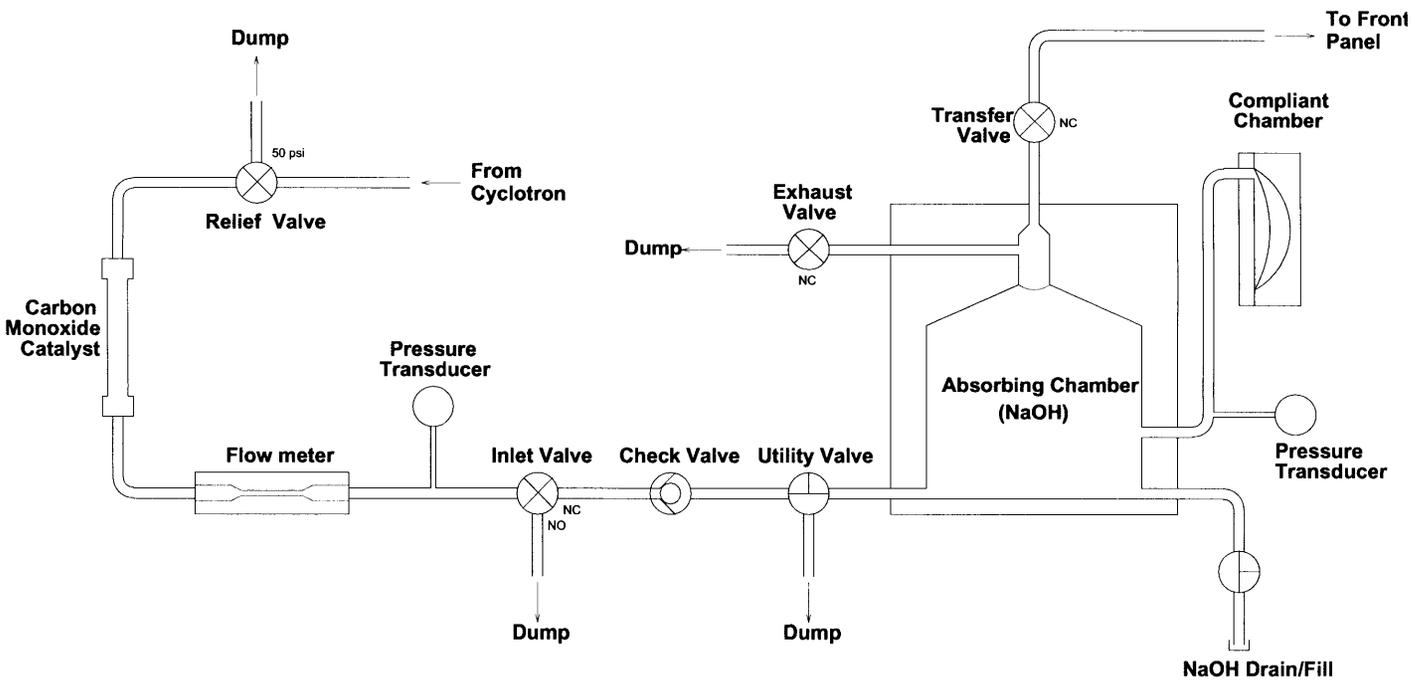


Figure 2.7: Schematic representation of components housed in the rear compartment.



Figure 2.8: Oblique view of main unit. Enclosure is divided into three compartments. The front compartment contains the sterile components which are replaced for each subject. The central compartment houses the electronics. The rear compartment contains the absorbing chamber, dump tank, and other flow components .

- Radioactive gas is admitted to the absorbing chamber and stirred until all carbon dioxide is absorbed.
- The bubble of gas remaining at the top of the absorbing chamber is then transferred to the injector syringe which is filled with degassed saline.
- The mixture in the injector is dissolved by repeated ejection back and forth from the the passive syringe
- Before injection into the patient, a sample of the injectate is expelled into the sampling syringe.
- The injector gives a rapid bolus of tracer solution to the patient.

2.2.5 Design and Function of the Absorbing Chamber

The absorbing chamber is composed of two parts: the main vessel, which is essentially rigid, and a second smaller compartment, which is compliant. The compliant compartment incorporates an elastic membrane that expands to accommodate a volume of gas from the cyclotron. Pressure in the chamber rapidly equilibrates with the gas supply, so that gas enters the absorbing chamber at the rate it reacts with sodium hydroxide. Without this mechanism, the volume of gas admitted to the chamber would be minimal, and hence the absorption rate drastically lower.

The chamber is filled with sodium hydroxide solution, leaving a cylindrical pocket of flush gas at the top of the vessel, where the nitrogen-13 collects. When gas is admitted, the fluid in the chamber is magnetically stirred to promote absorption of carbon dioxide. The agitation is carefully controlled so that the pocket of gas at the top of the chamber remains intact. In this way, no liquid infiltrates the tubing leading to the rest of the apparatus, where small droplets of liquid might cause false triggering of the liquid detector or blocking of the hydrophobic filter.

An important aspect of the design of the compliant compartment is that it is only compliant to positive volumes. That is, volume can be added to the chamber, but not withdrawn. Once the carbon dioxide is absorbed, and the bubble of nitrogen withdrawn, the membrane is pressed flat, and the chamber becomes rigid. Thus it is impossible to suck sodium hydroxide directly out of the absorbing chamber and into the rest of the system.

2.2.6 Encapsulation and Transfer of the Nitrogen-13 Bubble

1. After system preparation is completed, the state of the apparatus is as depicted in figure 2.10.1. The tubing upstream (to the left) of stopcock #1 is filled with flush gas, and the tubing downstream (to the right) is filled with degassed saline.
2. The bubble of radioactive gas is drawn up out of the absorbing chamber, and towards the injector syringe. Sodium hydroxide solution is also drawn out of the absorbing chamber on the trailing edge of the bubble.

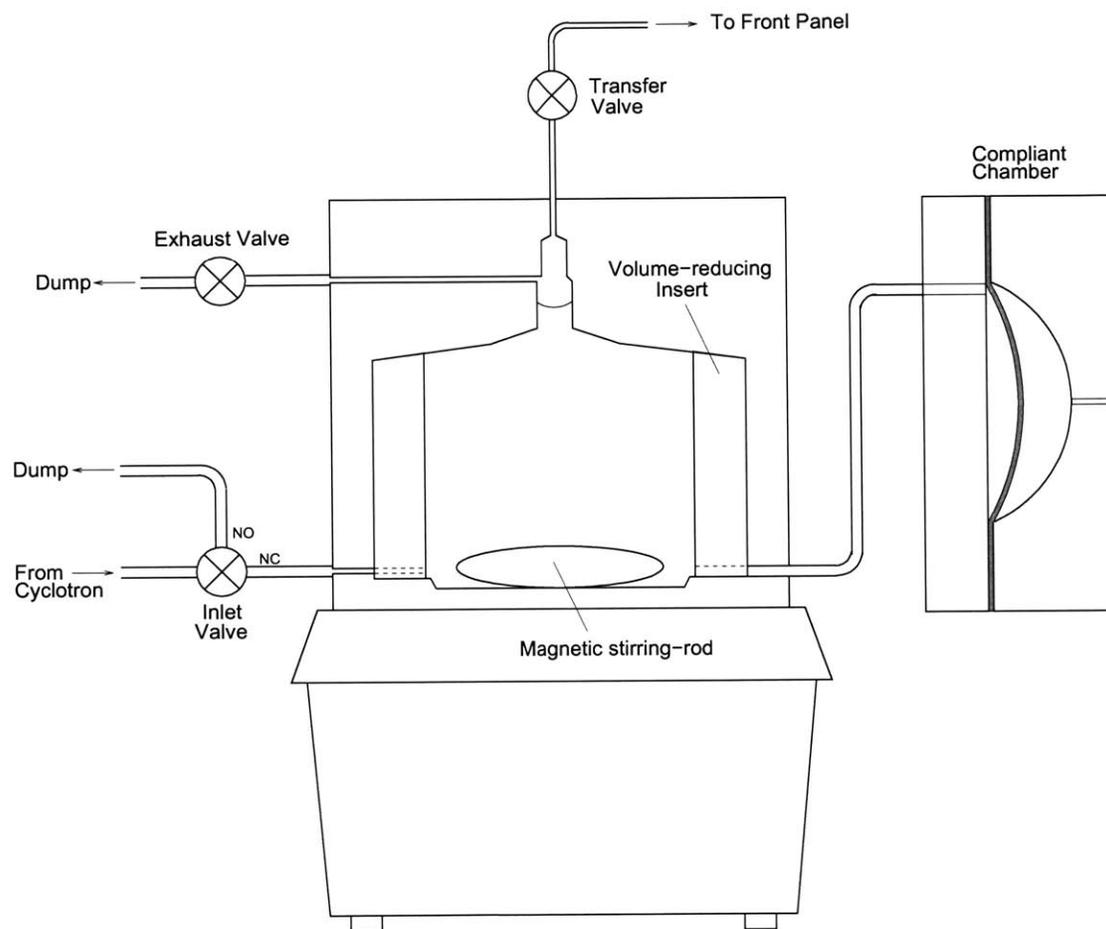


Figure 2.9: The absorbing chamber.

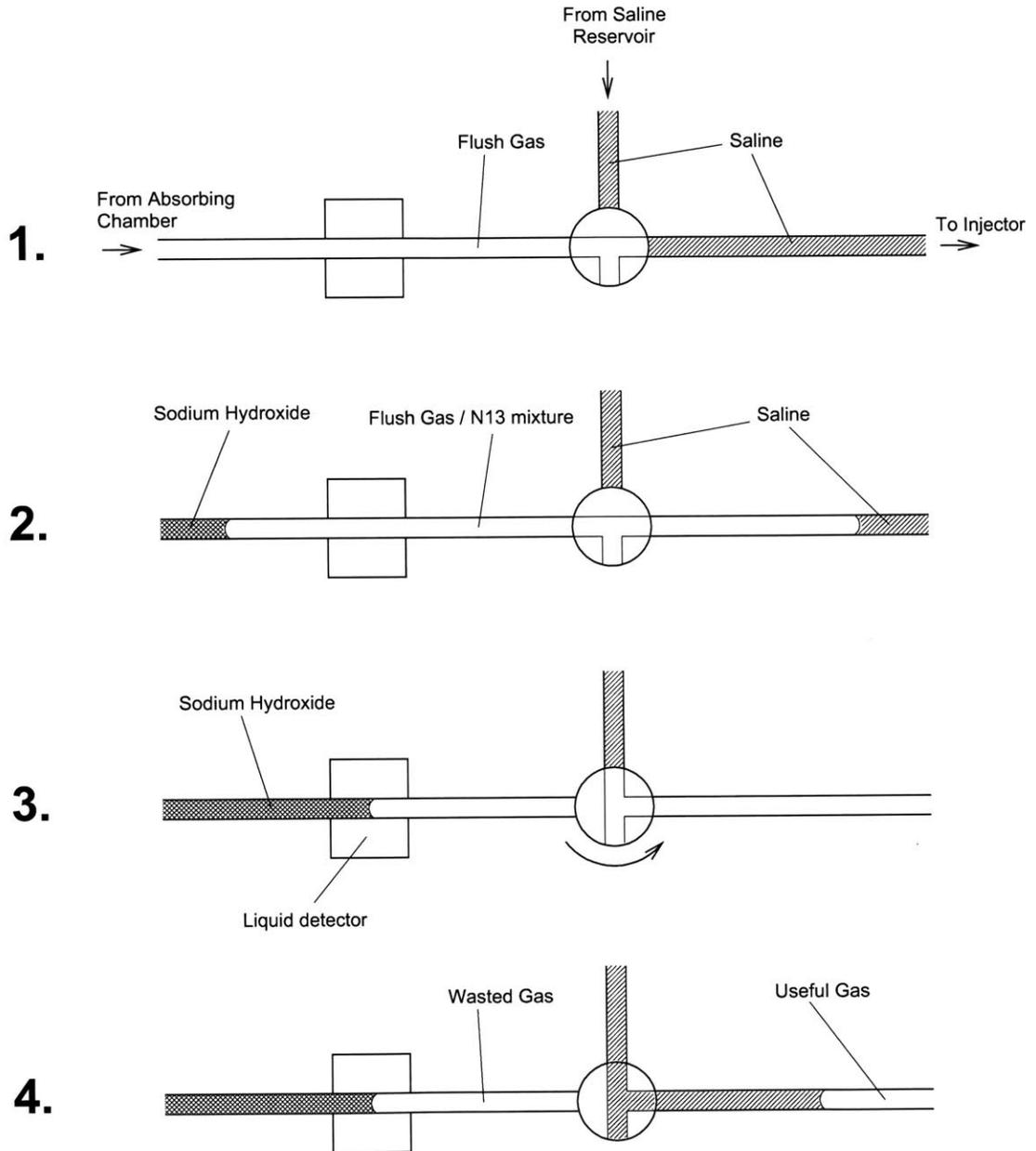


Figure 2.10: Bubble encapsulation and transfer sequence.

3. A liquid detector is installed on the line, just upstream of the first stopcock. When the sodium hydroxide reaches this point, the transfer valve is closed and the stopcock is turned to fill in behind the bubble from the saline reservoir.
4. The bubble, thus encapsulated by saline solution, can be transferred through the apparatus to the injector syringe.

The volume of gas that is wasted is approximately 0.7 ml, of which the majority (0.4 ml) is due to the large dead volume of the luer connections to the hydrophobic filter. In animal studies, where we can tolerate a slightly increased risk of NaOH contamination, the filter can be removed entirely to improve yield.

2.2.7 Dissolution of Bubble in Saline

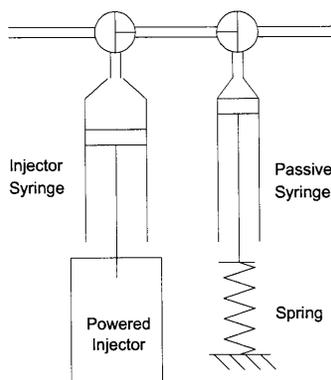


Figure 2.11: The syringe driven by the powered injector is connected to a passive, spring-loaded syringe. The bubbles is dissolved by repeated expulsion and withdrawal of the mixture back and forth between the two syringes.

After the bubble of gas is completely drawn into the injector syringe, the stopcocks are switched so that the injector communicates with the passive syringe.

The mixture is then vigorously expelled into the passive syringe. This promotes dissolution of the gas in several ways. Firstly, the surface area of the interface is dramatically increased by breaking the bubble of gas into many smaller bubbles. Secondly, the pressure is raised, so increasing the potential for diffusion. Finally, the strong currents and highly turbulent flow mix the liquid very well, reducing concentration gradients which retard the process.

The mixture is repeatedly expelled and withdrawn from the passive syringe until all of the N13 is dissolved. After the mixing process, a volume equal to the volume of gas originally drawn into the injector syringe is expelled to the dump reservoir. This ensures that any undissolved gas is ejected from the system.

2.2.8 Percussing Mechanism

Problem Description During testing, we observed that, after mixing, small bubbles of gas remain around the upper rim of the injector syringe.

Although the direct effect on the system yield is minimal (as the fraction of gas that is not dissolved is slight), these bubbles are a cause for concern for two reasons. Firstly, they have a destabilizing effect because previously-dissolved gas evolves out of solution into the bubbles. Secondly, as the bubbles swell, they have a tendency to detach from the walls of the syringe and to enter the stream of liquid that is infused to the patient.

The presence of these bubbles poses minimal risk to the patient as they are very small (<0.1 ml) and are routinely detected by the bubble detector on the infusion line. Furthermore, if any bubbles were to escape detection, they would routinely lodge in the pulmonary capillaries and gradually dissolve. (It is consequently important to note that right-to-left cardiac shunt is a contraindication for the technique.)

However, these bubbles pose a significant experimental inconvenience, as their detection in the infusion stream triggers a system shutdown and immediately arrests the infusion. Consequently the PET imaging sequence is aborted, and must be repeated using a reduced volume of tracer solution (and consequently degraded image quality) after the lines have been flushed. Alternatively, a large delay is introduced while a new batch of the tracer is prepared in the cyclotron, which may compromise the validity of an experiment whose protocol calls for a precise schedule.

Remedy Our remedy to this problem is to install a mechanism to dislodge these bubbles during the mixing. Instead of sitting on the walls of the syringe, that they are entrained in the fluid that is exchanged between the two syringes, and are dissolved with the rest of the nitrogen-13.

The mechanism is depicted in figure 2.12, and consists of a small motor with a freely-hinged arm which swings around and lightly impacts the top of the injector syringe. This provides a percussive, tapping action which dislodges the bubbles.

In addition to its use during the mixing and dissolution stage, the tapping mechanism is also activated before the infusion to the patient. Any bubbles that might remain (or have formed) in the syringe will be dislodged and removed in the final flush that is always performed immediately prior to infusion.

2.2.9 Sampling

The patient is connected to the unit by twin 72" lines, one for outward, the other for returning flow. This enables the lines to be flushed with tracer as close as possible to the IV site, minimizing the dead volume of the infusion. The tracer is manually sampled at this location, slightly downstream of the IV site, where the distance from the main unit decreases radioexposure to the operator.

Once the dissolution of the gas is complete, the lines to the patient are flushed, a small (1ml) sample is taken, and its radioactivity measured in a well-counter cross-calibrated with the PET camera. As

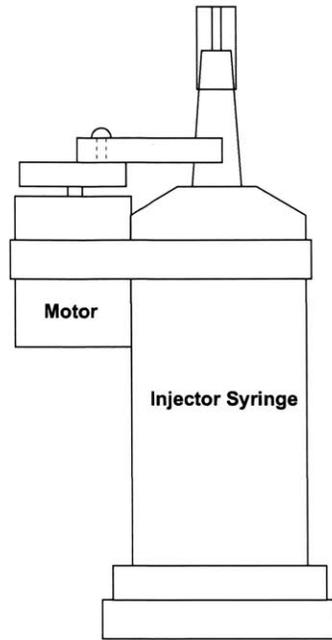


Figure 2.12: Percussing mechanism. A hinged arm, driven by a small DC motor, taps the top of the injector syringe. The repeated impacts from the arm dislodge small bubbles around the top rim of the syringe.

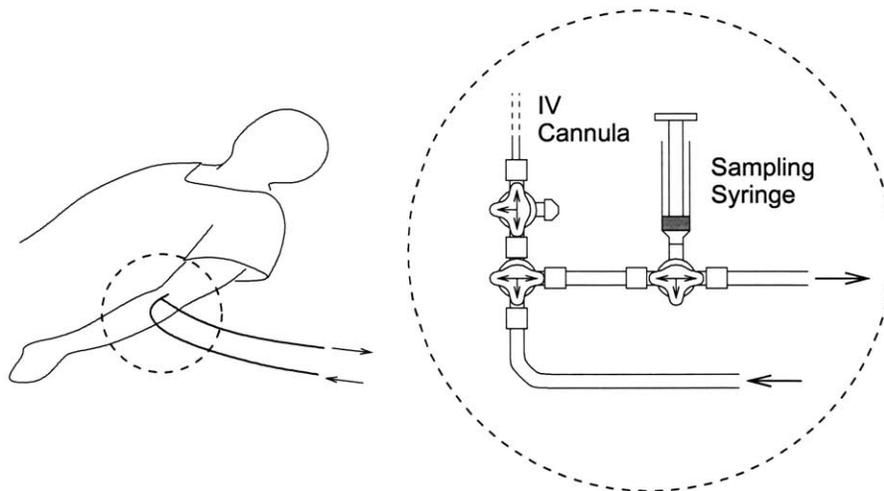


Figure 2.13: Arrangement of infusion lines and stopcocks at infusion site.

a final safeguard against sodium hydroxide contamination, the pH of the tracer solution is checked manually at this point.

2.2.10 Infusion

Immediately prior to infusion to the patient, the infusion lines are flushed for a last time to clear any microbubbles that may have formed as the saline warms up to room temperature.

In human studies, 30-35 ml of tracer are typically infused at 5 ml/s.

Immediately following the infusion, the IV cannula is manually flushed with regular saline. This ensures that the full bolus of tracer solution is fully convected out of the IV site and out of the arm.

Chapter 3

Tracer Preparation System: Detail

3.1 Software Architecture

3.1.1 Platform

The control software is written entirely in LabView (National Instruments, Austin, TX). This is a graphical programming language intended mostly for laboratory automation.

LabView has many strengths: it is intuitive, and quick to learn; interfacing with data acquisition hardware is painless (particularly with National Instruments hardware); parallel programming is easy and building graphical interfaces very rapid. However, the graphical, simplified nature of the language also has disadvantages: many conventional debugging tools and techniques cannot be applied; algorithms execute slowly; sequential programming is awkward. Some fundamental issues are startling: for example, it is impossible (in general) to make a printout of a LabView program.

3.1.2 General Architecture

Since the apparatus is directly connected to patients, safety is of critical importance. We need to be able, at any time, to instantly abort execution and switch the system to a safe state.

The software has multiple threads that are executed in parallel. Since a thread cannot guarantee that it is the only part of the program accessing the hardware, each hardware access must be atomic. This is accomplished by accessing each device only through a single, non-reentrant interface routine. In normal use, resource conflicts are not a problem. For example, it is hard to envisage a useful program sending two simultaneous motion commands to the injector. However, regulatory mechanisms are vital to ensure correct behavior during a system shutdown.

A shutdown cannot be safely accomplished merely by interrupting program execution. We need to ensure that solenoid valves are closed, that the injector is stopped, and that stopcocks are turned to prevent flow to the patient. To achieve this, the shutdown commands must not only be atomic, but also must immediately disable further access.

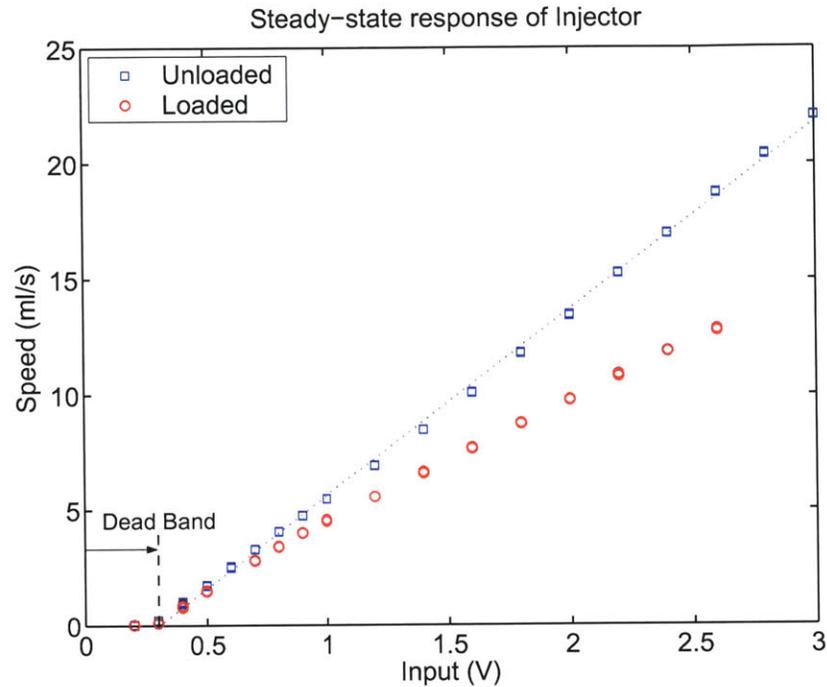


Figure 3.1: Steady-state response of Injector. For small voltages (and hence motor torque) the injector does not move: an input of >0.3 V is required to overcome friction. Thereafter the response is linear.

For example, if the abort button is pushed midway through a injector motion, the abort thread starts running in parallel with the injector thread. Since the injector motion thread will be continuously sending commands to the injector, the abort sequence must not only stop the injector, but also must disable subsequent access, to prevent the still-running injector thread from restarting the injector motion (which would then continue *even after program execution is halted*). This is accomplished by unsetting global enable flags that are checked for each hardware access.

3.1.3 Injector control algorithm

The electronics of the injector were modified so that the internal feedback error signal is replaced by an analog signal generated by the computer. The injector electronics implement a simple proportional controller, and thus motor current (\sim torque) is proportional to the error signal. With the feedback interrupted, then, the motor current is proportional to the computer-generated input.

The friction in the injector head is significant, and results in a dead band in the response. Once the static friction is overcome, the response is fundamentally linear. In figure 3.1, steady-state response is plotted for two conditions: ‘unloaded’, in which the injector syringe is air-filled, and not connected

to other components, and ‘loaded’, in which the syringe is water-filled and expelling through two 72” infusion lines in series.

The algorithm implemented is a PD controller with a non-linearity on the output to compensate for the dead band, improving the dynamics for small movements. Commands to the injector are given as specified excursions (i.e. ‘move $+x$ at rate y ’). These are executed with the PD controller running in a loop while the set point is changed according to the scheduled motion. The loop terminates when the injector reaches its target final position to within a set tolerance.

The software implements a minor enhancement to this method. As described above, the desired motion will always take longer than scheduled (reflecting the fact that a constant error develops during steady-state linear motion), and will always under-reach the final position (since the loop terminates as soon as the injector gets within tolerance). To remedy these problems, a ‘settling-time’ (~ 3 time-constants) is subtracted from the scheduled duration of the motion, so that the actual peak velocity is slightly higher than specified. The PD loop subsequently runs for the settling-time without checking. Thereafter, control continues until the injector position is within tolerance, which under normal circumstances is immediately.

3.1.4 Screenshots

In the sections above, I have firstly given an overview of how the system software is architected to ensure safe operation, and secondly summarized the injector control algorithm (which is a traditional mechanical engineering subject). I do not propose to discuss the details of the software further. However, I include a few screenshots of the user interface in order to give the reader some idea of how the system appears to the operator.

3.2 Stopcock Actuators

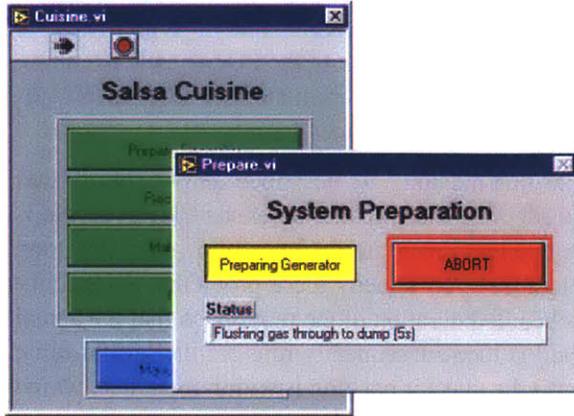
The design of the system requires computer-controlled, sterile, three-way valves. A key challenge in the system development was to devise an effective and economical means to accomplish this.

Cheap, disposable, three-way valves, called ‘stopcocks’, are ubiquitous in medical institutions. An obvious possibility was thus to develop a method to manipulate these valves automatically.

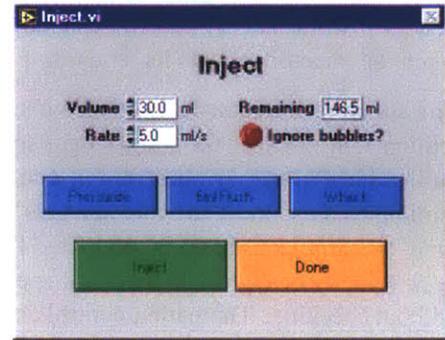
The system we developed uses a servomotor to control the position of the stopcock. These servos are designed for use in radio-controlled models, and are inexpensive and easy to interface. The body of the stopcock is mounted in a machined block, and a flexible moulded coupling mates with the handle of the stopcock. These components are shown in figure 3.5.

The simple assembly allows a stopcock to be positioned accurately (~ 1 degree) and rapidly (~ 0.2 seconds). Additionally, a trivial modification to the servomotor provides external position feedback so that correct functioning of the device can be monitored.

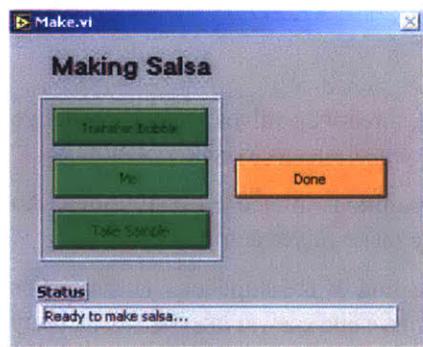
The stopcock actuators are described above as they were originally developed, to actuate individual stopcocks. In a design revision of the tracer-preparation apparatus, the individual stopcocks were



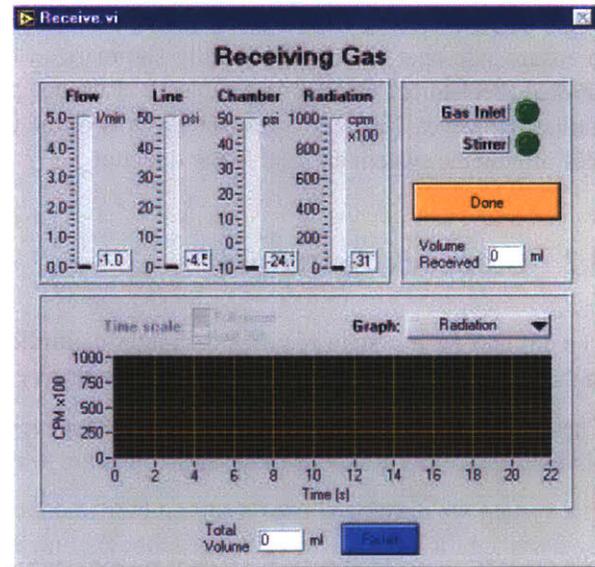
(a) Root window, with preparation window open.



(b) 'Inject' window



(c) 'Make Salsa' window



(d) 'Receive' window (data shown not valid)

Figure 3.2: Various screenshots of the control software.

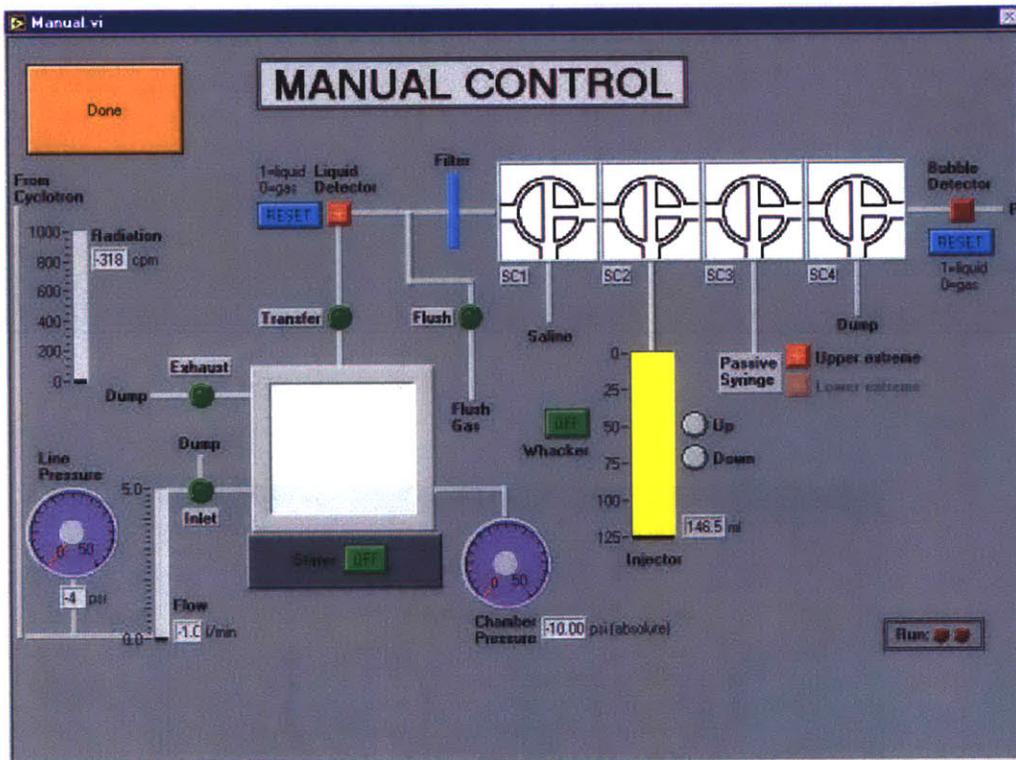
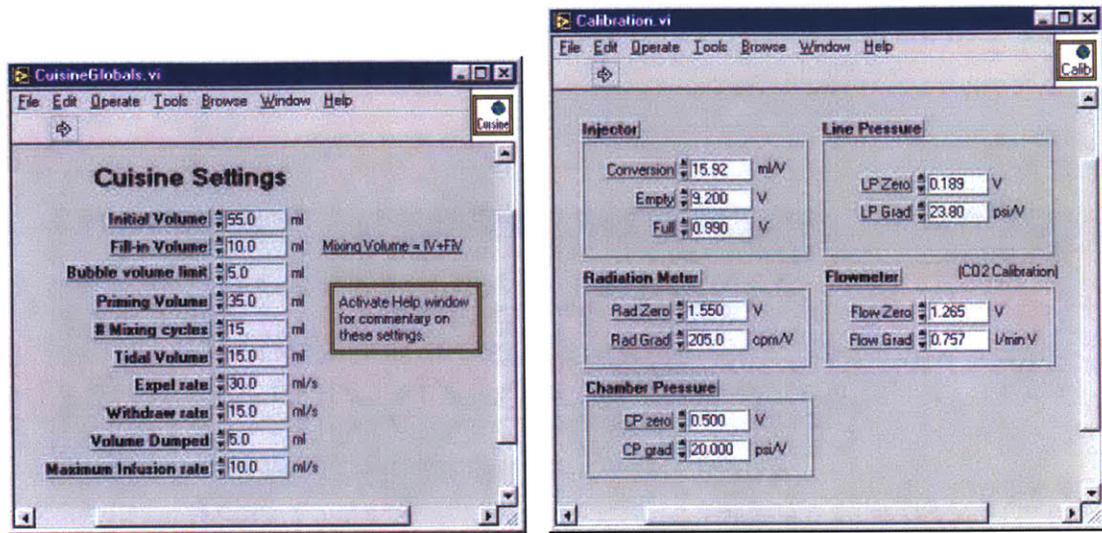
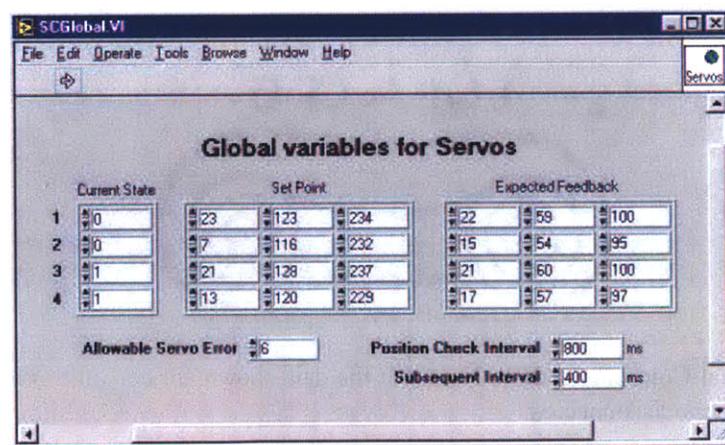


Figure 3.3: 'Manual Control' window. Note that the data shown are not valid: the screenshot was taken with the sensors disconnected.



(a)

(b)



(c)

Figure 3.4: Various configuration windows from control software.

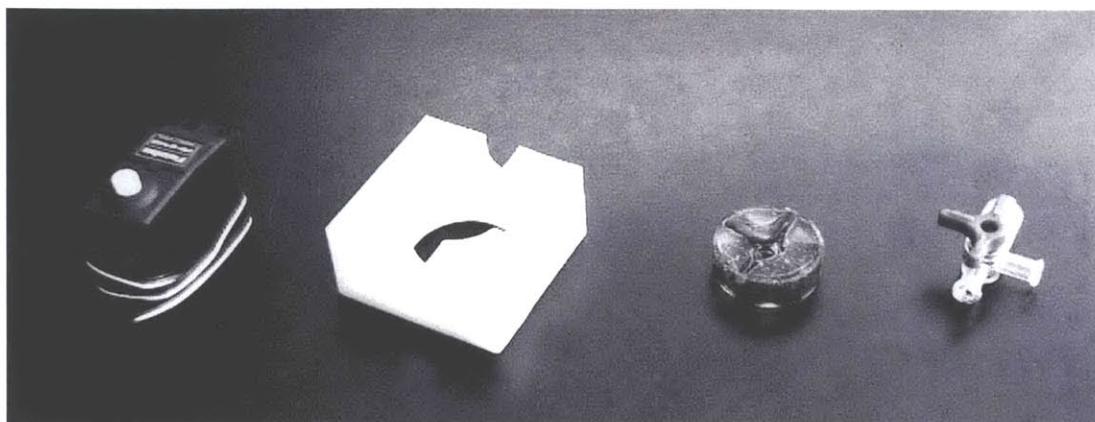


Figure 3.5: Components of the stopcock controller: (from L to R) servomotor, mounting block, coupling, and stopcock.

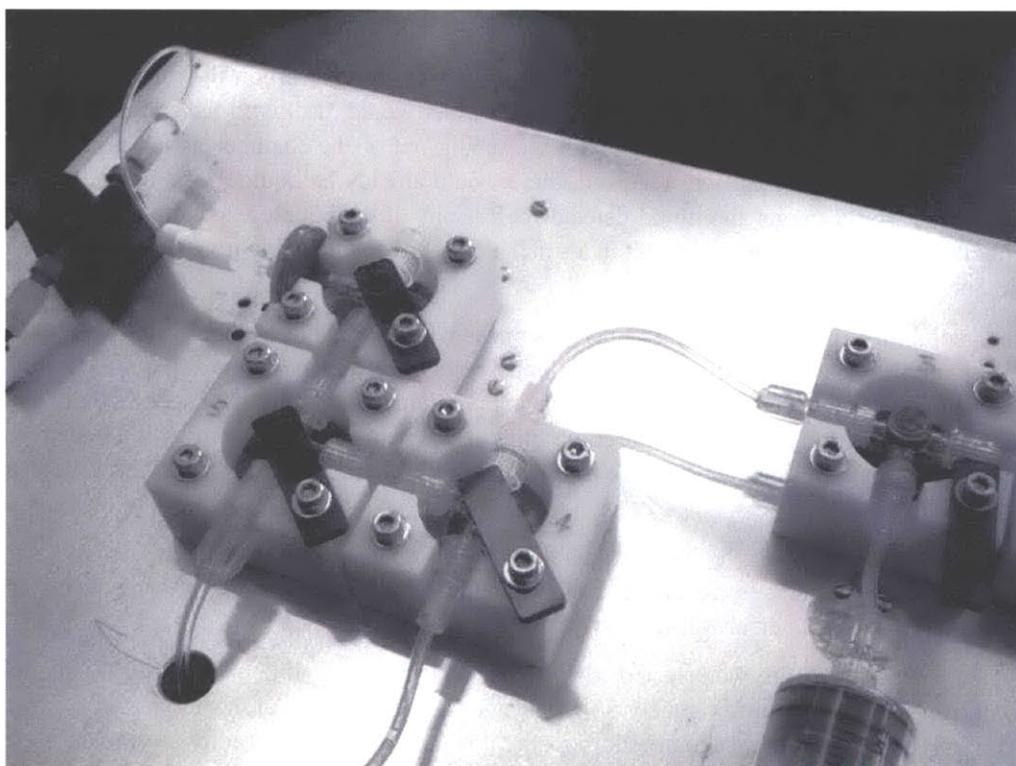


Figure 3.6: Close-up of original front panel showing stopcock controllers with discrete components.

replaced with an integrated, four-stopcock 'manifold'. [Figure 3.7.] This optimization obviates the need for interconnecting components and significantly reduces the time taken to assemble the sterile components. Additionally, the simplification reduces the chance of contamination during assembly.

The servomotors are controlled from an SV201 servomotor controller board (Pontech, Rancho Cucamonga, CA) which utilizes a PIC16C73 microcontroller to provide six channels of pulse-width modulated output (only four are used in the current design). Communication to the board is via a RS232 serial connection. The board also provides five channels of low-resolution (8-bit) A-to-D, which are employed to read the potentiometer output, giving position feedback from the servos.

3.3 Choice of Flush Gas

The primary mechanisms to prevent sodium hydroxide contaminating the injectate (liquid detector, hydrophobic filter) require that the tubing between the absorbing chamber and the sterile components be filled with gas. A tank of pressurized gas is used to flush this tubing, and to leave a small initial volume at the top of the absorbing chamber.

The principal requirements for this gas are that (1) it does not react with sodium hydroxide, (2) that it has good solubility, and (3) that it is known to be bio-compatible and non-toxic.

If the first condition is not satisfied, for example if we were to use carbon dioxide as a flush gas, then the initial volume at the top of the absorbing chamber would react with the sodium hydroxide, and the sodium hydroxide solution would fill the tubing out of the chamber as far as the transfer valve. When the cyclotron gas was received, there would already be liquid in the line ahead of the bubble, which would trigger the liquid detector before any of the bubble was transferred. (Simply ignoring the liquid detector would result in a small amount of NaOH being transferred, ahead of the bubble, into the injectate.)

The high solubility is required because flush gas mixes with the bubble. A significant amount (~1.4 ml) of the bubble is composed of flush gas, and must be dissolved in the saline solution. It is already hard enough to dissolve the nitrogen, without having problems dissolving other species that we have introduced! More seriously, although the solubilities of different gases are essentially independent at low concentrations, if one species is close to its solubility limit, the solubility of others is decreased [3].

Finally, the flush gas should be bio-compatible and non-toxic. In order to facilitate approval of the system for use in humans, its biochemistry should be well-understood, non-controversial and, ideally, familiar to the medical world.

Motivated by these criteria, we selected nitrous oxide as flush gas. It does not react with sodium hydroxide, has solubility approximately 30 times higher than nitrogen (see section 3.5.0.5), and has longstanding use as anesthetic agent. Additionally, it is inexpensive and readily available.

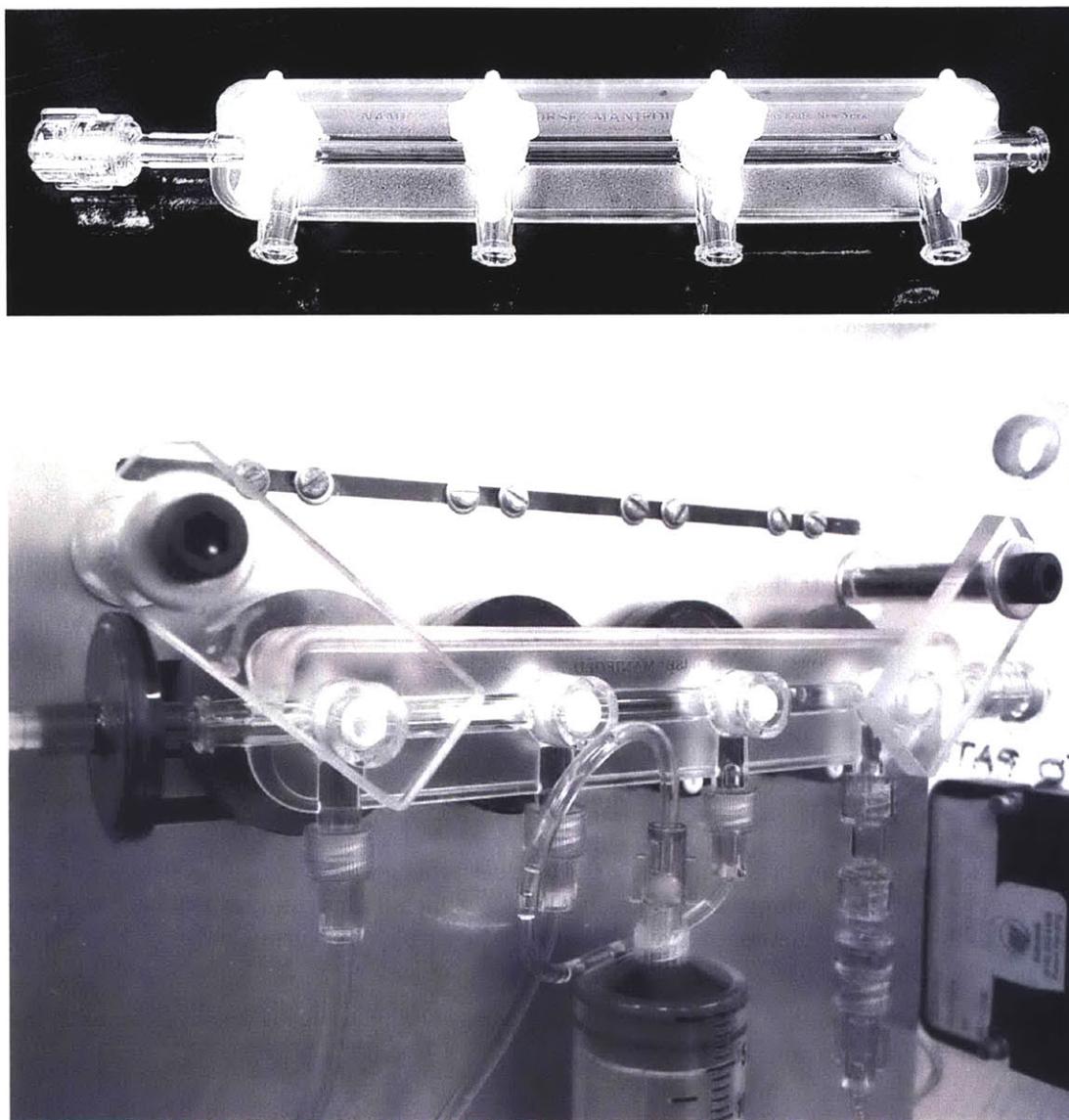


Figure 3.7: Photographs of integrated stopcock manifold: in isolation (top) and in position on front panel (below).

3.4 Bubble composition

Of fundamental importance is the composition of the bubble of concentrated gas that we are attempting to dissolve. As we shall see in the following section, the volume fraction of actual nitrogen-13 is tiny: the bubble is essentially entirely impurities.

3.4.1 Nitrogen-13 volume

We can use the amount of activity received to estimate the actual volume of nitrogen-13 in the bubble.

The total activity received from the cyclotron is of the order of 100mCi, which is 3.7×10^9 Becquerel (disintegrations per second), and the half-life of nitrogen-13 is almost exactly 598 seconds.

Radioactive decay is exponential:

$$N(t) = N_0 e^{-\alpha t}$$

The rate of decay is thus:

$$\frac{dN}{dt} = -\alpha N_0 e^{-\alpha t} = -\alpha N$$

Hence we can estimate N , the amount of nitrogen-13:

$$\begin{aligned} N &= -\frac{1}{\alpha} \frac{dN}{dt} = -\left(\frac{t_{1/2}}{\ln 2}\right) \frac{dN}{dt} \\ &= \left(\frac{598}{0.69}\right) 3.7 \times 10^9 = 3.2 \times 10^{12} \text{ atoms} \\ &= 5.3 \times 10^{-12} \text{ moles} \end{aligned}$$

At room temperature and pressure, 1 mole of gas molecules occupies approximately 24 liters. So the expected volume of nitrogen-13 is around 1.3×10^{-10} liters. This assumes that each molecule contains only one atom of nitrogen-13, i.e. that the gas molecules are heterogeneous ($^{13}\text{N} - ^{14}\text{N}$). If the gas is diatomic nitrogen-13, the volume is half this quantity.

This calculation suggests that the actual nitrogen-13 content of the bubble is absolutely minute: less than one millionth of its volume.

3.4.2 Analysis of Bubble Composition

We made multiple attempts to assess the composition of the bubble of irradiated gas with mass spectrography. Early results (with apparatus that was sensitive only to the fundamental peak) were confusing, primarily due to the coincident molecular masses of both carbon dioxide and nitrous oxide (44), and of carbon monoxide and nitrogen gas (28). A more successful analysis is described in detail in appendix A.4, and is summarized here.

In this analysis a sample of the gas was taken to a custom-built mass spectrograph in the department of molecular biology at MIT. The spectra of two samples were measured: the first (sample A) was a 'dummy run', in which gas was received from the cyclotron that had not been bombarded; the second sample (B) was a bubble from a normal, twenty-minute bombardment.

The spectra were compared with reference spectra for air, nitrous oxide, and carbon dioxide. These are thus the only species that this analysis has the capacity to resolve. The most obvious omission is carbon monoxide, which was not tested due to the difficulty in obtaining an adequately pure sample. However, according to other tests (see appendix A.1), there should be virtually no carbon monoxide remaining once the cyclotron gases have passed through the CO catalyst.

Sample	CO ₂	N ₂ O	Air	Total
A	24.81%	25.26%	49.02%	99.08%
B	18.34%	53.21%	27.82%	99.37%

A linear combination of the candidate components describes, with very good accuracy, the spectra of the samples. Any other species present in significant quantity must, at the minimum, have both fragments of coincident mass, and fundamentals that are either coincident with these peaks, or have molecular mass greater than 70 (the upper limit of the spectrometer used).

3.5 Solubility of gases.

3.5.0.1 Henry's Law

The solubility of a gas in a liquid is most commonly described by Henry's law, which can be stated as follows: the amount of gas that dissolves in a liquid at equilibrium is proportional to the partial pressure of the gas in communication with the liquid. The coefficient of proportionality is the *solubility* of the gas, s .

$$n_l = s.V_l\rho_l P_g = s.V_l\rho_l x_g P \quad (3.1)$$

In the above definition, s is defined in terms of the following quantities: n_l is the amount of substance (number of moles) of gas dissolved in the liquid phase; V_l and ρ_l are the volume and density of the liquid, respectively; x_g is the fraction of the gas phase that is comprised of the species described; P is the pressure of the gas phase. (The partial pressure is thus the product of the latter two quantities.)

3.5.0.2 Partition Coefficient

Solubility can be quantified in a startling diversity of ways, for example as Bunsen coefficient, Kuenen coefficient, mole fraction solubility, and Henry's constant. I shall try to minimize unnecessary proliferation of such measures (which are in most cases trivially related).

Henry's constant solubility, s , is defined above. It will also prove convenient to utilize the *partition coefficient*, λ , which is defined as the ratio of amount of substance in gas phase to the amount of substance dissolved in liquid phase. when the volume of the two phases are equal. (Note that the gas phase need only be partially occupied by the species in question.)

$$\lambda = n_g/n_l \quad (3.2)$$

Note that this measure only describes the solubility for a specific temperature and pressure. The two measures are related as follows.¹

$$\lambda = R_0 T \rho_l s \quad (3.3)$$

Where R_0 is the universal gas constant, and T is the temperature of the gas and liquid (which are in equilibrium).

3.5.0.3 Variation of solubility with temperature

The solubility of gases in water decreases with increasing temperature, tending to zero at boiling point. A graph of solubility vs. temperature for nitrogen is shown in figure 3.8

Other gases have this characteristic sigmoidal form, although the curves are not geometrically precisely similar, as shown in figure 3.9.

3.5.0.4 Variation of solubility with dissolved salts

The presence of dissolved salts in water also decreases the solubility of gases. For example, data for the solubility of nitrogen in water with varying salinity are shown in figure 3.10.

Saline for IV use is labeled "0.9% Sodium Chloride... 900mg/100ml", i.e. a 0.225M solution, or 9 ‰(parts per thousand). I do not have data on the temperature dependence of this effect: consequently, I shall assume that the fractional decrease is constant over the range of temperatures relevant here (0-37°C). Based on data shown in figure 3.10, we expect the salinity of this solution to lower the solubility of nitrogen to ~93% of its value in pure water.

3.5.0.5 Empirical data at physiological conditions

There is a wealth of literature concerning solubility of gases in biological fluids. Reported values vary significantly from study to study, depending on the method used to estimate solubility. The following tables are based on the review paper by Lango et al [10], which summarizes results from literally hundreds of investigators.

¹When using this formula to convert between λ and s , care must be taken to ensure consistency of units. The commonly-used SI definition of R_0 ($8.314 \text{ mol kg}^{-1} \text{ K}^{-1}$) implicitly uses Pascals (Pa) as units of pressure, whereas s is usually expressed (e.g. in CRC handbook) in units of $\text{mol kg}^{-1} \text{ bar}^{-1}$. Either atmospheric pressure (101.2 kPa) must be explicitly included, or a consistent definition of R_0 used ($8.206 \times 10^{-5} \text{ m}^3 \text{ atm mol}^{-1} \text{ K}^{-1}$).

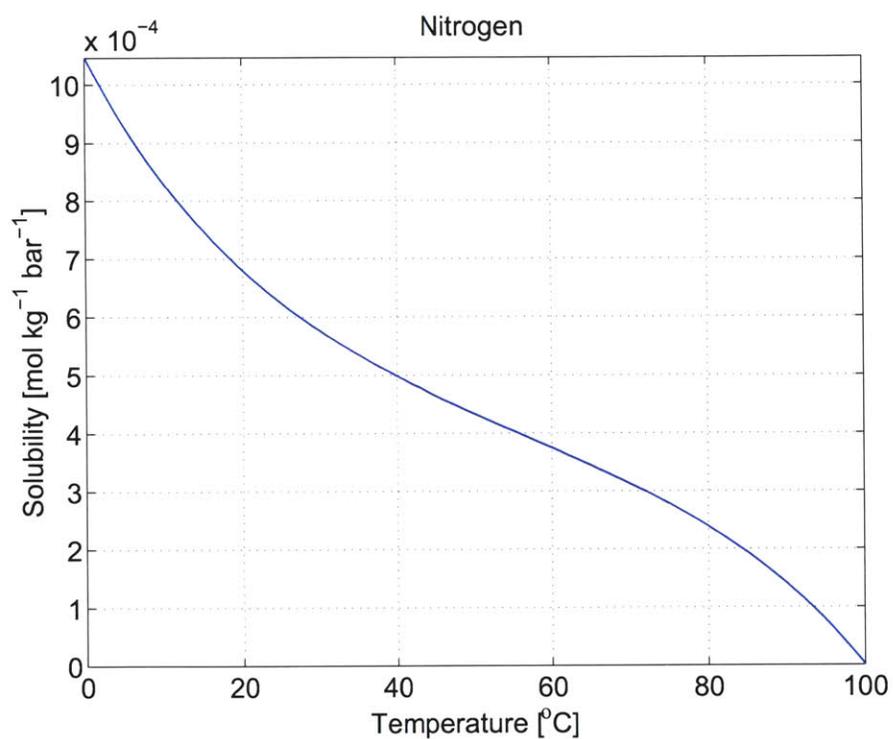


Figure 3.8: Solubility of nitrogen in water: variation with temperature [12]. Data measured at atmospheric pressure (101.3 kPa). At lower pressures, the boiling temperature is decreased and the solubility curve is shifted comensurately.

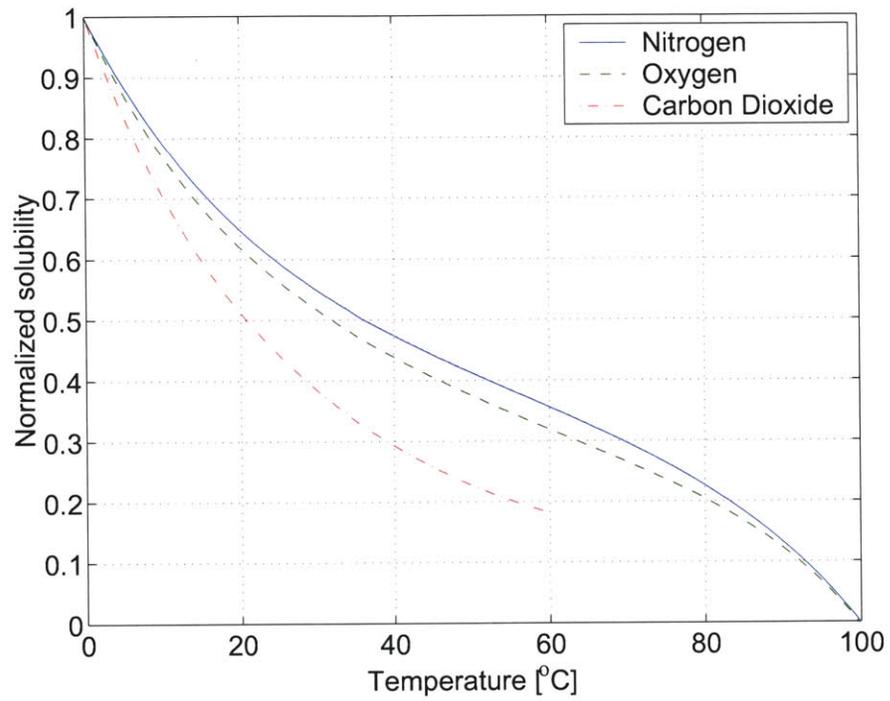


Figure 3.9: Temperature variation of solubility for various gases, normalized by their solubility at 0°C. [12]

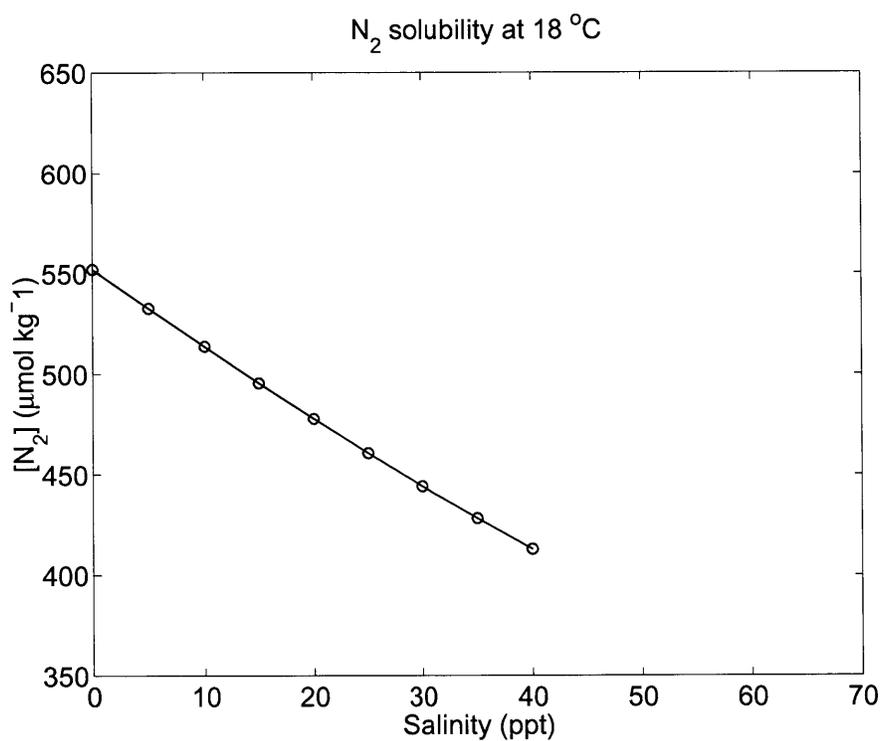


Figure 3.10: Variation of solubility with salinity. The solubility of nitrogen is lower in saline solution than in pure water.

Nitrogen

$\lambda@ 37^{\circ}\text{C}, 1 \text{ atm}$	Nitrogen
0.155M Saline	0.0138
Whole blood	0.015 [0.0135-0.018]
Water	0.0145
Lipids	0.07
Plasma	0.0137

For obvious reasons, data for nitrogen are most relevant here. In water at body temperature, the data are notably consistent that $\lambda \approx 0.0145$ (range 0.0142-0.0146). In whole blood, the data are much more variable, but with a similar mean. In saline and plasma, it appears that the solubility is marginally lower. However, we shall be lucky if our analysis is accurate enough to warrant fretting over more than two significant figures: a consensus value of 0.015 will suffice. We should note also that nitrogen is approximately fivefold more soluble in lipids, although the solubility is still low enough that this is unlikely to be a confounding factor in the analysis.

Other Gases

$\lambda@ 37^{\circ}\text{C}, 1 \text{ atm}$	Carbon Dioxide	Oxygen	Nitrous Oxide
0.155M Saline	0.60	0.025	0.45
Whole blood	*	*	0.46 [0.44-0.48]
Water	0.62	0.027	0.45 [0.43-0.48]
Lipids	1.4	0.13	1.4
Plasma	0.582 (no bicarbonate)	0.0243	0.45

For other gases, we are principally concerned with solubility in water and saline solution. Solubility in physiologic saline appears marginally reduced, but as above, not to an extent that is significant in our analysis. What is most relevant is a broad appreciation of how much more soluble are carbon dioxide and nitrous oxide than nitrogen, by factors of more than 40 and 30 respectively.

3.6 Degassing

Before we attempt to dissolve any radioactive gas, we need to ‘degass’ the saline solution. That is, to remove any gases that may already be dissolved. This is accomplished by heating the saline and then placing it in a vacuum chamber for a duration as described below.

3.6.0.6 Mechanism

The solubility of gases in water decreases with increasing temperature, tending to zero at boiling point. A graph of solubility vs. temperature for nitrogen is shown in figure 3.8. In this case, the solubility is expressed as moles per kilogram per bar: moles of gas dissolved in unit mass of water

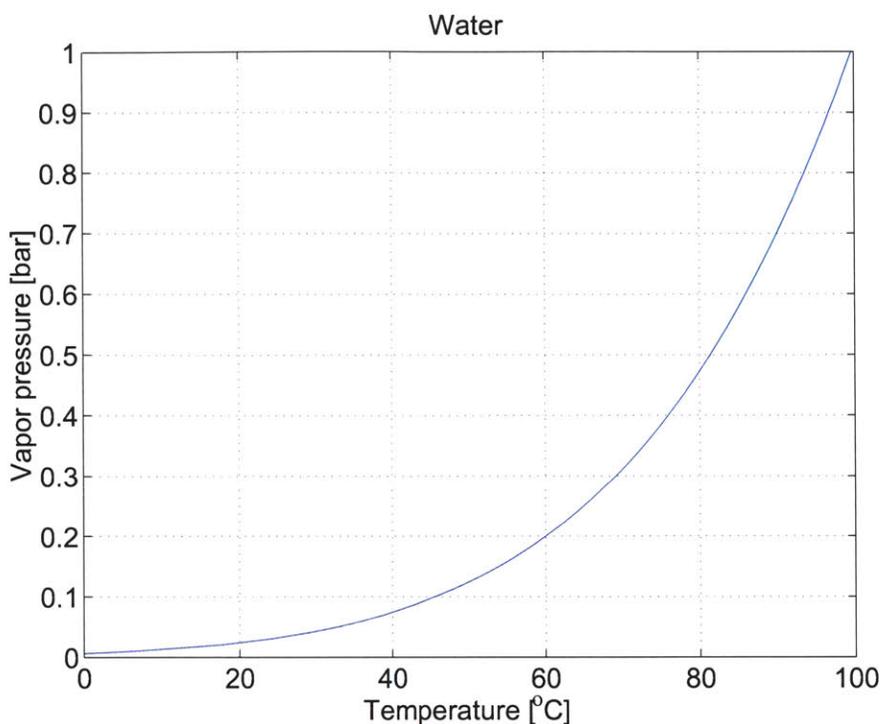


Figure 3.11: Variation of vapor pressure of water with temperature. Raising the temperature of saline from 20 to 60 increases the vapor pressure approximately 8-fold.

Since nitrogen is both the largest component of air, and the principal gas that we are trying to dissolve in saline, I shall limit my discussion to this gas.²

The saline solution is heated in a microwave (to maintain sterility) for six minutes, raising its temperature to approximately 60°C. The warming has two relevant effects. The primary effect is to increase the vapor pressure dramatically (by a factor of ~8). A secondary effect is to lower the solubility of nitrogen to ~55% of its value at 20°C [Figure 3.11]. The saline bag is then placed in a vacuum chamber, which is connected to a vacuum pump [Laboport UN830.3 FTP, KNF Neuberger Inc.]. As the pressure falls, small bubbles of gas form on the surface of the bag. When the pressure within the bag reaches to vapor pressure of the saline, the liquid boils, and the bag inflates. Dissolved gases bubble to the surface and are convected out of the saline bag with the water vapor.

To maintain good degassing, the pressure outside of the bag should be significantly below the vapor pressure, ensuring that there is a steady flow of water vapor out of the bag. If the flow is not maintained, gases will come out of solution, but remain within the bag.

Although at equilibrium dissolved gases should be nil, it is likely that equilibration is not complete,

²The solubility curves of other gases, although not geometrically similar, all possess the same sigmoidal shape.

and at the end of the degassing period, a small quantity of gas remains in solution.

3.7 Sodium Hydroxide

The gas received from the cyclotron is almost entirely (~99.7%) carbon dioxide, with a small amount of other gases. The first step in the preparation process is to remove the carbon dioxide by bubbling the gas mix through aqueous sodium hydroxide.

The absorbing chamber contains approximately 250ml of 5M NaOH solution, which is sufficient for about eight runs before it needs to be replaced. [See appendix A.2.]

3.7.0.7 Change in Volume

Absorption of carbon dioxide by the sodium hydroxide solution increases the volume of the solution. This was noted during testing of the absorbing chamber, when absorbing rate is critically dependent on the pressure-volume relationship.

We estimate that the volume increases by approximately 7ml per 4.5 liters of gaseous carbon dioxide. The volume of cyclotron gas received in each run is typically 1300-1500ml, and thus the increase in absorber volume is around 2ml per run.

This effect is actually useful, as the compliant chamber becomes rigid to negative volumes [section 3.8.1]. When transferring the bubble, the increase in absorber volume enables us to withdraw the bubble from the chamber without the pressure dropping significantly.

Between runs, however, this increased volume of the absorber (~1% of the total) must be vented to ensure repeatable starting conditions, and to avoid progressive accumulation of liquid in the absorbing chamber.

3.7.1 Dissolution of gas in absorber

One of the disadvantages of using a liquid absorber is that at the same time as removing the unwanted carbon dioxide, we dissolve both nitrogen-13 and other gases in the NaOH solution. After receipt is complete, and the bubble has been transferred, significant activity remains in the absorbing chamber due to the dissolution of N13.

3.7.1.1 Analysis

How much gas dissolves into the absorber? We assume that the species in question is present in small concentrations, so that the amount that dissolves (or comes out of solution) does not significantly affect the volume of the gas bubble. This assumption should be valid for N13, where the volume of the bubble is of the order of 5ml, and the volume of N13 is of the order 0.01 μ l. (See section 3.4.1.)

At equilibrium, we expect that

$$\frac{n_g}{n_l} = \frac{V_g}{\lambda V_l} \quad (3.4)$$

where n_g , n_l are the *amount of substance* of the species in the gaseous and liquid phase, respectively. V_g is the volume of the gas bubble, V_l is the volume of the liquid absorber, and λ is the partition coefficient.

Validity and Relevance The above equation is valid only if the system is equilibrated, but in this case equilibration is incomplete. During receipt of the cyclotron gases (~1 min), the gas bubble is large and the pressure in the absorbing chamber is high. After the inlet valve is closed, the bubble volume decreases as stirring in the chamber continues (~15 s). V_g is constant (determined by the capacity of the compliant chamber) during receipt, but equilibrium is not reached as gas is continually entering the absorbing chamber and the nitrogen concentration in the bubble steadily rises. The subsequent stirring is timed so that the bubble reaches a minimum volume, but this does not necessarily indicate that equilibrium is approximated for nitrogen (or other similar species), for which the time-constant is longer than for carbon dioxide.

Nevertheless, the above expression has some utility. Even if we do not know V_g , we observe that to maximize yield, we want this to be large. This suggests we should admit a large volume to the chamber during receipt, and that we should abbreviate subsequent stirring as much as possible. (Ideally, we would stop stirring immediately that the inlet valve is closed, and transfer a large bubble of mostly carbon dioxide.) Similarly, we want V_l to be small: we should minimize the volume of NaOH solution in the absorbing chamber. Finally, we want the partition coefficient to be as small as possible.

Of particular interest is the fact that pressure does not appear in equation 3.4. Intuitively, we might expect that raising the pressure, by increasing the partial pressure of N13, will increase the amount that dissolves in the absorber. However, if V_g remains constant, the concentration of N13 in the bubble is compensatorily reduced, and the ratio n_g/n_l does not change. There is thus no advantage to receiving the cyclotron gases at low pressures. Higher pressures, on the other hand, will increase the rate at which carbon dioxide is absorbed, so that we can receive more quickly, losing less activity due to radioactive decay.

3.7.1.2 Optimization of absorbing process

Modification of Partition Coefficient Can we lower the partition coefficient in order to improve yield? Equation 3.2 suggests that we may be able to change λ by changing the temperature of the absorber. However, since λ increases with T , and s decreases with T , it is not immediately obvious whether we should attempt to heat or cool.

Differentiating equation 3.2 w.r.t. T , we get:

$$\frac{\partial \lambda}{\partial T} = R_0 \rho_l \left(T \frac{\partial s}{\partial T} + s \right)$$

Using appropriate values for these quantities, we find that the bracketed term is negative (so that we need to *increase* temperature to lower λ), and that $\partial s/\partial T \sim 0.0001 K^{-1}$.

Realistically, we could elevate the absorber temperature by about 40-50 °C before we run into boiling-related concerns. With this sort of ΔT , we would expect a decrease in λ of around 30%. However, this does not mean that our yield would improve by the same proportion. Equation 3.2 is written in terms of n_g/n_l , but a more useful formulation for describing yield is to cast the equation in terms of n_g (the amount of N13 in the bubble, which is transferred and dissolved) as a fraction of the *total* amount of N13.

$$\frac{n_g}{n_g + n_l} = \frac{V_g}{\lambda V_l + V_g}$$

Now, V_g is approximately an order of magnitude larger than λV_l . Consequently, a 30% decrease in λ results in a more modest improvement in yield, for example from 90% to 93%.

Based on this analysis, we opted not to install heating apparatus for the absorbing chamber.

3.7.1.3 Modification of $V_g : V_l$ ratio

The other obvious approach to decrease the amount of activity lost to the absorber is to increase the ratio $V_g : V_l$.

A first step that we took was to decrease V_l , the volume of sodium hydroxide solution. In the original design of the absorbing chamber, this volume was approximately 500ml. We fabricated an insert which decreased the volume to 250ml, while retaining the existing design. This improves the $V_g : V_l$ ratio by a factor of two. The current volume provides capacity for approximately ten runs before the sodium hydroxide needs to be replaced.

A second step was to replace the compliant chamber. This was done for a number of reasons: principal among them were reliability issues. The redesign, however, allowed us to increase the compliant capacity (and thus V_g) from ~25ml to 60ml. This increase also reduces *effective* V_l because sodium hydroxide solution is sequestered in the compliant chamber during receipt and does not participate in gas exchange. $V_g : V_l$ is again improved by a factor of 3.8. Note that improvement only applies to the phase where the membrane of the compliant chamber is fully distended.

3.7.1.4 Retention of dissolved gases

When we receive gas from the cyclotron and pass it through the absorber, gas dissolves in the liquid, and remains in solution at the time of the next run. We need thus to consider the state of the absorber and its history.

Analysis Let us start by considering a pristine absorber, in which no gases are initially dissolved. Before cyclotron bombardment, we add a small volume, V_1 , of some additional gas to the target. After receipt of the cyclotron gas, a fraction of this gas dissolves in the absorber. The fraction that dissolves, x , is approximately constant, provided that the bubble size, V_g , is constant. This condition

is not satisfied if we add a large volume of gas, as the undissolved part will significantly increase V_g .

$$x = \frac{n_l}{n_l + n_g} = \frac{1}{1 + \left(\frac{V_g}{V_l \lambda}\right)}$$

The additive gas is redistributed as follows.

$$V_1 \begin{cases} (1-x)V_1 & \text{in gas phase} \\ xV_1 & \text{in liquid phase} \end{cases}$$

The gaseous part, $(1-x)V_1$, is carried out of the chamber, leaving xV_1 in solution.

Before the second batch of tracer is prepared, we add V_2 to the target. After receipt, we have a total volume of $(V_2 + xV_1)$ of nitrogen gas. This then redistributes as

$$(V_2 + xV_1) \begin{cases} (1-x)(V_2 + xV_1) & \text{in gas phase} \\ x(V_2 + xV_1) & \text{in liquid phase} \end{cases}$$

The n^{th} run thus produces a gas bubble containing

$$(1-x)V_n + (1-x)xV_{(n-1)} \dots + (1-x)x^i V_{(n-i)} \dots + (1-x)x^{(n-1)}V_1$$

If the same amount of gas is added to the target each time, $\forall i : V_i = V_N$, then this is a simple geometric series which converges to V_N . That is, the system reaches a steady-state in which the amount of extra gas in the bubble equals the amount added to the target.

Quantification Based on measurements of the residual activity of the sodium hydroxide solution [appendix A.3], x for N13 (and with current absorbing chamber configuration) is approximately 0.13. Using this number, the bubble for the n^{th} run has the following additional content:

$$0.87V_n + 0.11V_{(n-1)} + 0.015V_{(n-2)} + \dots$$

The term for the previous, $(n-1)^{\text{th}}$, run is significant, but all others prior to this are negligible.

For more soluble gases than nitrogen, x is larger, for example, ~ 0.83 for nitrous oxide. In this case, the terms decrease much more slowly:

$$0.17V_n + 0.14V_{(n-1)} + 0.117V_{(n-2)} + \dots$$

Significance The importance of this analysis is twofold. Firstly, during test-runs, for example to determine the effect of target additives, in which the volume added is varied, we observe that we need to account for the history of the absorber.

Secondly, this analysis allows us to understand what happens to gases initially dissolved in the absorber before it is used. The sodium hydroxide solution is mixed using water that has equilibrated

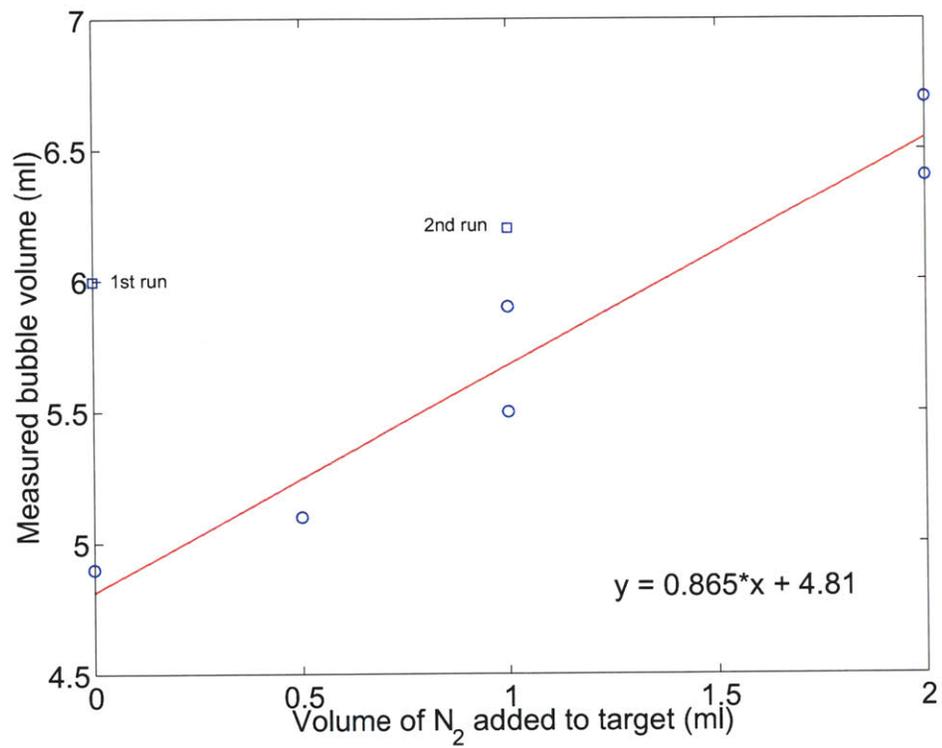


Figure 3.12: Variation in measured bubble volume with the amount of nitrogen used to prime the target. Of particular interest is the fact that the first (and to a lesser extent, the second) run has very different bubble volume due to atmospheric gas initially dissolved in the absorber.

with the atmosphere and thus has significant amounts of nitrogen dissolved in it. This gas evolves out of solution during bubble receipt, and causes readily measurable increases in the bubble volume, as shown in figure 3.12.

The linear regression from this empirical data suggests that around 13.5% of the nitrogen added to the target dissolves in the absorber. We might reasonably presume that this number is also typical of nitrogen-13. This figure is approximately consistent with the measured values detailed in appendix A.3.

3.8 Design of Absorbing Chamber

The absorbing chamber is constructed out of PMMA, and is mounted directly on top of a magnetic stirrer. Since the chamber is essentially rigid, it needs to be connected to a compliance to allow a significant volume of gas to be introduced. This compliance of the chamber is provided by a separate compartment in which an elastic sheet contains the sodium hydroxide solution, but distends under pressure.

As shown in 2.9, the original design of the absorbing chamber was modified with an annular insert to reduce the volume of sodium hydroxide solution from 500 ml to 250 ml. This modification decreases the amount of nitrogen-13 which is lost by dissolution in the sodium hydroxide solution. (See section 3.7.1 for details.)

Gas from the cyclotron is pumped to the apparatus along a long piece of tubing enclosing a volume of approximately 1100ml. During receipt of the cyclotron gas, the amount of radioactivity in the line is monitored; when the radioactive gas arrives, the inlet valve is opened, admitting gas to the chamber. Once the elastic sheet in the compliant chamber has distended significantly, the mixture is stirred magnetically. (We wait for a large bubble to form before stirring to minimize the chances of splashed liquid droplets entering the tubing at the top of the chamber, and prematurely tripping the liquid detector.) At this stage, the flow-rate into the chamber (typically around 2 l/min) is limited by the rate at which the sodium hydroxide can absorb carbon dioxide. After the radiation level in the line starts to fall, the inlet valve is closed, and stirring continues (for ~15 seconds) until pressure in the chamber falls to a stable level, slightly elevated above the initial pressure, reflecting the presence of a bubble of unreactive gas.

3.8.1 Compliant Chamber

The compliant chamber is designed with a spherical surface on the reverse side of the elastic sheet, so that as the elastic sheet distends, it simultaneously contacts the entire outer surface simultaneously. At this point the chamber becomes effectively rigid.

The reverse side of the compliant chamber is ported, and connected to an air-filled reservoir of fixed volume that is pressurized slightly. The backside pressure (~100cm H₂O) is too small to significantly affect the absorbing process, but when the system is depressurized, presses the elastic membrane flat against the inner wall. This allows us to achieve repeatable conditions in the absorbing chamber following a preparation sequence.

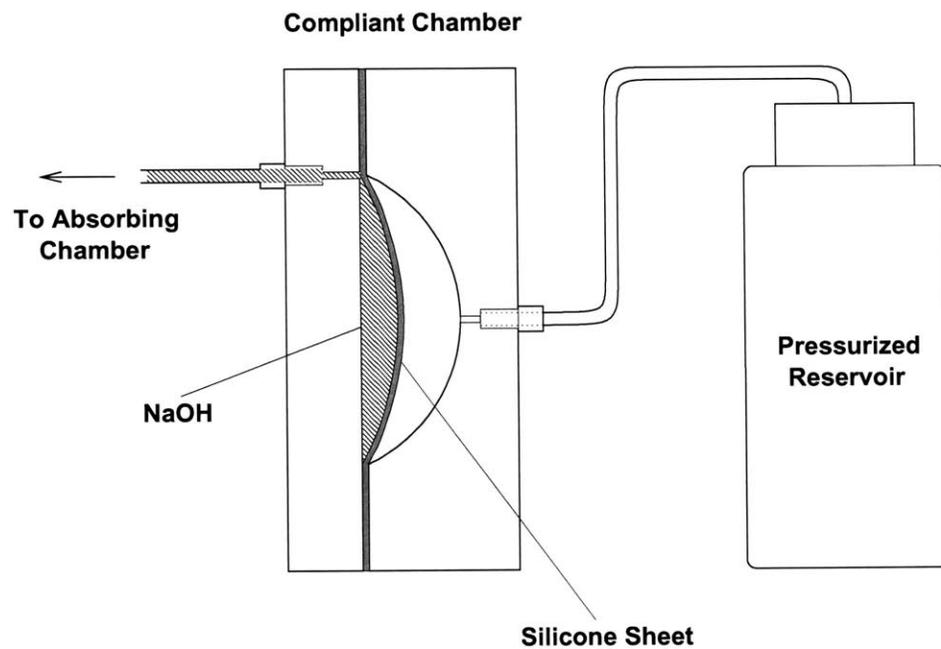


Figure 3.13: Detail of compliant chamber (not to scale). The second part of the chamber is connected to a larger (1000ml) reservoir, which is pressurized to 100cm of water (1.5psi). This provides a small pressure on the reverse side of the elastic sheet, ensuring that the compliant chamber empties fully.

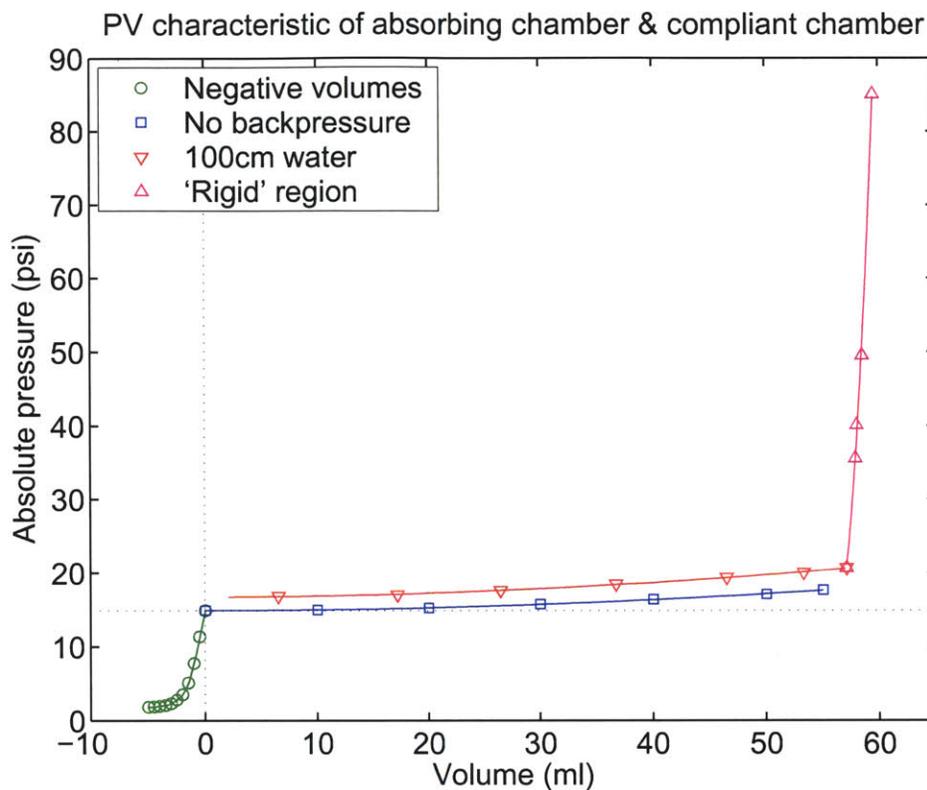


Figure 3.14: Pressure-volume characteristic of the absorbing chamber (while connected to the compliant chamber). Initially ($V=0$), the elastic membrane lies flat against its inner wall; as fluid is admitted to the absorbing chamber, the sheet is stretched outward

A second important effect of the fixed reservoir volume is that the pressure within it increases as the elastic sheet distends, and falls as volume is withdrawn from the compliant chamber. At the point where the inlet valve is closed, the volume of the bubble in the absorbing chamber decreases as the remaining carbon dioxide is absorbed. By monitoring the pressure on the back side of the sheet, we can estimate the volume of the bubble that remains, and assess when the bubble has fallen to a minimum volume, at which point stirring should stop and the bubble should be transferred. Without this mechanism, we would continue stirring unnecessarily, which decreases the tracer yield by dissolving more nitrogen-13 in the absorber.

3.8.2 Pressure-Volume Characteristic

After a preparation sequence, the absorbing chamber is filled with sodium hydroxide solution up to the level of the exhaust port. The volume above is filled with flush gas, and the pressure is

equilibrated with the atmosphere. The compliant chamber is at minimum volume, with the elastic sheet lying flat against its inner wall.

To measure the pressure-volume characteristic of the system, however, the chamber was completely filled with liquid. Gas was excluded from the system in order to avoid artefacts due to the compressibility of the gas. Volume changes were recorded relative to a reference point with the compliant chamber empty and the system at atmospheric pressure.

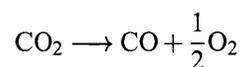
When liquid is admitted to the system, the elastic sheet distends easily, and the system is very compliant: this is the almost flat region in the middle of figure 2.9. When the elastic sheet reaches its outer wall, the system becomes suddenly rigid – any slight compliance is due to the imperfect rigidity of the walls and interconnecting tubing. Thus the PV curve abruptly turns steeply upward at $V \sim 58\text{ml}$.

Two sets of positive volume data are shown in figure 2.9, one with the reverse side of the compliant chamber open to atmosphere, and a second with the port connected to a reservoir pressurized to 100cm H₂O. This latter characteristic reflects the behavior of the system during normal use.

Since at $V = 0$ the elastic sheet lies flat against the inner wall, the system becomes almost immediately semi-rigid to negative volumes, and the pressure drops very rapidly. In this region of figure 2.9, the P-V relationship is seen to curve, apparently approaching an asymptote around 1 psia. This does not reflect the ‘true’ P-V characteristic, but is an artefact of the experimental methodology. The liquid used for the test was incompletely degassed, and at low pressures, dissolved gas evolves out of solution, increasing the effective volume of the system, so that withdrawal of liquid does not decrease the actual volume remaining in the chamber. (Even if the liquid had been completely degassed, we would have observed a similar effect when the pressure dropped to the vapor pressure, forming cavities of water vapor.) The unexpected curvature (of the measured P-V relationship) at pressures above the partial pressure of dissolved gas is mostly likely due to the existence of very small residual pockets of gas that were not excluded from the system.

3.9 Carbon Monoxide catalyst

One side-effect of the cyclotron bombardment of carbon dioxide is the degradation of carbon dioxide into carbon monoxide and oxygen.



Although carbon monoxide is a poison, the quantities involved are too slight for this to be a concern. The significance of this degradation is that neither carbon monoxide nor oxygen react with sodium hydroxide. Since neither are removed in the absorbing chamber, they are transferred and dissolved as part of the bubble, where their relatively poor solubility increases the risk of creating a super-saturated solution.

After a 20 min bombardment at beam current of $30 \mu\text{A}$, the concentration of CO in the cyclotron gas is approximately 0.1%.³ Since the volume of the target (when decompressed to atmospheric pressure) is $\sim 1500\text{ml}$, the volume of carbon monoxide that we expect to see in the bubble is $\sim 1.5\text{ml}$, accompanied by $\sim 0.75\text{ml}$ of oxygen.

Early systems for preparation of N13 solution oxidized the CO by piping the flow from the cyclotron through a furnace filled with copper beads [7]. This approach, although effective, is inconvenient, and not well-suited to clinical application. Recently, a catalyst has been developed for use in high energy carbon-dioxide lasers [REF], which is capable of recombining CO and O₂ at room temperature.

We packaged some of this catalyst [STC Catalysts, Inc] in a cylindrical enclosure, installed in the rear compartment of the main unit. In tests, the catalyst has performed admirably: CO recombination was complete up to a maximum tested flow of 2.5 l/min. These tests are detailed in appendix A.1.

Note that the catalytically-induced recombination requires a stoichiometric mixture of reactants. If oxygen were removed from the system, for example by reacting with the walls of the cyclotron target chamber, the resultant excess of CO would not be removed. Our measurements, however, suggest that this effect is negligible.

3.10 Cyclotron Preparation

The following procedure is used for preparation of nitrogen-13 in the hospital cyclotron. An aluminium target chamber (of 250 ml volume) is filled with high purity carbon dioxide gas at 92 psi (6.3 atm). The target gas is then bombarded with deuterons, at a current of $30 \mu\text{A}$, through a thin window in the chamber. During the bombardment, the pressure in the target chamber rises to approximately 114 psi.

3.10.1 Duration of bombardment

The duration of the irradiation is not fixed. Typically, we aim to have the cyclotron bombard for twenty to thirty minutes. In practice, the timing is usually dictated by the study being performed: experimental protocol might call for imaging at, say, thirty minutes after an intervention; we might need to wait for a subject's condition to stabilize; or imaging may need to wait for other measurements to be completed.

There is a constant probability of a deuteron-carbon interaction resulting in formation of nitrogen-13, so the total amount of nitrogen generated will be proportional to the product of the beam current and the duration of irradiation. However, due to the radioactive decay of the nitrogen-13, its concentration does not rise continuously, but approaches a steady-state level.

³At low concentrations, the amount of carbon monoxide produced increases linearly with integrated cyclotron beam current [5].

3.10.1.1 Analysis

We assume that rate of ^{13}N production is proportional to beam current, I . The rate of radioactive decay is A/τ , where τ is the time constant of the decay, and A is the amount of nitrogen-13.

This gives us the differential equation:

$$\frac{dA}{dt} = kI - \frac{A}{\tau}$$

Applying the boundary condition that $A = 0$ at $t = 0$, we get:

$$A = kI\tau \left(1 - e^{-\frac{t}{\tau}}\right)$$

This result implies that the total activity within the target asymptotically approaches $kI\tau$. Since τ is constant, we can improve this only by increasing the beam current or by increasing the likelihood of a generative interaction.

The rate at which we approach this asymptote is determined entirely by the half-life of the isotope, and cannot be improved. Since the half-life of nitrogen-13, conveniently, is almost exactly ten minutes, we expect that a ten minute bombardment should produce half the ultimate activity; a twenty minute bombardment should produce three quarters; a thirty minute bombardment seven eighths, and so on.

3.10.2 Target Priming

At one stage in our studies, we observed a sudden and consistent drop in yield. Investigation finally tracked the cause down to the replacement by the cyclotron staff of the carbon dioxide cylinder that they were using to fill the target. Although the original cylinder had been returned and was could not be specifically identified, we hypothesized that the newer cylinder might be of a higher purity, and that impurities in the original tank were somehow improving the production of nitrogen-13.

We experimented with adding oxygen, water, and nitrogen to the target gas, eventually concluding that only the latter was useful. Injecting 1ml of nitrogen into the target gas was observed to produce an increase of greater than 50% in the specific activity of the gas received from the cyclotron. [Full results of these tests are presented in appendix A.5.]

This is a controversial finding. Physical theory predicts that the likelihood of a generative interaction is constant, and thus the total amount of radioactive material produced should be proportional to the integrated beam current, and not affected by the very small concentrations of contaminants introduced.

One possible mechanism of this improvement is as follows. Nitrogen-13 is generated in the cyclotron as single atoms, which are unstable until they associate with another to form diatomic molecules of nitrogen. Since the concentration is slight, the chance of the nitrogen-13 atoms associating with each other is small, and instead the nitrogen-13 bonds, undesirably, with other species

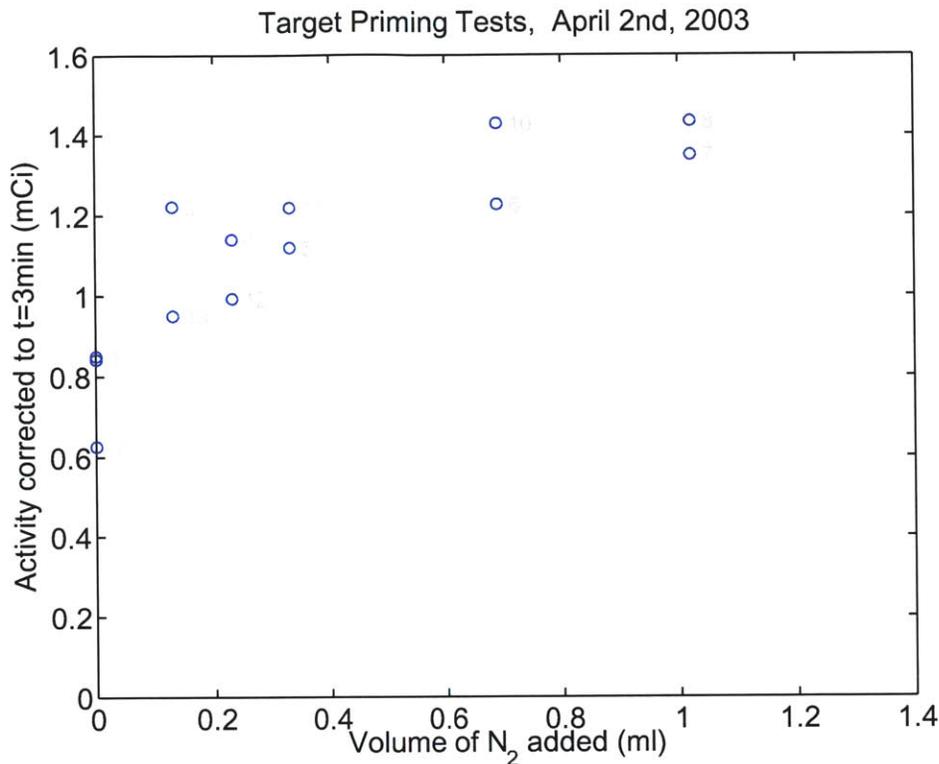


Figure 3.15: Preliminary results of target priming tests.

in the target chamber (possibly with the walls of the chamber itself). By increasing the nitrogen content of the target, we increase the formation of molecular nitrogen-13, presumably in the form of $^{13}\text{N} - ^{14}\text{N}$.

Although the addition of nitrogen to the target gas causes a monotonic increase in the generated activity, this effect plateaus relatively quickly [Figure 3.15]. Adding large quantities of nitrogen is deleterious to the system yield, as the total amount of this gas that can be dissolved is very limited. (Around 1ml nitrogen in 60ml of saline at NTP.) In practice, we have settled on a compromise of priming the target with 0.33ml of N₂. This is large enough to significantly increase the activity coming from the cyclotron, but without introducing a destructive quantity of relatively insoluble gas.

3.11 Infusion

To maximize the total amount of activity that we introduce into the subject, we want both the specific activity of the tracer solution, and the volume infused to be as high as possible.

The infused volume is limited by the maximum flow rate which is tolerable, and the maximum time that is deemed reasonable.

Maximum infusion rate is of the order of 5-10ml/s. Several subjects have complained of discomfort during the infusion, but it is unclear how dependent this is on the injection rate. Several subjects have tolerated the first infusion well but complained subsequently; it may be that a 'sensitization' phenomenon is occurring. An alternative explanation is that the temperature of the saline is dropping during the course of the study, since the bag is sitting in an ice bath. The discomfort may be alleviated by (i) the use of local anesthesia, (ii) maintaining a flow of saline, either continuously or shortly prior to infusion, (iii) encouraging the subject to move his/her hand and arm around to increase perfusion.

Maximum acceptable duration of infusion is around 10s. We need to strike a compromise between how long a subject is able to breath-hold, and how much activity reaches the lung during the apnea.

Alternatively, assuming that we will inject as much activity as we can,

Activity Injected = Activity dissolved in Saline - Activity wasted in flushing System.

=Activity transferred(Mixing volume - Flush volume).

Activity wasted in flushing system is typically 17ml (two 4ml flushes; 2ml to sample syringe and dumped; 1ml sample; 6ml final flush immediately before infusion.) One way of decreasing this is to decrease the volume of the lines: discussion required. An important observation here is that subsequent samples are typically something like 10% hotter than the first sample. Clearly, as we flush more liquid through the lines, our sample activity asymptotically approaches the 'true' activity of the salsa. Is this a predictable relationship? If so, we might be able to get away with flushing less, and scaling up the sample activity. Analysis required.

Maximum volume infusible = 25 to 100ml.

As we increase this volume, we decrease the proportion lost to flushing, and we increase the volume of gas that we can dissolve.

3.12 Passive Syringe

The passive syringe provides a mechanism to dissolve the bubble of radioactive gas. Once the bubble has been transferred to the injector syringe, the gas/liquid mixture is mixed by vigorous ejection into and withdrawal from the passive syringe.

In order to prevent the pressure from dropping below atmospheric during the withdrawal (which would result in evolution of gas out of solution), the passive syringe is mounted on a spring-loaded holder. This also accelerates the dissolution by increasing the mean mixing pressure.

The passive syringe holder incorporates micro-switches which indicate when the syringe is at its extremes of movement. The switch at the lower end is a safety feature that prevents potential over-filling and explosion. The upper end switch is used after mixing to determine when the syringe is empty. (We cannot just withdraw the starting volume, because the total volume of the mixture has decreased by an unknown amount.)

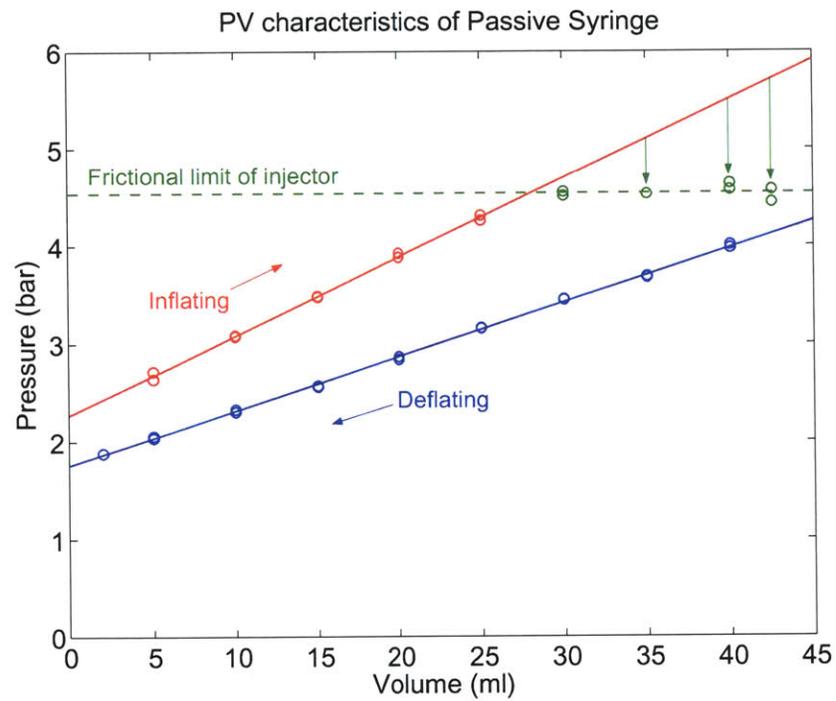


Figure 3.16: Pressure-volume characteristics of the passive syringe. Due to friction, the inflating and deflating limbs are not identical. Measurements were made with the injector non-energized: at high pressures, the injector fell back when power was removed.

3.12.1 Bubble vol/surface area/pressure calculation.

The dissolution (and evolution) of gas mixtures is a complex and extensively-studied field. However, it is worth doing some simplified modeling to understand the factors involved. The first and greatest assumption we make is that the bubble is pure, that it is wholly comprised of a single species.

We assume that the total amount of gas, n_g , dissolves at rate proportional to the product of the surface area, A , and the difference in partial pressure between the gas and the liquid.

$$\frac{dn_g}{dt} = -kA(P_g - P_l) \quad (3.5)$$

The mean partial pressure in the liquid⁴, P_l , may be expressed in terms of the solubility, s , the volume of liquid, V_l , and the (Henry's constant) solubility, s .

$$P_l = \frac{n_l}{sV_l} = \frac{(N - n_g)}{sV_l} \quad (3.6)$$

If the total volume of gas, V_g , is broken into m spherical volumes, then the total area available is given by

$$A = 3\sqrt{36\pi}m^{\frac{1}{3}}V_g^{\frac{2}{3}} \quad (3.7)$$

Particularly as the bubbles become small, we also need to consider the surface tension of the liquid, since this increases the pressure in the bubbles, P_g , above the ambient pressure, P_∞ .

$$P_g = P_\infty + \frac{\sigma}{2r} = P_\infty + 3\sqrt{\frac{m\pi}{6V_g}} \quad (3.8)$$

Combining these expressions gives a differential equation for n_g .

$$\frac{dn_g}{dt} = -k\left(\left(P_\infty + 3\sqrt{\frac{m\pi}{6V_g}}\right) - \left(\frac{N - n_g}{sV_l}\right)\right)\left(3\sqrt{36\pi}m^{\frac{1}{3}}V_g^{\frac{2}{3}}\right) \quad (3.9)$$

Before we can solve this, we also to substitute for V_g using the ideal gas law:

$$P_g V_g = n_g R_0 T = \left(P_\infty + 3\sqrt{\frac{m\pi}{6V_g}}\right) V_g \quad (3.10)$$

Note that this is effectively a cubic equation in $V_g^{\frac{1}{3}}$. Solving this and substituting into equation 3.9 gives rise to an differential equation that cannot be solved analytically.

We can at least use a simplified form of equation 3.9 to consider the initial rate of dissolution. Here the amount of dissolved gas, n_l , is zero. Discounting additionally the effect of surface tension, then

$$\frac{dn_g}{dt} = -k(P_\infty - 0)\left(3\sqrt{36\pi}m^{\frac{1}{3}}\left(\frac{n_g R_0 T}{P_\infty}\right)^{\frac{2}{3}}\right) = -k'\left(n_g P_\infty^{\frac{1}{3}} T^{\frac{2}{3}}\right) \quad (3.11)$$

⁴Strictly speaking, the concept of a partial pressures is not applicable to liquids. However, here we follow the common practice among physiologists to extend this concept, defining the partial pressure of gas in a liquid as the partial pressure of that gas in an atmosphere in communication with that liquid at equilibrium.

An important question that arose during development is whether increasing the pressure during mixing is beneficial. At first sight this seems trivial: more pressure should cause the gas to dissolve faster. However, increasing pressure decreases the surface area of the gas bubbles, decreasing the area available for mass transfer. If the temperature is held constant, then we see from the above equation that increasing pressure results in a modest improvement in the initial rate of dissolution.

3.13 Performance History

Figure 3.17 shows some global measures of how the preparation system has performed over its lifetime. As is evident from this data, the yield of the system is highly variable. Apart from intrinsic variation in the preparation process, the variation seen can be directly attributed due to a number of factors.

Firstly, there is considerable variation in the amount of radioactivity in the gas we receive from the cyclotron. This depends principally on the duration of bombardment (see section 3.10.1), the amount of nitrogen added to the target (section 3.10.2), and the beam current used. Typically, the cyclotron will bombard continuously until we are ready to receive the gas, and this is highly dependent on the study being performed: for example, we might need to wait until an animal has stabilized after an insult, until a transmission scan has finished, or until another measurement has been made. The amount of nitrogen added to the target has also changed substantially as we have experimented with different levels. Furthermore, early apparatus to perform the target priming had poor repeatability, and techniques varied with different cyclotron operators.

Secondly, limitations on the total amount of activity that we can inject have affected the volume and concentration of the tracer prepared. Some of the early animal studies were performed in an old, single-slice camera in which the detectors were easily saturated by large amounts of radiation. Thus we did not attempt to generate a lot of tracer. In early human studies, the experimental protocols approved by the hospital stipulated a dose per infusion that was much lower (10 mCi) than currently used (30 mCi). It also took some time for us to realize the full importance of infusing large doses in order to maximize image quality. Consequently the volume infused, as well as the specific activity, has also increased.

Thirdly, many incremental improvements have been made to the tracer-preparation apparatus. Principal among these are changes to the nature and layout of the front-panel components, the switch to a higher-volume compliant chamber, and the installation of the injector tapping mechanism. There have also been many smaller enhancements to the operational sequence, timing, and other parameters.

Finally, these data also reflect experimental and/or operator error (for example, the unusually high tracer concentration measured in mid-2000, was almost certainly a bad sample), and malfunction of the preparation system (the abnormally low volumes at the start of 2003, for example, are infusions that were arrested prematurely because bubbles were detected).

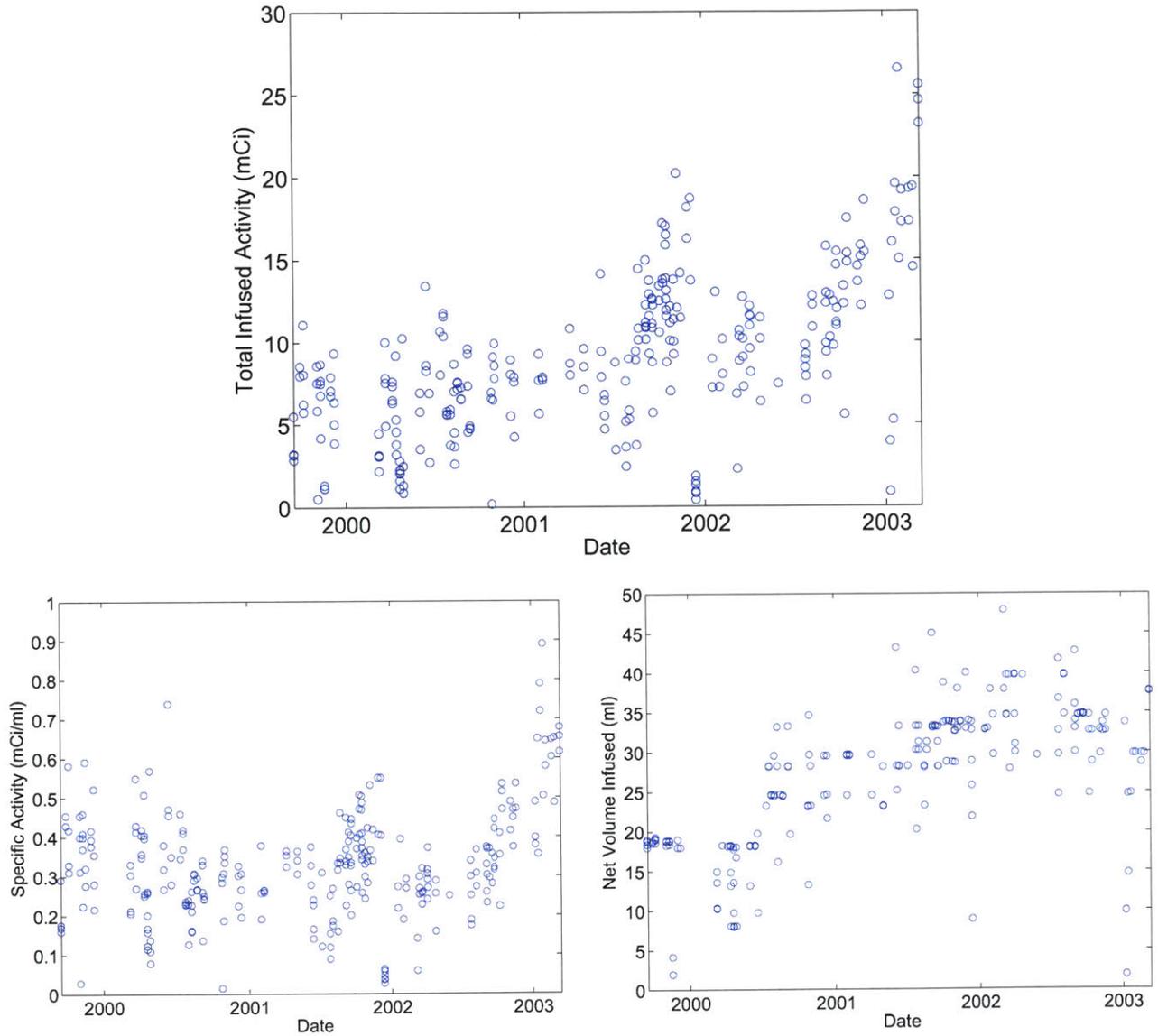


Figure 3.17: Performance history of the system. A number of enhancements to the system have resulted in gradually increasing amounts of radioactivity infused [top]. The improvement comes partly due to improvements in specific activity [bottom left], and partly due to increases in the volume infused [bottom right].

Chapter 4

Lung Model

In this chapter, a theoretical model is developed of the kinetics of the tracer in the heart and lungs. All of the subsequent analysis will be based on this description of the system.

The system is essentially linear, and can be described conveniently using transfer functions and analyzed using conventional methods.

4.1 Model Formulation

We start by modeling the infusion of the tracer itself, which will be the input to the system. To describe the passage through the heart and to the lungs, we model the right heart, vena cava, and pulmonary artery as a well-mixed chamber in combination with a delay. Next, we model the gas-exchange in the lung, assuming that the tracer in the blood equilibrates with a gas-filled volume.

4.1.1 Bolus Infusion

In human studies, the tracer is injected into a peripheral vein at a constant flow-rate. Typically, we inject around 30 ml of tracer at a concentration of ~ 0.5 mCi/ml, and at flow-rate of 5 ml/s.¹

The infusion is modeled as a step-up to a constant mass flow-rate at time $t = 0$, followed by step-down at time $t = t_{inf}$. The step height is A_{inf}/t_{inf} , so that the bolus has area equal to the total amount of activity infused, A_{inf} .

$$C_{inf}\dot{Q}_{inf}(s) = \left(\frac{A_{inf}}{t_{inf}}\right) \frac{1}{s} (1 - e^{-st_{inf}}) \quad (4.1)$$

It is convenient to consider the product of infused tracer concentration, C_{inf} , and flow-rate, \dot{Q}_{inf} , as a single variable, as it does not matter whether we inject a large volume at low concentration, or a small volume at high concentration.

¹In animal studies we inject similar volumes and concentrations, but through a central venous catheter which allows us to inject at higher flow-rates (10 ml/s) and closer to the heart. Consequently, tracer reaches the lung more rapidly, and we are able to extract more information from the apnea period.

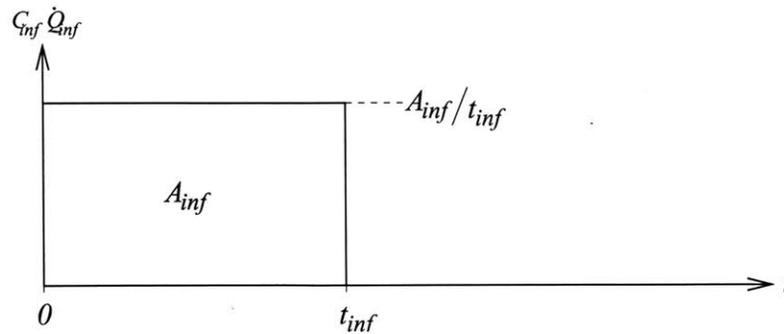


Figure 4.1: Tracer is infused as a bolus. The input to the system is thus modeled as a step-up to constant mass flow-rate at $t = 0$, followed by a step-down at $t = t_{inf}$. The area of the bolus is equal to the total amount of activity infused, A_{inf} .

4.1.2 Heart Model

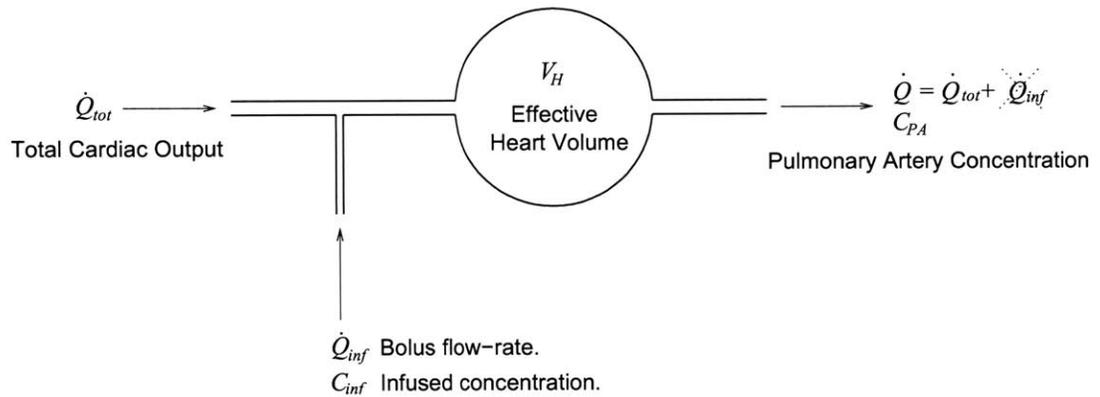


Figure 4.2: Model of right heart. Venous blood, with flow-rate \dot{Q}_{tot} and zero tracer concentration, mixes with tracer infused at rate \dot{Q}_{inf} and concentration C_{inf} . The right heart is considered to be a well-mixed chamber with volume V_H . Blood exits into the pulmonary artery with concentration C_{PA} equal to the concentration in the heart.

Venous blood, in which the tracer concentration is assumed to be zero, arrives with flow-rate equal to the total cardiac output, \dot{Q}_{tot} . This mixes with tracer infused at rate \dot{Q}_{inf} and concentration C_{inf} . We model the right heart itself as a well-mixed (i.e. concentration is uniform) chamber with volume V_H . Tracer concentration within the heart is equal to C_{PA} , the concentration at which blood exits into the pulmonary artery. [Figure 4.2]

Rate of change of mass of tracer = mass flow rate entering - mass flow leaving:

$$V_H \frac{dC_{PA}}{dt} = \dot{Q}_{inf} C_{inf} - (\dot{Q}_{tot} + \dot{Q}_{inf}) C_{PA} \quad (4.2)$$

If $\dot{Q}_{tot} \gg \dot{Q}_{inf}$, i.e. if injected volume is small compared to cardiac output, then we can linearize this equation, and describe the heart by a transfer function. For notational convenience, we group the terms $1/\dot{Q}_{tot}$ and V_H/\dot{Q}_{tot} into constants B and C .

$$\frac{C_{PA}(s)}{C_{inf}\dot{Q}_{inf}(s)} = \frac{1}{V_H s + \dot{Q}_{tot}} = \frac{(1/\dot{Q}_{tot})}{(V_H/\dot{Q}_{tot})s + 1} = \frac{B}{Cs + 1} \quad (4.3)$$

4.1.3 Convective Delay

There is a delay between the start of tracer infusion and its appearance in the heart, due to the time taken for the tracer to be convected through the veins. Consequently, we need to include a time-shift term, e^{-st_c} .

Mixing and diffusional artifacts of the convection are modeled implicitly in the V_H term, which describes the effective mixing volume of the heart and great vessels.

4.1.4 Aerated Lung

Referring to figure 4.3, we model the gas-exchange in an aerated lung unit as follows.

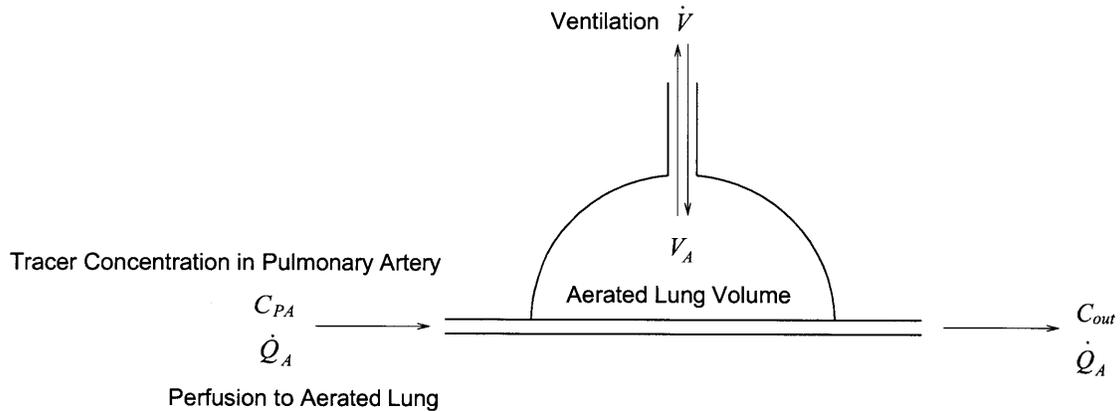


Figure 4.3: Model of aerated lung unit. Blood enters from the pulmonary artery with tracer concentration C_{PA} and flow-rate \dot{Q}_A . Gas exchange takes place within a lung unit of aerated volume V_A and mean tracer concentration C_A , that receives steady ventilation of \dot{V} . End-capillary blood leaves with concentration C_{out} .

Blood arrives with pulmonary artery concentration C_{PA} . The perfusion to the the unit is \dot{Q}_A . This blood is in communication with an inflated volume, V_A . The lung unit is tidally ventilated at a rate equivalent to a continuous flow of \dot{V} of gas at inspired concentration. Blood leaves with concentration C_{out} , and (by continuity) flow-rate \dot{Q}_A .

We assume that the inspired concentration is zero, and that the volume of the lung unit is constant. Conservation of mass requires that:

$$V_A \frac{dC_A}{dt} = \dot{Q}_A (C_{PA} - C_{out}) - \dot{V} C_A \quad (4.4)$$

If blood gases have equilibrated with alveolar airspace, then $C_{out} = \lambda_A C_A$. Here λ_A denotes the partition coefficient for nitrogen at body temperature. [See section 3.5.0.2 for an explanation of this quantity.]

$$V_A \frac{dC_A}{dt} = \dot{Q}_A (C_{PA} - \lambda_A C_A) - \dot{V} C_A \quad (4.5)$$

The system may be considered to have two time-varying inputs, C_{PA} , and \dot{V} . The system is linear in C_{PA} , but not in \dot{V} . However, since ventilation is zero during the apnea and then steady during the washout, we can treat this as piecewise constant, and describe the lung with a single input transfer function:

$$\frac{C_A(s)}{C_{PA}(s)} = \frac{\dot{Q}_A}{V_A s + (\dot{V} + \lambda_A \dot{Q}_A)} \quad (4.6)$$

During apnea, \dot{V} is zero, so we can simplify the above transfer function. For clarity, we group parameters into a gain D , and a time constant E .

$$\frac{C_A(s)}{C_{PA}(s)} = \frac{(\dot{Q}_A/V_A)}{s + (\lambda_A \dot{Q}_A/V_A)} = \frac{D}{s + (1/E)} \quad (4.7)$$

For healthy lung, the time constant of an aerated compartment, E , is of the order of 1000 seconds.

In this section, I have deliberately not discussed the the size, nor the physical manifestation of a ‘lung unit’. We are describing the average behavior of an arbitrarily-sized lung compartment. Although such a description may describe the behavior of a single alveolus very well, it equally describes a collection of alveoli with homogeneous dynamics.

4.1.5 ROI-level description

The equations above describe C_A , the concentration of tracer within the alveolar airspace. Importantly, this is not the same thing as the mean concentration within the volume of the lung unit, since lung volume is composed of the airspace, blood, and tissue. We need thus to include in our equations the fact that the airspace is only a fraction of the total volume.

Consider also that when we model diseased lung, a single-compartment description is often insufficient to describe the dynamics of even very small regions. We need to consider a voxel or ROI to be composed of two compartments, one aerated, with volume V_A and kinetics $C_A(t)$, in which the

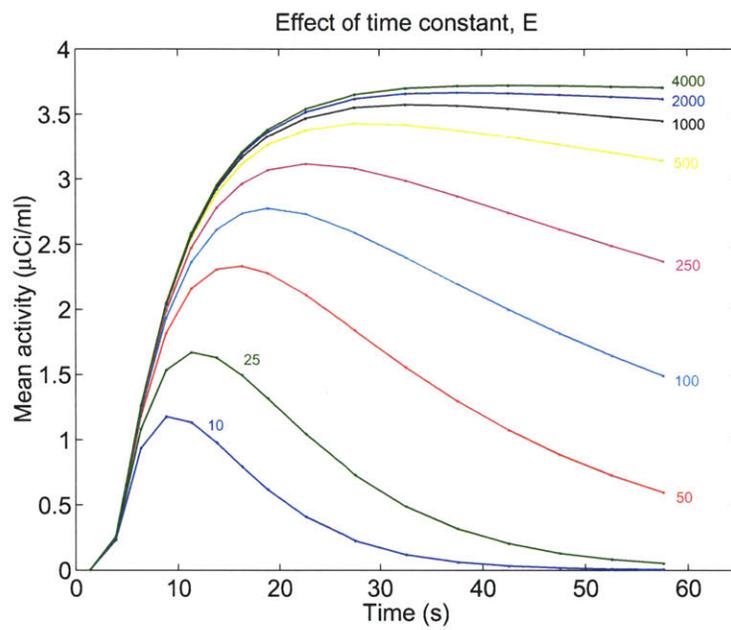


Figure 4.4: Effect of lung time constant, E , on model output, using typical values of other parameters. As the time constant becomes large (relative to the apnea duration), the kinetics become insensitive to this parameter.

tracer redistributes into airspace, and another, non-aerated compartment, V_{NA} , $C_{NA}(t)$, in which the blood does not participate in gas exchange, and in which the tracer redistributes minimally.

$$C_{ROI}(t) = \left(\frac{V_A}{V_{ROI}} \right) C_A(t) + \left(\frac{V_{NA}}{V_{ROI}} \right) C_{NA}(t) \quad (4.8)$$

We will discuss the kinetics of the non-aerated compartment, $C_{NA}(t)$, later, but a couple of observations may be relevant at this stage. Firstly, in a healthy lung, $C_{NA}(t)$ is negligible. Secondly, even in a diseased lung, because there is no preferential redistribution of the tracer, $C_{NA}(t)$ is short-lived, and rapidly (within 60s) tends to zero.

Let us assume for now that we are dealing with a healthy lung, and ignore the non-aerated compartment. We see from the above equation that to describe the kinetics of a whole ROI, we need to modify parameter D , introduced in equation 4.7 above.

$$C_{ROI}(s) = \left(\frac{V_A}{V_{ROI}} \right) C_A(s) = \frac{(\dot{Q}_A/V_{ROI})}{s + (\lambda_A \dot{Q}_A/V_A)} C_{PA}(s) \quad (4.9)$$

$$= \frac{D'}{s + (1/E)} C_{PA}(s) \quad (4.10)$$

4.1.6 Summary of parameters

In the development above, I have introduced lumped parameters B through E . This notation is summarized below.

$$\begin{aligned} B &= 1/\dot{Q}_{tot} \\ C &= V_H/\dot{Q}_{tot} \\ D &= \dot{Q}_A/V_A \\ D' &= \dot{Q}_A/V_{ROI} \\ E &= V_A/\lambda_A \dot{Q}_A \\ E' &= V_A/(\dot{V} + \lambda_A \dot{Q}_A) \text{ [introduced later]} \end{aligned}$$

4.1.7 Camera behavior

The PET scanner records the mean of specific activity over the imaging duration, t_i .

$$C^m = \frac{1}{t_i} \int_{t-t_i}^t C(t) dt \quad (4.11)$$

In transfer function form:

$$\frac{C^m}{C}(s) = \frac{1}{st_i} (1 - e^{-st_i}) \quad (4.12)$$

Note that, in this form, t is the time at the *end* of the imaging period of each frame.

4.1.8 Combined description of heart

We can now compose these elements to describe the measured activity in the right heart.

$$C_{heart}^m(s) = \underbrace{\left(\frac{A_{inf}}{t_{inf}}\right) \frac{1}{s} (1 - e^{-st_{inf}})}_{\text{Bolus}} \underbrace{(e^{-st_{c1}})}_{\text{Convection}} \underbrace{\left(\frac{B}{Cs+1}\right)}_{\text{Heart}} \underbrace{\frac{1}{st_i} (1 - e^{-st_i})}_{\text{Scanner}} \quad (4.13)$$

We will use this description in our identification of cardiac parameters.

4.1.9 Combined description of aerated lung during apnea

To form a complete description of the dynamics of the measured activity in aerated lung (equation 4.10), we merely need to insert a lung term to equation 4.13.

$$C_A^m(s) = \underbrace{\left(\frac{A_{inf}}{t_{inf}}\right) \frac{1}{s} (1 - e^{-st_{inf}})}_{\text{Bolus}} \underbrace{(e^{-st_{c2}})}_{\text{Convection}} \underbrace{\left(\frac{B}{Cs+1}\right)}_{\text{Heart}} \underbrace{\left(\frac{D'}{s + (1/E)}\right)}_{\text{Lung}} \underbrace{\frac{1}{st_i} (1 - e^{-st_i})}_{\text{Scanner}} \quad (4.14)$$

The only other change is that t_{c1} , the delay associated with convection of the tracer to the heart, is here replaced with t_{c2} , which represents the combination of delay t_{c1} , and a second delay associated with convection through the pulmonary artery to the lungs.

4.2 Shunt

In disease, regions of the lung may not ventilated: blood to these regions is not oxygenated, and is said to *shunt*. The fraction of the cardiac output that perfuses these areas is called the *shunt fraction*.

The tracer kinetics of a shunt compartment depend dramatically on whether the compartment is inflated or not. In asthma, for example, airway closure may leave an area of the lung unventilated but still inflated. In such regions, the tracer will stably redistribute out of the blood and into the airspaces, but not be washed out when breathing is resumed. This behavior is well-described by the equations developed above.

On the other hand, shunt may occur due to edema (flooding) or atelectasis (collapse). In such regions of non-aerated lung, the tracer kinetics do not exhibit an apnea plateau. The tracer has limited and transient redistribution into the surrounding tissue, and the kinetics rise to a peak before rapidly dropping back to zero.

In a lung with significant shunt, such as shown in figure 4.5, the whole-lung kinetics reflect the combination of a shunt compartment (in which the tracer concentration reaches an early peak and returns to zero) and an aerated compartment (in which the tracer concentration builds to a pseudo-stable plateau).

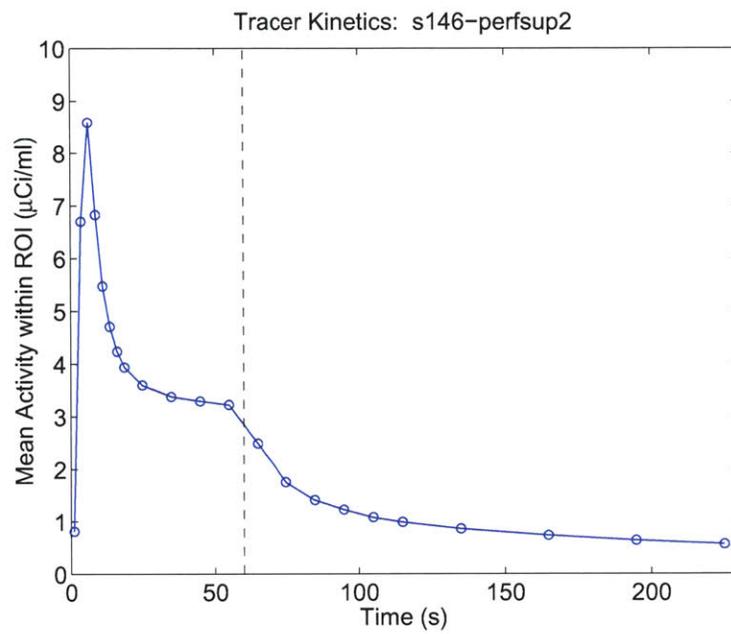


Figure 4.5: Kinetics from lung with very large amount of shunt. During apnea, the tracer reaches an early peak before dropping down to a plateau, reflecting the combination of a shunt compartment and an aerated compartment.

We can regard these non-inflated regions as one extreme of a continuum of lung states described by of the aerated lung model (equation 4.7). Instead of an aerated volume, we consider V_A to be a generic volume of distribution, and replace λ_A with λ_{NA} , a partition coefficient appropriate for the blood/tissue or blood/liquid repartitioning (consequently much closer to unity).

$$\frac{C_{NA}(s)}{C_{PA}(s)} = \frac{(\dot{Q}_{NA}/V_{NA})}{s + (\lambda_{NA}\dot{Q}_{NA}/V_{NA})} = \frac{D_{NA}}{s + (1/E_{NA})} \quad (4.15)$$

4.3 Theoretical Results

4.3.1 Integrating effect of well-inflated lung

In equation 4.10, the time constant, E , is of the order of 10^3 seconds in well-inflated lung. When E is large, the $1/E$ term in the denominator is almost insignificant. In this case, we can form a simplified approximation:

$$C_{ROI}(s) \approx \frac{D'}{s} C_{PA}(s) = \frac{(\dot{Q}_A/V_{ROI})}{s} C_{PA}(s) \quad (4.16)$$

That is, the lung behaves like an integrator with gain \dot{Q}_A/V_{ROI} . Now, as $t \rightarrow \infty$, the integral of C_{PA} is the total activity injected divided by cardiac output, A_{inf}/\dot{Q}_{tot} , assuming no recirculation of tracer. Thus

$$C_{ROI}(t \rightarrow \infty) \sim \frac{\dot{Q}_A}{V_{ROI}} \frac{A_{inf}}{\dot{Q}_{tot}} \quad (4.17)$$

This is an important result. In well-aerated lung units during apnea, the tracer will reach a pseudo-steady state where the total activity in a voxel (i.e. the product $C_{ROI}V_{ROI}$) is equal to the fraction of total cardiac output reaching that voxel times the total injected activity.

$$A_{voxel}(t \rightarrow \infty) \sim \frac{\dot{Q}_A}{\dot{Q}_{tot}} A_{inf} \quad (4.18)$$

This result can be derived more formally by substituting equations 4.1 and 4.3 into equation 4.16, and then calculating $\lim_{s \rightarrow 0} sC_{ROI}(s)$.

4.3.2 Apnea peak

To produce an expression for the peak activity during the apnea phase, we could transform equation 4.14 to the time domain, differentiate, and evaluate at the stationary point. However, the complexity of the resultant expression makes it difficult to interpret, and limits its usefulness. I have thus adopted a simplified form here.

Firstly we ignore the convectional delay, which does not effect the peak value. Secondly, we ignore the integrating effect of the camera. Thirdly we approximate the bolus as an impulse of area A_{inf} . Equation 4.14 is now reduced to

$$C_{lung}(s) \sim A_{inf} \left(\frac{B}{Cs+1} \right) \left(\frac{D'}{s+(1/E)} \right) \quad (4.19)$$

In the time domain, this is equivalent to

$$C_{lung}(t) \sim A_{inf} \left(\frac{BD'E}{C-E} e^{-t/C} + \frac{BD'E}{E-C} e^{-t/E} \right) = A_{inf} \frac{BD'E}{E-C} \left(e^{-t/E} - e^{-t/C} \right) \quad (4.20)$$

The stationary point occurs when $t = \frac{CE}{E-C} \ln \left(\frac{E}{C} \right)$. This leads to

$$C_{peak} \sim \frac{A_{inf}BD'E}{E-C} \left(\left(\frac{E}{C} \right)^{-\left(\frac{C}{E-C}\right)} - \left(\frac{E}{C} \right)^{-\left(\frac{E}{E-C}\right)} \right) \quad (4.21)$$

This expression is easier to interpret if we make the substitution $x = E/C$, the ratio of the time constants of the lung and the heart. Substituting also for parameters B, D' , we get

$$C_{peak} \sim \frac{A_{inf}\dot{Q}_A}{\dot{Q}_{tot}V_{ROI}} \left(\frac{x}{x-1} \right) \left(x^{\left(\frac{-1}{x-1}\right)} - x^{\left(\frac{-x}{x-1}\right)} \right) = \frac{A_{inf}\dot{Q}_A}{\dot{Q}_{tot}V_{ROI}} f(x) \quad (4.22)$$

We can trivially plot $f(x)$ over a representative range. Typical values of C are around 1 second, whereas E can vary from around 5 seconds (in collapsed lung) to 1000 in well-inflated regions with limited perfusion. As x becomes very large, $f(x)$ tends to unity as the lung behaves more and more like a perfect integrator.

This result is most relevant to assessment of shunt fraction. In lungs with significant shunt, the apneic tracer kinetics reach an early peak before dropping down to a plateau level, reflecting the two-compartment behavior of the lung. The difference in height between peak and plateau, expressed as a fraction of the peak height, is a rough measure of the shunt fraction. [13] This analysis gives us some idea about the error involved in this measure, which is invariably an underestimate of the true shunt fraction.

The ratio E/C is typically around 7 for a shunt compartment, and several hundred for an aerated compartment. Based on this development we expect the peak-to-plateau method to underestimate the shunt fraction as 72% of the true value.

Note that the finite imaging duration of the PET camera is not taken into account here, and this will decrease the measured value of the peak, further increasing the underestimation of the shunt fraction. Similarly, we have assumed that the tracer is infused in an instantaneous impulse: the longer the duration of the bolus, the lower will be the peak value.

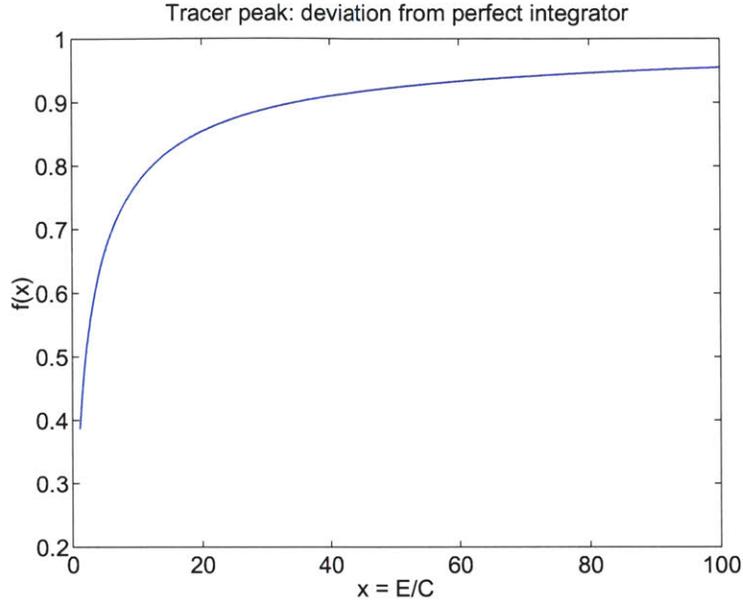


Figure 4.6: Dependence of apnea peak on the ratio of the time-constants for the heart and lung, E/C . As this ratio becomes large, the lung behaves like a perfect integrator.

4.3.3 End-apnea level

Equation 4.20 provides an approximation to the kinetics of an ROI during the apneic phase. If we are only interested in the end-apnea value, we can make further simplifications. The time constant of the heart, C , is around one second, so that late in the apnea the second exponential term is negligible. Additionally, the time constant of inflated lung, E , is 10^3 seconds, so that $E/(E - C)$ is approximately unity. Making these approximations, equation 4.20 becomes

$$C_{lateap}(t) \sim \frac{A_{inf}\dot{Q}_A}{\dot{Q}_{tot}V_{ROI}} e^{-t/E} \quad (4.23)$$

This equation is only valid for inflated compartments: shunt compartments have short enough time constants (5-10s) that they retain minimal tracer at the end of the apnea period.

If E is large relative to t , so that the exponential term is approximately unity, equation 4.23 reduces to equation 4.17. Applying the latter equation results in an over-estimate of the end-apnea concentration by about 10%, using typical parameter values.

4.3.4 Apnea integral

The integral to infinity of the activity within a single lung compartment (equation 4.14) has a conveniently simple form:

$$\int_0^{\infty} C_A^m(t) dt = A_{inf} B D' E \quad (4.24)$$

This has limited relevance to healthy lung. However in non-aerated lung, the time-constant, E , is short enough that the tracer concentration has normally fallen to zero by the end of the apnea, so that the integral to infinity is approximated by the integral to t_{ap} :

$$\int_0^{\infty} C_{NA}^m(t) dt \sim \int_0^{t_{ap}} C_{NA}^m(t) dt = A_{inf} B D' E = \frac{A_{inf} V_{NA}}{\dot{Q}_{tot} V_{ROI} \lambda_{NA}} \quad (4.25)$$

Note that this quantity is independent of the perfusion to the non-aerated compartment.

We can use this result in a number of ways. For example, if we select a pure shunt region (so that $V_{NA} = V_{ROI}$), and know the cardiac output, \dot{Q}_{tot} , we can estimate λ_{NA} , or conversely we can use an assumed value of λ_{NA} to estimate cardiac output.

4.4 Evaluation of Model Output

Equations 4.13 and 4.14 give a complete description of the expected kinetics of the heart and lungs in the Laplace domain. These must be transformed back to time-domain expressions which are evaluated at discrete time points (the end-frame times) and for distinct values of t_i (the frame durations).

Instead of writing a fully-expanded time-domain form, we can express these compactly by considering the description to consist of a ‘kernel’ function convolved with a series of delays. Rearranging 4.13 slightly, we have:

$$C_{heart}^m(s) = \frac{A_{inf} B}{t_{inf} t_i} \underbrace{(e^{-st_{c1}}) (1 - e^{-st_{inf}}) (1 - e^{-st_i})}_{\text{Delays}} \underbrace{\frac{1}{s^2 (Cs + 1)}}_{\text{Kernel}} \quad (4.26)$$

The last term is transformed back to the time domain as follows.

$$\frac{1}{s^2 (Cs + 1)} \longleftrightarrow t - C + C e^{-t/C} \quad (4.27)$$

Next we define a delay operator:

$$\text{delay}(f(t), t_d) = \begin{cases} 0 & \text{for } t < t_d \\ f(t - t_d) & \text{for } t \geq t_d \end{cases} \quad (4.28)$$

Putting these together results in a compact expression in a form suitable for rapid evaluation by the optimization algorithm used during parameter estimation.

$$C_{heart}^m(t) = \frac{A_{inf}B}{t_{inf}t_i} \{ \text{delay}(f_h, t_{c_1}) - \text{delay}(f_h, t_{inf} + t_{c_1}) - \text{delay}(f_h, t_i + t_{c_1}) + \text{delay}(f_h, t_i + t_{inf} + t_{c_1}) \}$$

where $f_h = t - C + Ce^{-t/C}$

(4.29)

We can do the same for the description of the lung during apnea (equation 4.14):

$$C_{lung}^m(t) = \frac{A_{inf}BDE}{t_{inf}t_i} \{ \text{delay}(f_l, t_{c_2}) - \text{delay}(f_l, t_{inf} + t_{c_2}) - \text{delay}(f_l, t_i + t_{c_2}) + \text{delay}(f_l, t_i + t_{inf} + t_{c_2}) \}$$

where $f_l = t - (C + E) + \frac{C^2}{(C - E)}e^{-t/C} + \frac{E^2}{(E - C)}e^{-t/E}$

(4.30)

4.5 Washout

Equation 4.6 describes the behavior of an aerated lung unit, which we simplified above by assuming that $\dot{V} = 0$. With non-zero, but constant ventilation, we can rearrange as follows.

$$\frac{C_A(s)}{C_{PA}(s)} = \frac{\dot{Q}_A}{V_A s + (\dot{V} + \lambda_A \dot{Q}_A)} = \frac{\dot{Q}_A/V_A}{s + (\dot{V} + \lambda_A \dot{Q}_A)/V_A} = \frac{D}{s + (1/E')} \quad (4.31)$$

Our transfer function is still first-order, but we have replaced the coefficient E in equation 4.7 with E' .

Commonly, in Laplace-domain descriptions of system dynamics, initial conditions are assumed to be zero, and can be neglected. In this case, however, the initial conditions are non-zero: there is tracer in the lung at the start of the washout. Including initial conditions, we have

$$C_A(s) = \frac{D}{s + (1/E')} C_{PA}(s) + \frac{1}{s + (1/E')} C_A(t = 0) \quad (4.32)$$

If the concentration in the pulmonary artery, C_{PA} , is zero throughout the washout, then we can neglect the first term. This condition will apply except when the apnea is very brief (so that tracer is still arriving in the lung at the start of the washout) or if there is significant recirculation of tracer.

Assuming $C_{PA}(t) = 0$ for $t > 0$, and converting to the time domain:

$$C_A(t) = C_A(0)e^{-t/E'} \quad (4.33)$$

Thus the washout (of a single compartment) is a simple exponential decay. We can express the washout rate, $1/E'$, as the sum of other quantities:

$$\frac{1}{E'} = \left(\frac{\dot{V} + \lambda_A \dot{Q}_A}{V_A} \right) = \left(\frac{\dot{V}}{V_A} + \frac{1}{E} \right) = \left(s\dot{V} + \frac{1}{E} \right) \quad (4.34)$$

In the above equation, E is the time-constant of aerated lung identified during the apneic phase (section 4.1.4), and $s\dot{V}$ is the specific ventilation, \dot{V}/V_A .² We see that the rate of tracer removal during the washout period is the sum of the specific ventilation, the rate at which tracer is removed by breathing, and $1/E$, the rate at which tracer is resorbed into the blood.

4.5.1 Washout Integral

If we sum the tracer concentration over the duration of the washout, i.e. integrate equation 4.33, we get the following result.

$$I_{wo} = \int_0^{t_{wo}} C_A(t) dt = E' C_A(0) \left(1 - e^{-t_{wo}/E'}\right) \quad (4.35)$$

Substituting for $C_A(0)$ the tracer concentration at end-apnea/washout-start (equation 4.23), and expanding E' :

$$I_{wo} = \left(\frac{V_A}{\dot{V} + \lambda_A \dot{Q}_A}\right) \left(\frac{A_{inf} \dot{Q}_A}{\dot{Q}_{tot} V_{ROI}} e^{-t_{ap}/E}\right) \left(1 - e^{-t_{wo}/E'}\right) \quad (4.36)$$

If we make the following assumptions:

- the washout is long, i.e. $t_{wo} \gg (s\dot{V} + 1/E)$
- the apnea is short, i.e. $t_{ap} \ll E$
- the lung is well-ventilated, i.e. $\dot{V} \gg \lambda_A \dot{Q}_A$

then we can neglect the $e^{-t_{wo}/E'}$ term, treat $e^{-t_{ap}/E}$ as unity, and neglect $\lambda_A \dot{Q}_A$, giving the following expression:

$$I_{wo} \sim \frac{A_{inf} V_A}{\dot{Q}_{tot} V_{ROI}} \left(\frac{\dot{Q}_A}{\dot{V}}\right) \quad (4.37)$$

Under these circumstances, the washout sum is proportional to \dot{Q}_A/\dot{V} , or the inverse of the VQ ratio. If we assume further that the degree of inflation is uniform in the lung (so that V_A/V_{ROI} is constant), and that both the apnea and washout are locally uni-compartmental, then the voxel-by-voxel washout sum will be an image proportional to \dot{Q}_A/\dot{V} .

These assumptions are very strong, however, and, notably, are violated in almost every diseased lung state. This result, then, should only be considered a first approximation, and is not generally applicable.

²This notation, $s\dot{V}$, for specific ventilation, can potentially cause confusion with the Laplace variable, s . Regrettably, it is standard in the field.

4.6 Equivalent Ventilation

When modeling the washout with multiple compartments, identification of the system generates multiple parameters. One problem with this approach is how to present the data effectively. With a single-compartment washout we can generate an image of specific ventilation. As soon as we move to a two-compartment model, we have at least three parameters to display. Although we can make images showing separate parameters, these are difficult to comprehend individually.

Ideally, we would like to be able to calculate a single parameter, which is equivalent in some sense to the multiple parameters identified for each voxel or ROI. The obvious ‘equivalent’ for ventilation would be the single-compartment ventilation that would result in the same oxygen content as the mixed blood leaving the multiple compartments. Note that this is not the same ventilation that would result in the same P_{CO_2} : we cannot match both measures with a single compartment.

4.6.1 Oxygenation from \dot{V}/\dot{Q} ratios

Consider the diagram of the lung unit shown in figure 4.3. Although this represents gas exchange of nitrogen-13, it applies equally to the gas exchange of other species. To describe oxygen exchange, equation 4.4 just needs to be modified to account for the fact that the inspired concentration, C_I , is non-zero. We also replace C_{PA} , the concentration in the pulmonary artery with $C_{\bar{V}}$, the mixed venous concentration.

$$V_A \frac{dC_A}{dt} = \dot{Q}(C_{\bar{V}} - C_{out}) - \dot{V}(C_I - C_A) \quad (4.38)$$

At steady-state, $\frac{dC_A}{dt} = 0$. Hence

$$\dot{V}(C_I - C_A) = \dot{Q}(C_{out} - C_{\bar{V}}) \quad (4.39)$$

As before, we assume that blood and gas equilibrate, so that the partial pressure in the aerated space equals the partial pressure in the blood leaving.

$$P_{AO_2} = P_{atm}C_A = P_{O_{2out}} = S(C_{out}) \quad (4.40)$$

Where S describes the O_2 dissociation curve. We can substitute for $C_A = S(C_{out})/P_{atm}$ into equation 4.39, but we are not able to produce an explicit expression for C_{out} in terms of \dot{V}/\dot{Q} . We can, however, do the reverse:

$$\frac{\dot{V}}{\dot{Q}} = \frac{C_{out} - C_{\bar{V}}}{C_I - S(C_{out})/P_{atm}} \quad (4.41)$$

Given this expression, we can tabulate \dot{V}/\dot{Q} ratios in terms of C_{out} , and interpolate these values to calculate C_{out} in terms of \dot{V}/\dot{Q} . [Figure 4.7]

The saturating nature of oxygen-hemoglobin association means that C_{out} is effectively limited no matter how large \dot{V}/\dot{Q} becomes. Oxygenation is minimally effected by high \dot{V}/\dot{Q} units. This is worth bearing in mind when attempting to validate the nitrogen-13 method by predicting blood gases. The predicted oxygenation is very insensitive to the shape of the higher parts of the \dot{V}/\dot{Q} distribution. This insensitivity increases as the inspired oxygen fraction is increased.

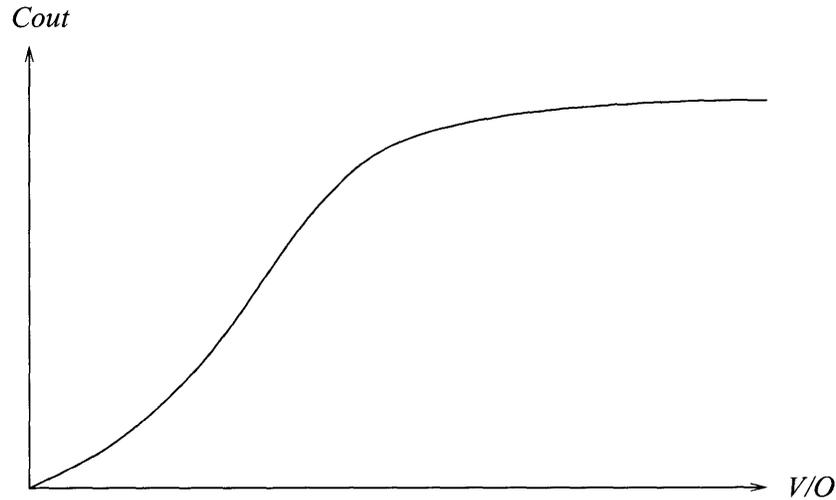


Figure 4.7: Relationship between the oxygen content of blood leaving a lung unit and the local ventilation-perfusion ratio. As \dot{V}/\dot{Q} becomes large, C_{out} tends to saturate: blood oxygenation is insensitive to high \dot{V}/\dot{Q} ratios. Note that the curve depends on the inspired oxygen concentration.

Because blood carbon dioxide content does not exhibit the same saturation, the C_{out} vs \dot{V}/\dot{Q} relationship for CO_2 does not flatten in the same way. Consequently, prediction of arterial P_{CO_2} depends more on the entire \dot{V}/\dot{Q} distribution than P_{O_2} .

4.6.2 Calculation of ‘equivalent’ ventilation

Given a relationship for C_{out} in terms of \dot{V}/\dot{Q} , we can calculate an equivalent ventilation easily. Given a number of parallel compartments with different ventilations and perfusions, we calculate the blood concentration that results after downstream mixing:

$$\overline{C_{out}} = \frac{\sum_i (Q_i C_{out i})}{\sum_i Q_i} \quad (4.42)$$

Next, using equation 4.41 we calculate the single-compartment \dot{V}/\dot{Q} ratio that would result in $\overline{C_{out}}$, and multiply by $\sum Q_i$ to give equivalent ventilation, $\overline{\dot{V}}$.

4.7 Parameter Identification

Referring to equation 4.14, the parameters that describe the system fall into three groups.

Firstly, we have parameters that are known *a priori*. These are the amount of activity injected, A_{inf} , and the imaging durations t_i .

Secondly, there are parameters which apply globally (i.e. are constant throughout the lung) and need only be estimated once. These are the cardiac parameters B , C , and combined convection delay, t_{c2} .

Finally we have parameters that vary spatially in the lung: D' and E , which will be estimated using the apnea data, and E' , which is estimated from the washout. This detail of this phase of the identification, however, depends greatly on the physiological state of the lung. In a well-inflated lung (e.g. asthma), we can reasonably model the apnea kinetics with a single, aerated compartment, but, as discussed in section 4.3.1, estimation of parameter E (the time constant) is difficult, as the apnea kinetics are very insensitive to this parameter when it is large. (I.e. when $t \ll E$.)

In a lung where collapse or flooding is present, then we almost certainly will need a two-compartment model for the apnea phase. Similarly, in a healthy lung, a single compartment is usually sufficient to describe the washout, but in a pathologic lung, we need at least two compartments per ROI. In all cases, however, the overall parameter identification strategy is the same.

4.7.1 Identification of Cardiac Parameters

We start by defining a mask for the right heart (and pulmonary arteries). In the first few frames of the image, tracer arrives in the heart before any has accumulated in the lungs. The heart shows up as a clear ‘hot’ region in these early frames, and thus does not require sophisticated methods to delineate it. We simply sum those frames in which tracer is obviously present in the heart, but before significant activity has arrived in the lung. We then generate a mask by thresholding this summed image. Plotting the kinetics of voxels within this mask provides a good approximation to the tracer kinetics through the heart.

The relatively large number of voxels within the heart mask ($\sim 10^3$), and the high tracer concentrations present ($\sim 10\text{-}50 \mu\text{Ci/ml}$) mean that the kinetics so derived are relatively clean, and do not require post-processing. However, the concentration in the heart is consistently underestimated by a small fraction because of partial volume effects (i.e. because voxels may be only partially occupied by intra-cardiac blood). As a consequence, we may overestimate pulmonary perfusion by a corresponding factor.

One way to deal with this issue is to examine the kinetics of a smaller ROI known to lie wholly within the right heart. Although the kinetics of such a region are undesirably noisy, we can use these kinetics to correct the absolute magnitude of our full heart mask kinetics.

Having made an approximation to the tracer kinetics through the right heart, we need to identify parameters B (gain, $1/\dot{Q}_{tot}$), C (time constant, V_H/\dot{Q}_{tot}), and convection delay t_{C1} . These parameters are identified using a non-linear optimization algorithm (Levenberg-Marquardt) to minimize the error (in a least squares sense) between the experimental data, and the simulated curve (calculated using the method described above in section 4.4). Figure 4.8 shows an example of heart kinetics, the model output, and the estimated cardiac parameters.

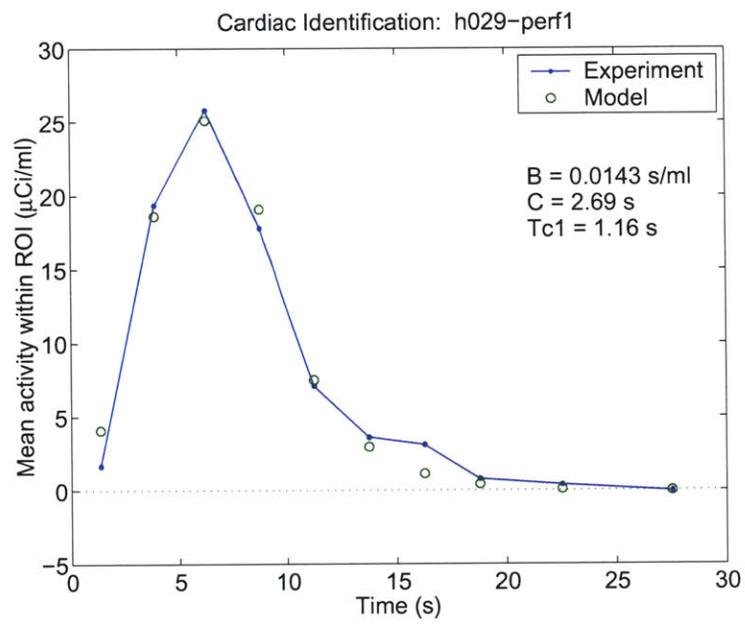


Figure 4.8: Example of tracer kinetics through right-heart ROI, shown with results of model fit and associated estimates or cardiac parameters.

4.7.1.1 Double Infusion Model

In animal studies, the tracer is infused through a catheter placed via the jugular vein which extends to the SVC. Due to the large caliber of this vessel, it is possible to infuse at higher flow-rates (10ml/s) than in human studies (5ml/s). The tracer is delivered very close to the right-heart, and is consequently visible almost immediately in the cardiac ROI.

In humans, the tracer is introduced via a peripheral (cubital) vein, and must be infused more slowly to avoid trauma to the vein at the IV site. To accelerate and maximize the delivery of the tracer, a saline flush is given immediately after the bolus infusion. Even with a flush, the tracer takes a substantial and highly variable amount of time (1-5 s) to reach the heart.

One problem with the post-infusion flush is that it is given manually, and is consequently variable in timing, volume, and pressure.³ The delay after the bolus, and before the flush is performed means that in many cases, a smaller, second peak in the cardiac kinetics is apparent as flushed tracer arrives late. This effect is just visible in figure 4.8; a more extreme example is shown in figure 4.9, in which the late flush causes a distinct ‘ripple’ in the lung kinetics.

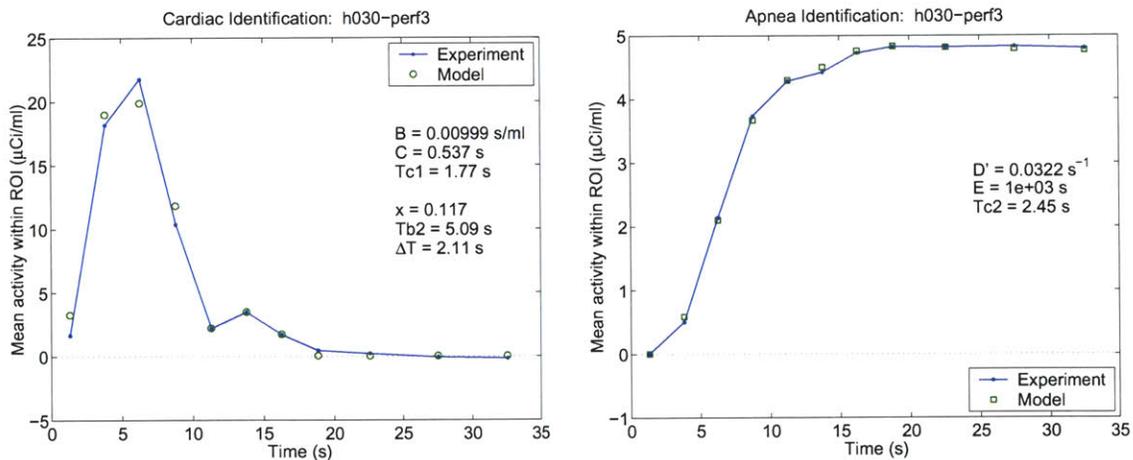


Figure 4.9: Tracer kinetics from image with late flush. On the left, a second peak is visible in the heart kinetics, which results in a ‘ripple’ in the lung kinetics, shown on the right. Response and parameter estimates from a double-infusion model (see text) are shown.

It is possible that this effect is not due to late or inconsistent flushing. It may be partially or entirely due to a ‘second compartment’ effect caused by compliance of the peripheral veins, or by delayed delivery of tracer that has been convected distally, against the normal venous flow. Whatever the mechanism, however, the heart kinetics look very much like a response to a second infusion of tracer, and consequently the following model is proposed.

³The consequences of flush variation have only recently been recognized, and we have recently taken steps to standardize the flush. In the studies discussed here, however, the flush was unregulated.

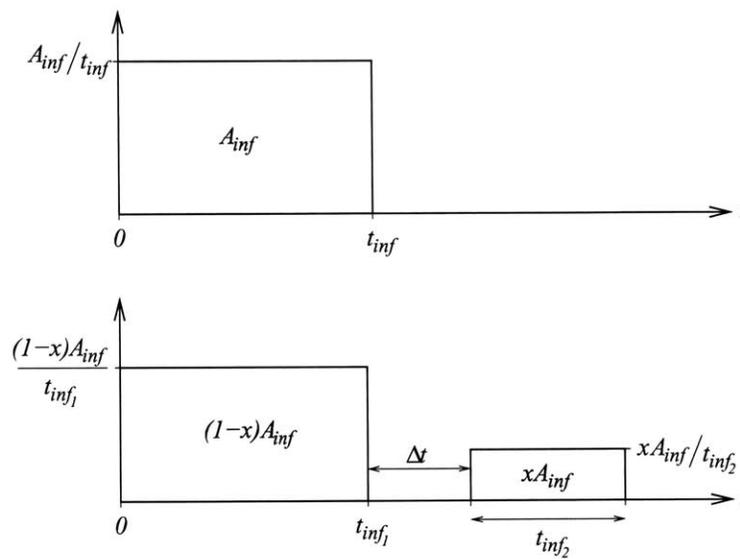


Figure 4.10: Double infusion model (bottom) compared to original (top). Following flushing of infusion line, an second bolus of tracer arrives in the heart. This is modeled as a double infusion: a fraction, x , of the total amount of tracer, A_{inf} , is infused in a second bolus, starting at time Δt after the end of the first, and with duration t_{inf_2} . A concomitantly reduced amount, $(1-x)A_{inf}$, of tracer is infused in the first bolus.

The total amount of tracer infused, A_{inf} , is divided into two infusions, with fraction x infused in a second bolus (figure 4.10). The second bolus commences Δt after the end of the first, and has duration t_{inf_2} . Of these parameters, A_{inf} and t_{inf_1} (the duration of the first bolus) are considered to be known *a priori*, while x , Δt , and t_{inf_2} are estimated from the cardiac kinetics.

The linearity of the heart and lung system means that evaluating the output for a double infusion is trivial, as the response to the second bolus can be superimposed on the response to the first. Additionally, describing the flushed tracer with another bolus simplifies matters still further, as we can use the same equations that describe the bolus response, and merely shift in time.

One might argue that such a model of the flush is unrealistic. But this model has the advantage of simplicity, minimizes the number of additional parameters that must be estimated, and gives output that satisfactorily matches the cardiac kinetics in every image analyzed.

4.7.2 Identification of Lung Parameters

4.7.2.1 Perfusion

The perfusion and inflation parameters are estimated from the local tracer kinetics during apnea, as described below.

Single-compartment In a healthy lung, a single compartment model suffices to describe the tracer kinetics during apnea. In this case, in principle, we need to estimate parameters D' and E for each ROI. In practice, however, estimation of the time constant, E , is considerably more difficult than D' , since the tracer kinetics are much less sensitive to this parameter.

Some examples of single-compartment parameter identification are shown in figure 4.11. The data shown in these examples are taken from human asthmatics. In each of these cases, the lung is generally well-inflated, so that parameter E is large. The insensitivity of the analysis to this parameter is reflected in the identified parameters shown on the graphs: in each case E is left unchanged from its initial estimate of 1000 seconds.

Note that these examples show whole-lung kinetics. Despite the improvement of signal-to-noise ratio as compared with small-ROI kinetics, we are unable to make a meaningful estimate of E . The situation is slightly improved in animal studies, where we have more information in the apnea plateau⁴ and can detect the delicate downward slope characterized by this parameter. However, even in these circumstances, the noise present in voxel-by-voxel data prevents high-resolution estimation of this parameter.

Although disappointing, this is not a great loss. The primary parameter of interest is the normalized perfusion, D' , which can be robustly estimated using noisy data. The only practical utility of the time constant, E , is in correcting the washout rate for tracer resorption. If the apnea plateau is long enough, we can make a global estimate of E from the apnea data. Alternatively, we can use

⁴Due to more rapid tracer arrival, longer apnea duration, and less chest-wall attenuation.

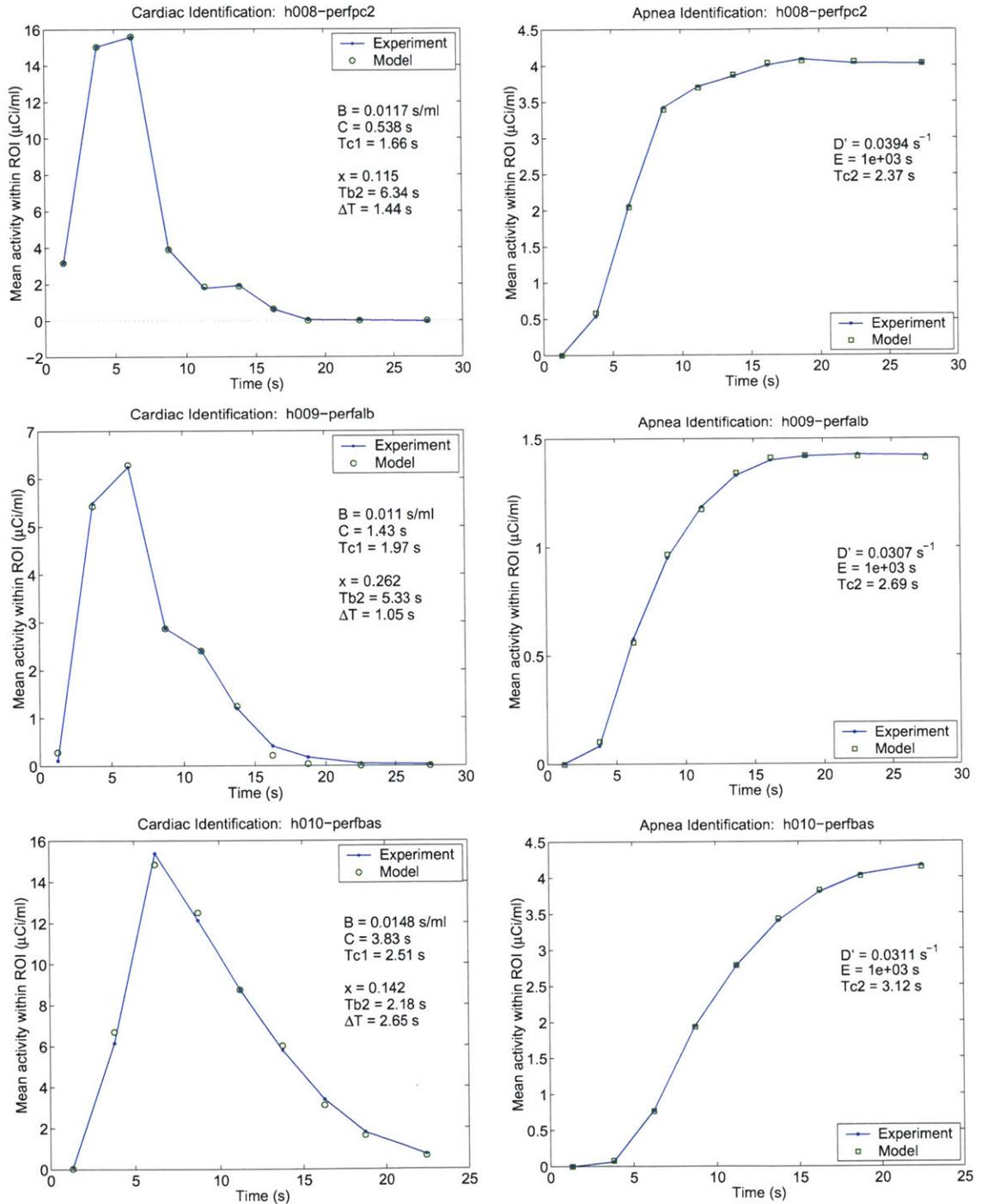


Figure 4.11: Examples of parameter identification applied to the apnea phase of images of human asthmatics. On the left-hand side are the heart kinetics (fitted with double-infusion model) and on the right-hand side, the perfusion results. Note that, in each case, parameter E , the time-constant of the lung is left at its initial estimate, 1000 seconds.

a transmission scan to estimate regional lung inflation, and then use the regional perfusion data to make a regional estimate of E .

I.e. from the transmission scan we estimate regional lung inflation, V_A/V_{ROI} . And, using regional values of parameter $D' = \dot{Q}_A/V_{ROI}$, and our *a priori* knowledge of the partition coefficient, λ_A , we can then estimate:

$$\hat{E} = \left(\frac{V_A}{V_{ROI}} \right) / \lambda_A \left(\frac{\dot{Q}_A}{V_{ROI}} \right)$$

Double-compartment In lung with significant non-aerated shunt, we need to model the apnea phase with two compartments. The first, long time-constant compartment, describes the kinetics of aerated lung, while the second describes the kinetics of a short time-constant, shunting compartment.

Examples of two-compartment identification are shown in figure 4.12. As with the single-compartment examples, we are unable to estimate parameter E_A , the time constant of the aerated compartment, so in practice, we are only estimating three parameters: D'_S, E_S, D'_A .

Since the two-compartment model requires three parameters to be estimated, voxel-by-voxel parameter estimation is not possible. Thus, in order to make regional estimates of shunt, we can either analyze large ROI's, or alternatively attempt to employ a group-wise parameter-estimation method such as that described in chapter 5.

4.7.2.2 Ventilation

Single-compartment If no more tracer is arrives in the lungs during the washout, the washout kinetics should be a simple exponential decay, with rate constant equal to the sum of the specific ventilation and the inverse of the (previously-estimated) time-constant, E [equation 4.33]. In this case, the simple form of the washout kinetics means that we can analyze simply by fitting an exponential to the washout curve. Specific ventilation is estimated by subtracting $1/E$ from the washout rate, although since a typical rate constant would be around 0.1 s^{-1} and $1/E$ on the order of 0.001 s^{-1} , this correction is usually negligible.

If tracer is still arriving in the lungs during the washout (if the apnea duration is very short, the tracer infusion is very long, or if there is significant tracer recirculation), we need to estimate parameters using a model-matching method like that used for the heart and apnea kinetics. (Using equation 4.32.) This approach is not investigated in this thesis.

Double-compartment In diseased lungs, particularly those that are highly bronchoconstricted like asthma, the washout of even a single voxel has multiple time-constants. Under these conditions, we need to use a multiple-compartment washout model. In practice, it proves not to be feasible to attempt to identify more than two compartments, as a two-compartment model is capable of matching all experimentally-observed kinetics to a high degree of accuracy. The choice of a two-compartment model is justified additionally by other research conducted in our laboratory that

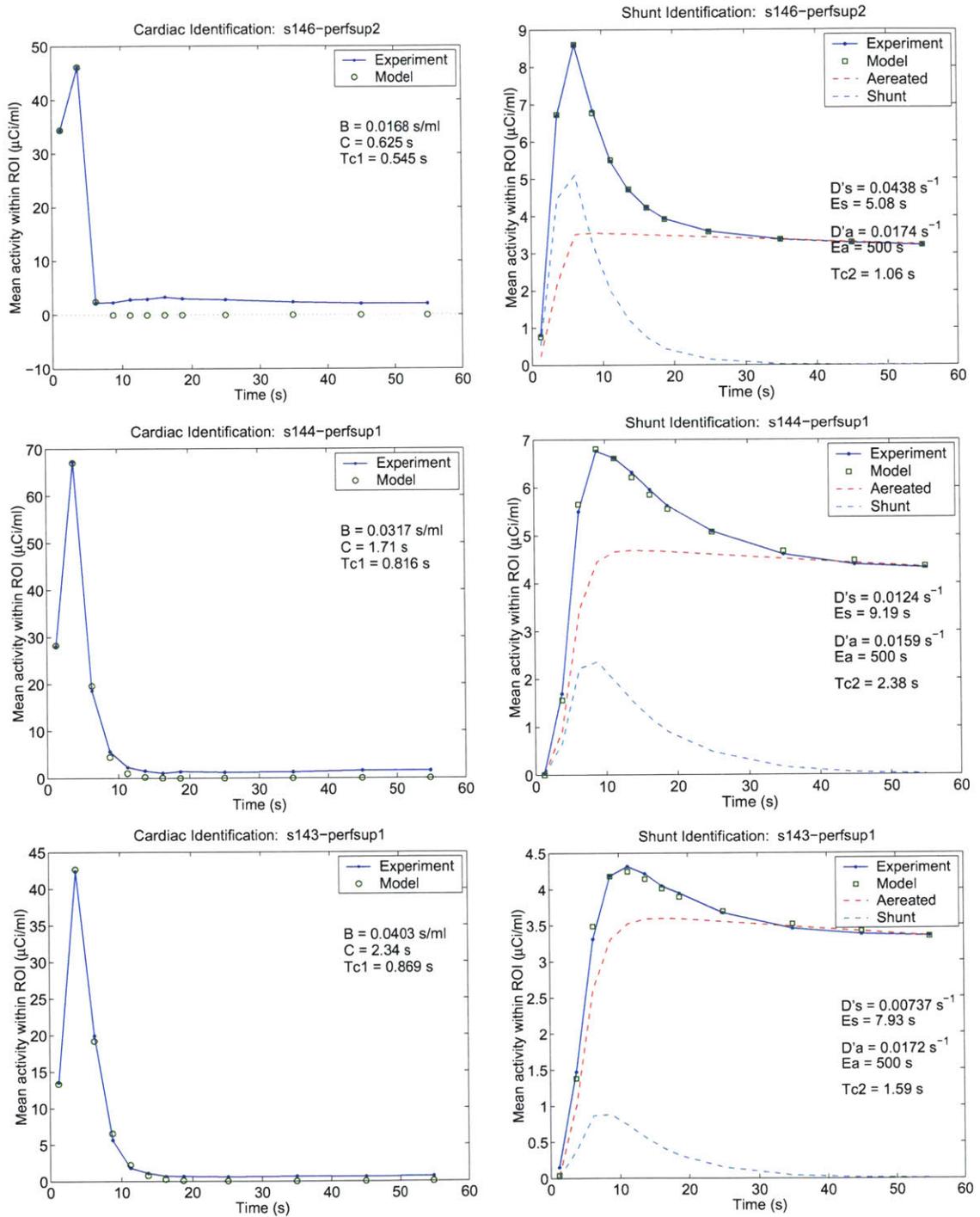


Figure 4.12: Examples of double-compartment parameter identification applied to the apnea phase of images of sheep injured by lavage. On the left-hand side are the heart kinetics (single-infusion model) and on the right-hand side, the perfusion results. Note that there is not enough information to estimate the time-constant of the aerated compartment, E_A , which is left at its initial estimate of 500 seconds.

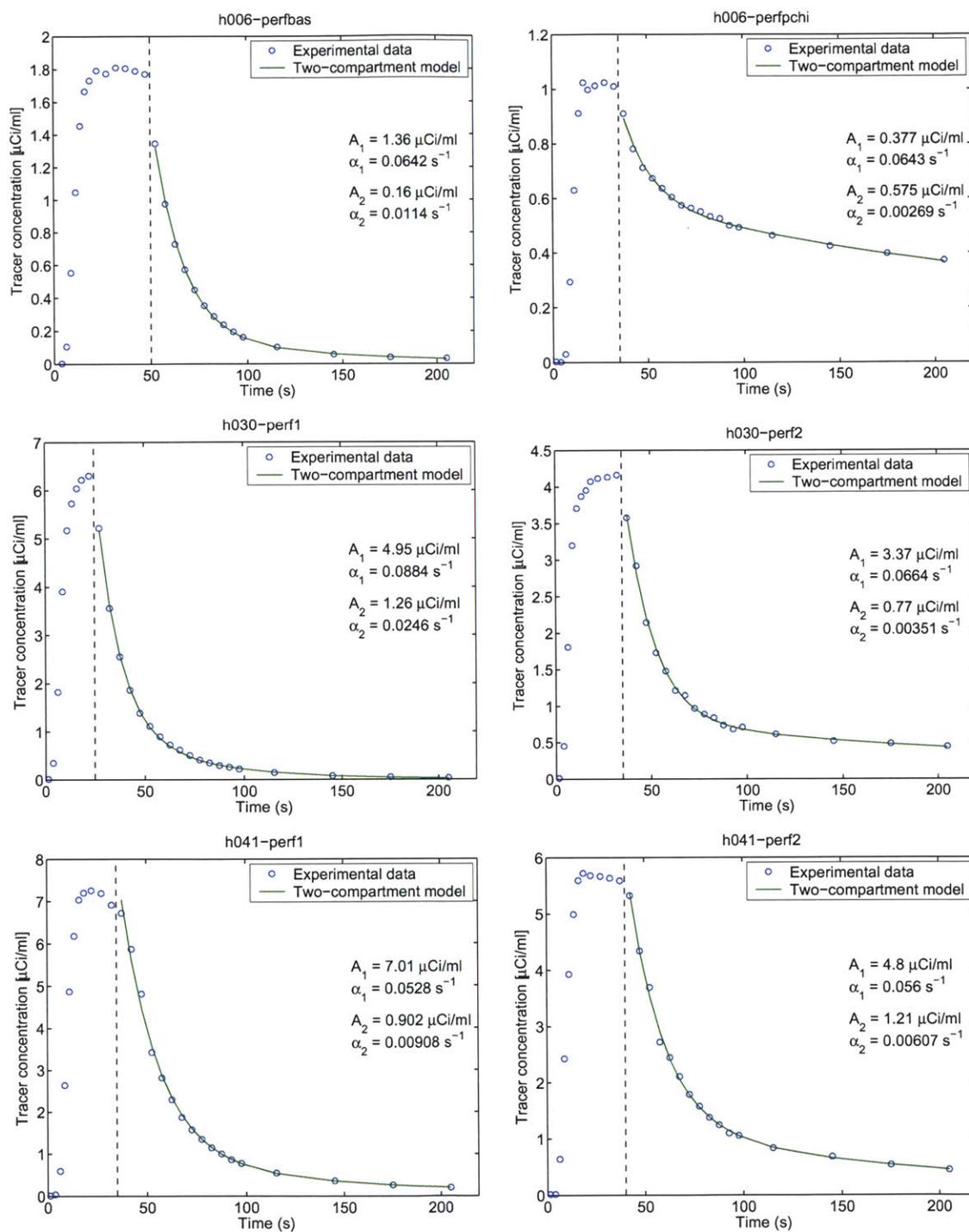


Figure 4.13: Examples of two-compartment parameter identification applied to washout kinetics. In all cases, the washout is well-described by a double-exponential. Consequently, it does not make sense to employ a washout model with more than two compartments.

suggests that bronchoconstricted lung has a bistable behavior, in which airways divide into highly-constricted or almost-unaffected populations. (This work, not yet published, extends the locally bistable airway model of Anafi and Wilson [2] to a whole-lung analysis.)

Ventilation analysis using a double-compartment model involves estimation of a minimum of three parameters. As with perfusion, voxel-by-voxel parameter estimation is not feasible, and we can only analyze large ROI's. The selection of spatially contiguous regions for analysis means that we lose much of the PET camera's resolving capability.

An alternative approach involves selection of groups of voxels that are not physically adjacent but instead have similar kinetics. This method is explored in the next chapter, with specific application to two-compartment washout analysis.

4.8 Summary

In this chapter a simple model is developed of the tracer kinetics in the right heart and lungs. This model allows us to understand the capabilities of the imaging technique and exactly what physical parameters we can expect to estimate.

The model is used to predict the approximations for \dot{V} and \dot{Q} that apply to healthy lungs, and shows why these do not apply to diseased lung.

Application of the model to parameter estimation is described and then demonstrated in whole-lung kinetics from injured animals and asthmatic humans.

Extension of the method to analysis of small regions of interest is discussed, and is further explored in the next chapter.

Chapter 5

Small ROI Parameter Estimation

5.1 Introduction

In the previous chapter, a model was developed of the expected tracer kinetics in the lung, and its use demonstrated in estimating physiological parameters from experimental data.

Unfortunately, while this method is suitable for analysis of large regions, which have good *signal-to-noise* (SNR) in their kinetics, it does not perform well when an image is broken into small regions, where noise precludes reliable parameter estimation.

In this chapter, a method for analysis of PET images is described which preserves spatial information, and can be applied voxel-by-voxel. The viability of the technique is explored by applying it to two-compartment analysis of washout kinetics from asthmatic volunteers.

Results from these asthma studies are presented in the following chapter. [Chapter 6]

5.2 Image Noise

PET images intrinsically exhibit a trade-off between resolution and noise [Section ??]. The kinetics of a small *region of interest* (ROI) is observed to be very noisy. The mean of many measurements, however, ‘dilutes’ the noise of individual measurements. If the noise is unbiased, the mean kinetics of a many-voxel ROI approximates the actual kinetics of that region more accurately as the number of voxels increases, *provided that the region is homogeneous*.

This effect is illustrated in figure 5.1, in which tracer kinetics of a number of ROI’s of decreasing size are plotted. For clarity, each curve has been vertically offset from the previous one. (Otherwise the curves would be almost coincident.) This image was reconstructed to 4mm x 4mm x 6.5mm voxels, and it is clear that the amount of noise visible in a single voxel ROI is enough to obscure all but the grossest trends in the data.

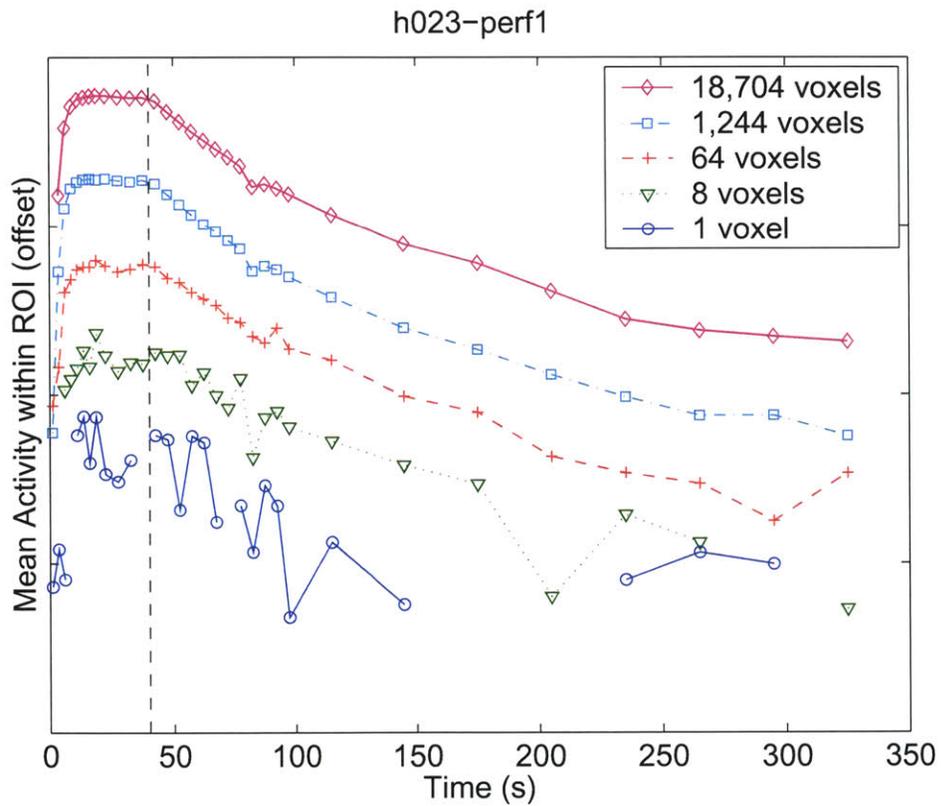


Figure 5.1: As a region of interest is made smaller and smaller, the amount of noise increases. (For clarity, the kinetics in this graph have been offset vertically.) Since this is a logarithmic scale, negative values are omitted.

These kinetics are plotted on a logarithmic scale, and the ‘missing’ points in the lower curves are negative values. Of course, a real tracer concentration can never be negative: such invalid data are an undesirable artifact of the convolution-back-projection algorithm used to reconstruct the image.

Note also that these data are taken from a study in which a relatively large amount of activity was infused into a subject with a thin chest-wall. Thus these kinetics have comparatively good SNR: in most images the data will be noisier still. [See section 1.5]

5.2.1 The ‘Pop-Up’ Effect

The amount of noise present in PET images can lead to unexpected artifacts.

Figure 5.2 is an image taken from a bronchoconstricted subject at the end of the washout period. Substantial amounts of tracer are retained in the lung, but as the image shows, the tracer is not retained uniformly across the lung, but is concentrated in highly-constricted areas.

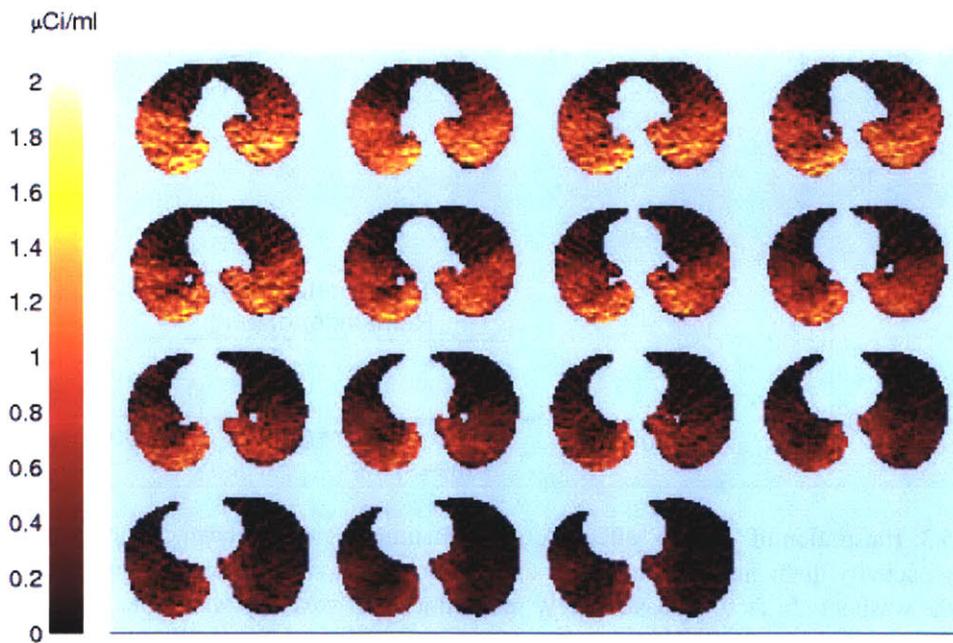


Figure 5.2: End-washout image showing tracer-retention in a bronchoconstricted subject. Tracer is not uniformly retained, but is concentrated in highly-constricted areas.

In order to plot the kinetics of these highly-constricted areas, we need first to make a mask that defines the ROI. Let us generate this mask by including voxels in which the measured activity in the last two washout frames is greater than 40% of the maximum activity in those frames. Figure 5.3 shows the tracer kinetics for this region and the remainder of the lung.

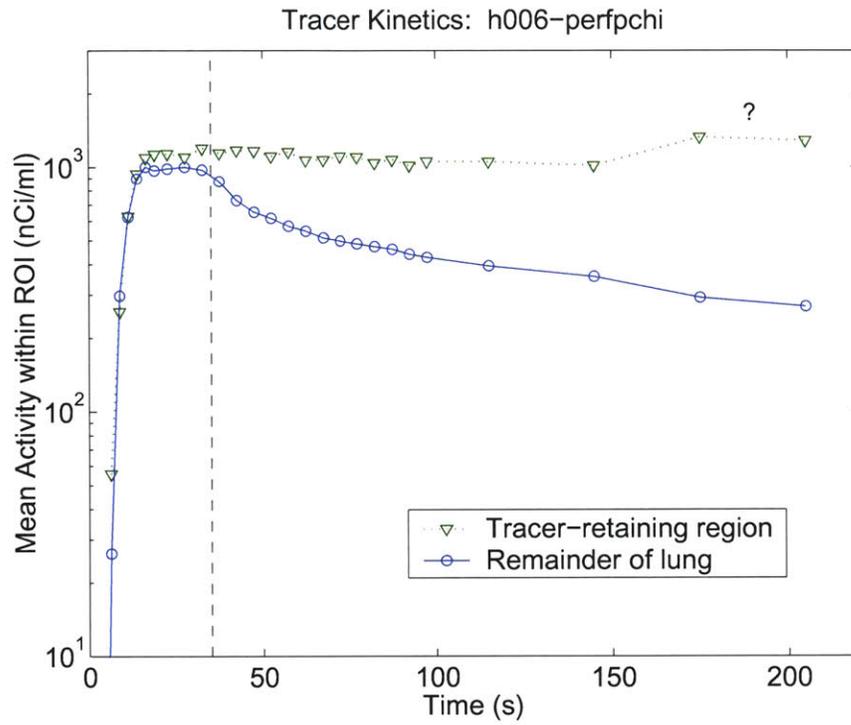


Figure 5.3: Illustration of ‘pop-up’ effect. A tracer-retaining region has been defined as voxels that have high activity in the last two frames. In this region, the activity now appears to ‘pop-up’ at the end of the washout. In fact, we have merely selected a set of voxels in which the noise in the last two frames is positive.

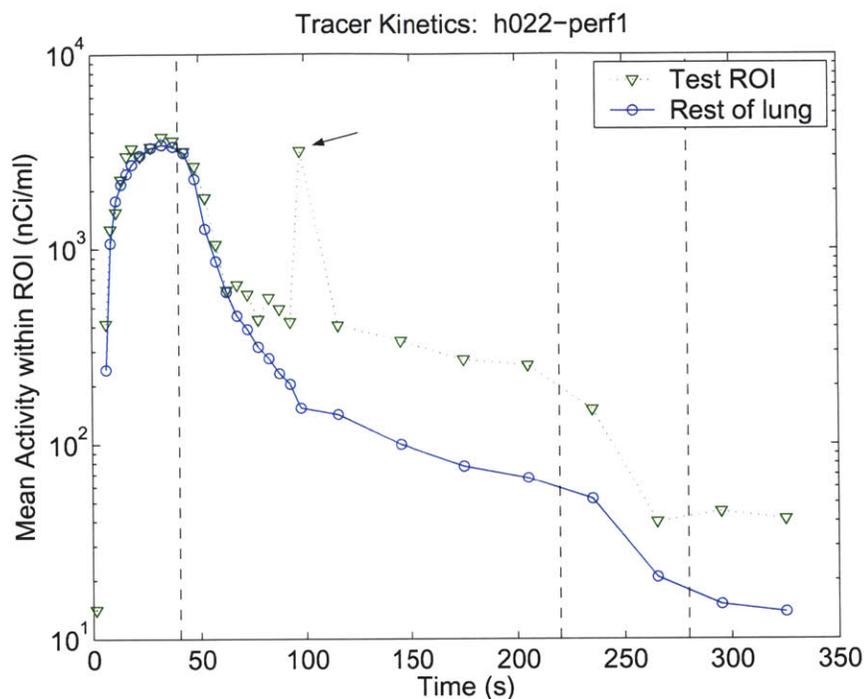


Figure 5.4: Extreme example of ‘pop-up’ effect. Using a 10-second frame, we define an ROI as those voxels with greater than 50% of the maximum activity measured in that frame. Due to the high level of imaging noise, it appears that the tracer content of this region jumps up suddenly by an order of magnitude in this frame.

What has happened here? At first glance, it appears that tracer is redistributing at the end of the washout: the tracer-retaining region acquires more tracer, while the average concentration in the rest of the lung is slightly reduced. This phenomenon is entirely artificial: in reality, we have selected a set of voxels in which (in addition to having high tracer-retention) the noise for these two frames happens to be positive. That the concentration in the remainder of the lung decreases is a manifestation of the unbiased nature of the noise.

This effect is more marked as the amount of noise in a region increases. In the example above, the amount of tracer in the region is large, and the frames involved of long duration. For short frames with low tracer content, the problem becomes huge, and can wreak havoc with attempts to categorize an image. An extreme example is shown in figure 5.4. This graph shows the kinetics of a 216-voxel region of interest that is defined to be those voxels that, in a single 10-second frame, have greater than 50% of the maximum activity measured in that frame. Anomalously, it appears that the tracer content of this group of voxels suddenly increases in this frame by an order of magnitude, before dropping back down in the following frame.

The relevance of this effect will be discussed below.

5.3 Parameter Estimation

Regional parameters may be estimated from the tracer kinetics by fitting an appropriate curve through the measured data, and deducing parameters from the fitted curve. Alternatively, the parameters can be directly estimated by finding the coefficients of a model that result in a best match between model output and measured kinetics. The two approaches are equivalent, as the ‘model’ in the former strategy is implicit in the form of the fitted curve. In either case, the problem becomes, in general, one of non-linear optimization.

Voxel-by-voxel parameter identification is made very difficult by the amount of noise in the measured kinetics. Not only is the signal-to-noise ratio low enough that the curve tends to fit noise rather than signal, but additionally, the actual fit so poorly matches the data that the optimization itself becomes slow and unreliable.

One approach is to combine adjacent voxels into small ROI’s which have cleaner kinetics, and thus more reliable parameter identification. However, this means that we sacrifice spatial information for reduced noise in our parameter estimates.

5.3.1 Non-spatial grouping

An alternative strategy was first proposed by Kimura et al [8]. These authors proposed that parameter identification could be performed on sets of voxels which are grouped not by their spatial proximity, but instead by the similarity of their kinetics – in other words, by their proximity in data space.

Groups of voxels that have similar kinetics are selected, and a set of parameters is identified for each group. Since location is not a factor in the grouping of the voxels, each group is typically scattered around the lung. The parameters identified from the group kinetics are assumed to apply equally to all members of the group.

If we divide the voxels in lung into n groups, then each parameter must take one of the n discrete values identified for each group. Thus spatial information is preserved at the expense of parametric discretization. However, if n is sufficiently large, the discretization is imperceptible.¹

To reiterate: instead of attempting to identify parameters voxel-by-voxel, we find we can make more reliable estimates by analyzing the mean kinetics of a group of voxels. If we group voxels that are physically adjacent, we are effectively increasing the voxel size and degrading spatial resolution. The alternative approach relies on selecting non-contiguous groups based on the similarity of their kinetics. Fundamentally, then, the validity of the method hinges on our ability to select voxels with truly similar behavior.

¹The discontinuous parameter distributions resulting from such analysis can be cosmetically ‘improved’ by low-pass filtering (i.e. blurring) parameter images. Although the validity of this ‘improvement’ is controversial, reasonable arguments, discussed later, can be made to justify such a method.

5.3.1.1 Work of Kimura et al

Kimura et al used this technique to extract metabolic parameters from the kinetics of ^{18}F -deoxyglucose (FDG) in the brain. They treated the kinetics of each voxel of a PET image with m frames as an m -element vector, and used *principal component analysis* (PCA, described in section 5.4 below) to reduce the dimensionality of the data. They then employed a clustering algorithm (Gaussian Mixture Model) to group voxels in m' principal component space. Once clustering was performed, they used conventional parameter identification methods to estimate parameters for each group.

A notable aspect of the study [9] was that the investigators tried two different degrees of dimensional reduction prior to clustering. Firstly they reduced the kinetics to the first two principal components (“2-PC”), and secondly they used as many PC’s as were required to describe 90% of the variance in the measured data (“Many-PC”). They found, unexpectedly, that the 2-PC approach gave better results than Many-PC. The reasons for this are discussed below.

5.4 Principal Component Analysis

Principal Component Analysis (PCA) is a well-known mathematical technique, used primarily to reduce the number of dimensions of a high-dimensional data set. In this method, the data are transformed to a different *basis* – that is, to a different orthogonal coordinate system. The set of basis vectors are ordered, and have the property that the first basis vector is in the direction in which the data has maximal variance. The second basis vector maximizes the remaining variance in the data once the component of the first has been removed, and so on. These vectors are termed the *principal components* of the data set.

Because they are ordered according to the amount of variance in the data, the principal components can be considered to encode the information in the data set in decreasing relevance. By truncating to the early principal components, we can reduce the dimensionality of a data set while minimizing the loss of information.

As a trivial example, for the two-dimensional data set shown in figure 5.5, the black arrow indicates the direction of the first principal component. Suppose that we wish to reduce this two-dimensional set to one dimension: we can do this optimally by taking only the component of the data in the direction of the arrow, which captures the majority of the trend.

5.4.0.2 Mathematics

The principal components are the eigenvectors of the covariance matrix of the data set, sorted by the magnitude of the corresponding eigenvectors.

The data matrix, \mathbb{X} , is oriented with variables arranged in columns, so that each row represents a set of observations.²

²In the application of PCA to transforming tracer kinetics, we treat each frame of an image as a variable, so that data describing n voxels over m frames should be arranged as an n -by- m matrix.

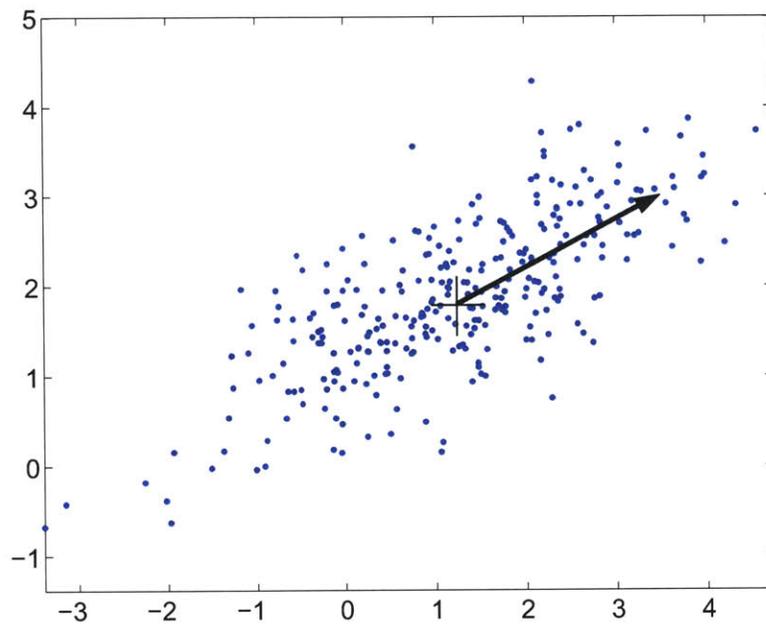


Figure 5.5: Illustration of the use of principal component analysis (PCA) to reduce the dimensionality of a data set. The two-dimensional data set shown above can be optimally reduced to one dimension by taking the component in the direction of the black arrow, which represents the first principal component.

Instead of computing the covariance matrix, which is computationally expensive in both storage and execution time, we can calculate the principal components directly. We first remove the mean values of each column from the data matrix.

$$\mathbb{X}_c = \mathbb{X} - \bar{\mathbb{X}} \quad (5.1)$$

And then use singular value decomposition to generate eigenvectors and eigenvalues. (At this point we are exploiting the fact that the eigenvectors of \mathbb{X}_c are identical to the eigenvectors of the covariance matrix of \mathbb{X} .)

$$\mathbb{U}, \mathbb{L}, \mathbb{P} = \text{svd}(\mathbb{X}_c) \quad (5.2)$$

Where \mathbb{U} and \mathbb{P} are unitary³ matrices, and \mathbb{L} is a diagonal matrix of eigenvalues, so that:

$$\mathbb{U}\mathbb{L}\mathbb{P}^\top = \mathbb{X}_c \quad (5.3)$$

This gives the principal components (eigenvectors) column-wise in \mathbb{P} .

$$\mathbb{P} = [\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n] \quad (5.4)$$

Transforming to feature space is thus simply a matter of computing the scalar product with \mathbb{P} .

$$\mathbb{Y} = \mathbb{X}_c \mathbb{P} \quad (5.5)$$

\mathbb{X} can be trivially reconstructed:

$$\mathbb{X} = \bar{\mathbb{X}} + \mathbb{Y}\mathbb{P}^\top \quad (5.6)$$

To approximate \mathbb{X} in fewer dimensions, we form \mathbb{P}' , by truncating \mathbb{P} from m columns to m' columns.

$$\hat{\mathbb{X}} = \bar{\mathbb{X}} + \mathbb{Y}'\mathbb{P}'^\top \quad (5.7)$$

where $\mathbb{Y}' = \mathbb{X}_c \mathbb{P}'$, the dimensionally-reduced image of \mathbb{X} in feature space. This is the space in which we will cluster our data.

5.4.1 Application to lung imaging

In order to illustrate the application of this PCA/clustering/parameter identification method to lung imaging, let us use consider the analysis of washout kinetics from a bronchoconstricted asthmatic.

Image h006-perfpchi has seventeen washout frames, composed of thirteen five-second frames followed by four thirty-second frames. Within the mask that delineates the lung are 26,826 voxels (reconstructed at 4mm x 4mm x 6.5mm resolution).

³A matrix is said to be *unitary* if its inverse is equal to the transpose of its complex conjugate.

Principal component analysis of this data results in seventeen principal components (eigenvectors) for which the corresponding eigenvalues describe the amount of variance accounted for by each principal component. These eigenvalues (left) and their cumulative sum (right) are shown in figure 5.6.

Data from a second image, h030-perf2, are also displayed. The second image was selected because there is a lower amount of noise in the data. (The decrease in noise is due to the fact that more radioactivity was infused, and because the chest-wall of subject h030 was thinner than h006.)

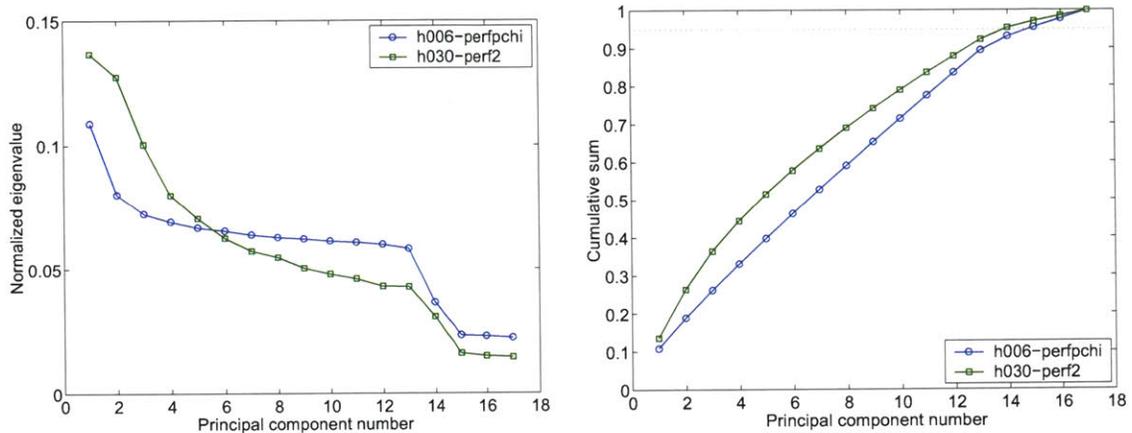


Figure 5.6: Eigenvalues from PCA of two different images. The normalized eigenvalue is the fraction of the total variance in the data accounted for by the corresponding principal component. These are plotted on the left. On the right, the cumulative sum of the eigenvalues is plotted.

Each successive component is seen to describe a smaller fraction of the total variance. It is clear that for both of these images, a large number of components are required to satisfactorily describe the data. To describe more than 90% of the variance, 14 components are required for h006-perfpchi, and 13 for h030-perf2. In general, we see that as the amount of noise in a PET image decreases, so the rate at which successive eigenvalues get smaller increases: less noisy images are better described by fewer principal components.

Since we model the washout of each voxel with four parameters [section 5.6.1], it is surprising that so many dimensions are required to describe the data. Is our model unrealistic? The answer lies in the principal components themselves. These are shown for each washout in figure 5.7.

In both images only the first two principal components describe a meaningful underlying trend in the data (and arguably only the first in the case of h006-perfpchi, the left-hand, noisier image). Thereafter, the shapes of subsequent components are incompatible with any expected kinetics, and only characterize noise in the image.

This is why Kimura et al achieved better results with the 2-PC method than with Many-PC. In the former case, voxels were grouped using the the first two principal components; in the latter, the authors used enough PC's to explain 90% of the variance in the image (approximately X PC's). The

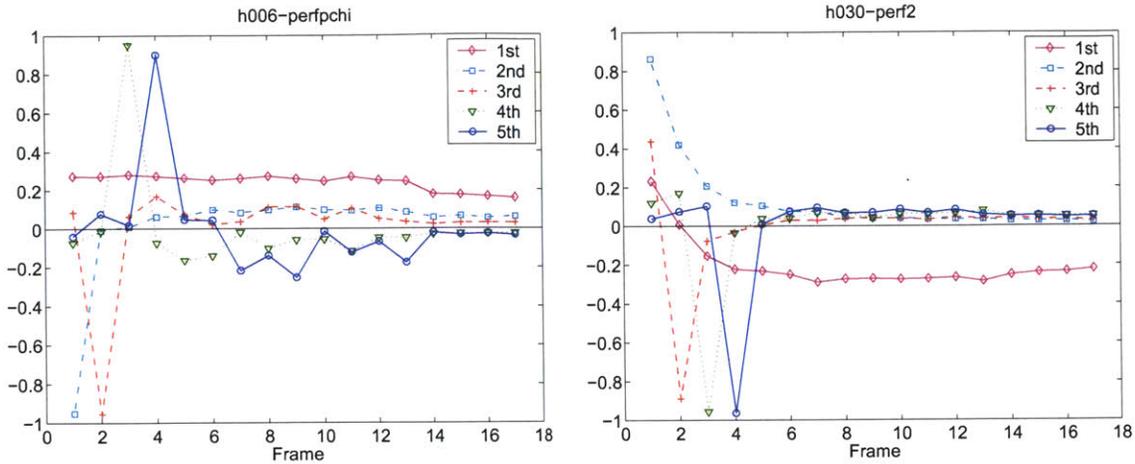


Figure 5.7: The first five principal components of washout frames from two different images, both of bronchoconstricted lungs. In both cases, only the first two principal components actually describe any meaningful trend in the data: subsequent components only characterize imaging noise.

reason the latter approach performs worse is that, by including more PC's, voxels are grouped not by similar underlying kinetics, but by similar noise.

In the example of h003-perf2, the inclusion of the 3rd principal component (see figure 5.7) means that we will start to group according to whether voxels happen to have noise which is highly negative in the second frame. This effect is analogous to the 'pop-up' phenomenon described above. In the 'pop-up' example, we inadvertently selected a population in which the activity appeared to increase in the last two frames. Here, if we were to select voxels by the 3rd component, we would select a population that appears to have a sudden dip in activity in the second frame.

5.4.2 Principal components from synthetic data

The viability of non-spatial grouping relies on our ability to group voxels that have similar underlying kinetics, while avoiding as much as possible the grouping of voxels that have similar noise characteristics. This is a delicate balance. If we use large numbers of principal components, we end up grouping by noise. If we use too few, we discard so much information that we group voxels that have fundamentally different kinetics, in which case it is invalid to ascribe identical parameters to all of the voxels in a group.

One problem with PCA is that this is a 'blind' operation. It is optimistic to hope that the process will automatically and magically discriminate between signal and noise based merely on variance in the data. Ideally, we should take advantage of our knowledge of the system, and use this information to group voxels more intelligently.

5.4.2.1 Concept

The following enhancement is proposed to the method of Kimura et al. We artificially generate a ‘perfect’ image, in which each voxel has noise-free kinetics defined by the same model used in the parameter-identification step. Instead of calculating the principal components of the experimental data, which results in PC’s that describe image noise, we calculate the principal components of the synthetic image. These PC’s only describe (what we expect to be) the ‘signal’ characteristics. We then transform the experimental image to the PC’s derived from the synthetic data, and group voxels by their proximity in this space.

What this does, in a sense, is to consider only those aspects of the system’s behavior that are consistent with our model. We ignore any departure from this behavior, assuming this to be noise.

One might worry that this is limiting our analysis: if our model is invalid, or is at least an oversimplification, then we will be discarding real information about the system. However, this is already done at the parameter identification step: our model is implicitly imposed at the point where we estimate parameters from the kinetics of each group. By imposing the model earlier in the analysis, we are losing nothing, and arguably improving the consistency of the method.

5.4.2.2 PCA of synthetic data

In order to generate a synthetic image, we make preliminary parameter estimates for each voxel within the lung, discarding physically implausible values (like positive time constants) and replacing these with values interpolated from neighboring voxels. Then we form an artificial emission image using these parameters to generate ‘noise-free’ kinetics.

Continuing our example, the normalized eigenvalues for synthetic data derived from h030-perf2 are shown in figure 5.8, alongside the values for experimental data.

In the case of noise-free data, we see that the earlier principal components dominate. Only three components are required to account for more than 90% of the variance; for experimental data, thirteen are required. This should reinforce our conviction that later principal components describe noise rather than signal.

The shape of these components is also promising: they have a form that appears more appropriate for describing washout kinetics.

5.4.3 Weighted PCA

The principal components describe directions of maximal variance in a data set. In our analysis, we hope that these represent ‘signal’ in the data. And indeed, that the technique works at all reflects the approximate truth of this assertion. However, to identify components which *optimally* achieve this goal, we need to consider other aspects of our experimental data.

The duration of the frames that comprise an image is not constant. We typically acquire 30-second frames at the end of washout (when the amount of activity in the lung is small), and much shorter (5

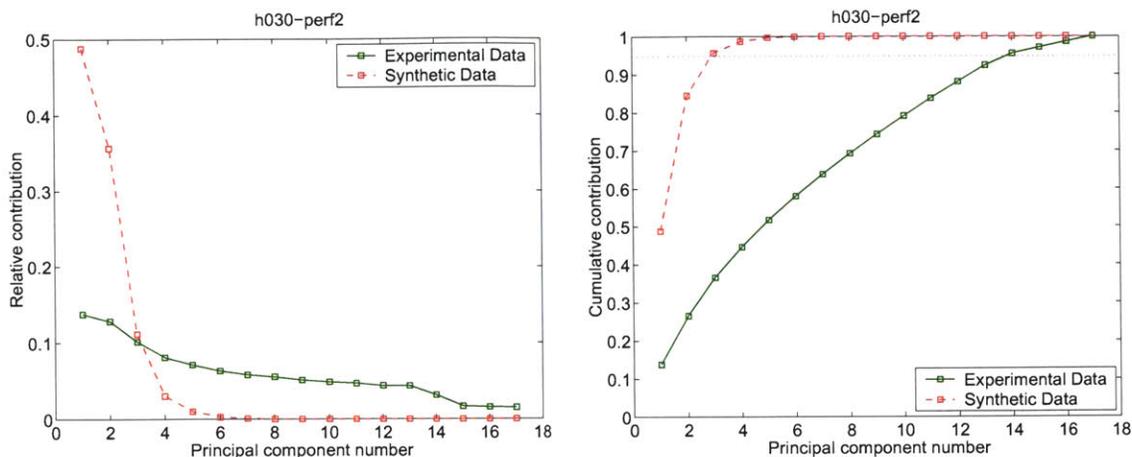


Figure 5.8: Eigenvalues from PCA of experimental data and synthetic noise-free data derived from the same image. In the case of the synthetic data, three principal components are able to account for >90% of the variance, whereas fourteen are required for experimental data.

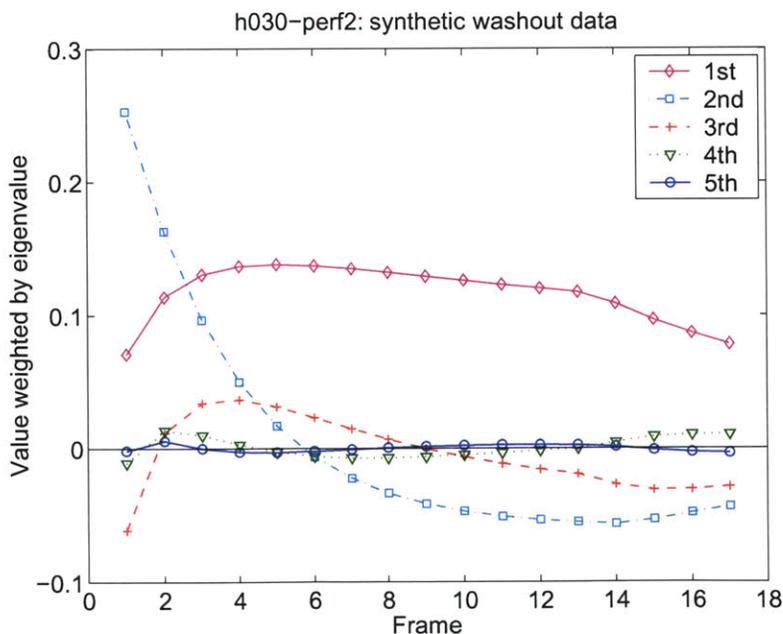


Figure 5.9: Weighted principal components of synthetic, noise-free washout data. These are smoother than those generated from experimental data (Figure 5.7), and have shapes more appropriate to kinetic trends.

or 10-second) frames earlier in the image (in order to capture faster transients). As a consequence of the variable frame durations and the changing amount tracer in the lung, the amount of noise in each frame is not constant, and this should be taken into account when we calculate principal components.

The variance of a frame is the sum of the variance in tracer content and the variance of the noise: a frame may have high variance because there is a lot of noise but relatively uniform tracer content, or because there is a lot of heterogeneity in tracer content and little noise. In the latter case, such a frame contains a lot of useful information, and should be weighted strongly in our analysis; in the former case, we should give the frame little weight, to diminish the significance of this data.

We can accomplish this goal by rescaling the data set prior to PCA. Each frame is multiplied by a weight, w_i . The weighted data set is transformed to its principal components, and voxels are grouped as before. The obvious scheme is to weight by the inverse of the standard deviation of the expected noise, $w_i = 1/\sigma_i$.

Application of this weighting scheme to experimental data imperfectly compensates for imaging noise, which is additive rather than multiplicative in nature [Section 1.5]. If instead we are calculating principal components from a synthetic data set, this weighting works extremely well, as the PCA now identifies directions which maximize signal-to-noise in the data.

5.5 Grouping

In their papers, Kimura et al talk about *clustering* their data following principal component analysis. In this thesis, a more neutral term, ‘grouping’, is used. Most clustering algorithms (e.g. K-means, mixture gaussian) are designed with the presumption that the data has an innately clustered structure, which the algorithm seeks to extract.

In the case of PET data from the lungs, however, there does not seem to be any such clustered structure [see, for example, figure 5.15]. Even in the case of data from a highly-bronchoconstricted subject, where one might imagine a bimodal distribution of voxels that either retain tracer or washout normally, no saddle is evident in the histogram of tracer retention. [Figure 5.10.]

Accordingly, it makes sense not to try to identify clusters in the data, but instead merely to partition the data orthogonally in feature space. Orthogonal partitioning is described in the following section.

5.5.0.1 Orthogonal partitioning

Our goal, loosely specified, is to divide the voxels of an image into n groups, with roughly equal number of points in each group, and where the points within a group are close to each other (in feature space), with no overlap between groups.

We form orthogonal partitions in feature space, which are spaced so that the number of points between adjacent partitions is equal. In order to form groups that are roughly square, we weight the

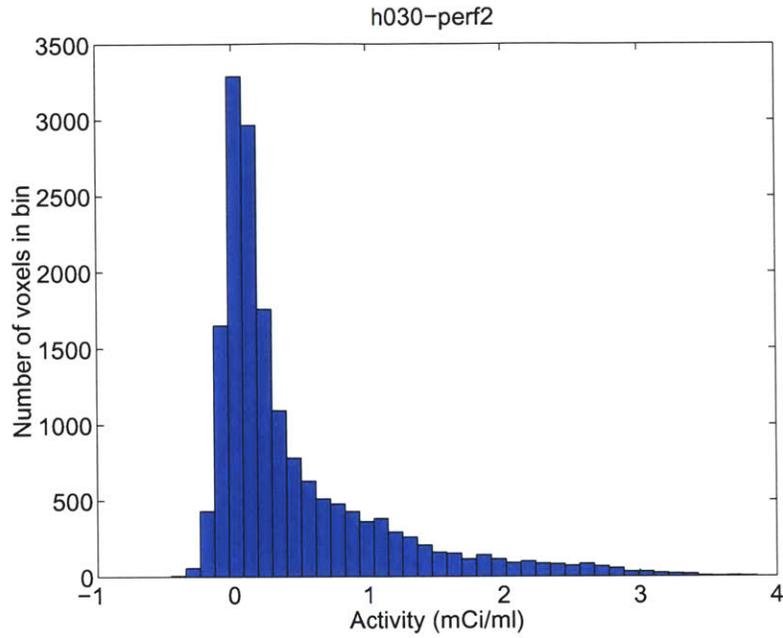


Figure 5.10: Histogram of tracer retention in a bronchoconstricted subject. Although the degree of retention is substantial, the distribution is not visibly bimodal.

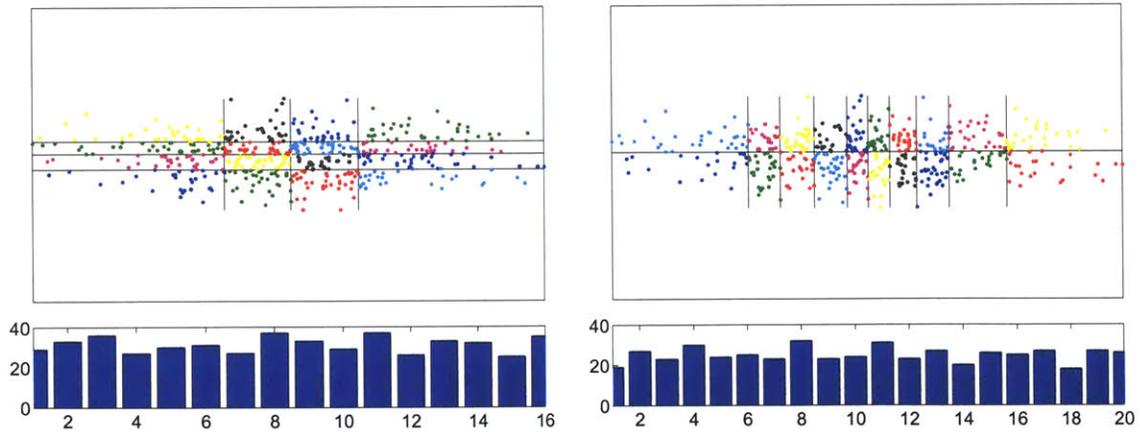


Figure 5.11: Example of partitioning an artificial, two-dimensional data set of 500 points into approximately 20 groups. On the left, we partition into equal numbers of groups along each axis. Because the data has much greater standard deviation along the x -axis, it is better to form commensurately more partitions in this direction (right). Underneath each plot is a histogram showing the number of points in each group.

number of partitions along an axis by the standard deviation of the data in that direction. [Figure 5.11.]

In this case, our goal to form groups with equal membership is reasonably well achieved. However, with this method groups around the corners of the grid are often under-filled. A worst-case example is shown on the left of figure 5.12. Here the same data set as in figure 5.11 is rotated through 45° . The partitioning is now perfectly mis-aligned with the data set, and the corner groups are either empty or over-filled. We can ameliorate the situation by performing PCA (again!) to identify the most appropriate partitioning direction.

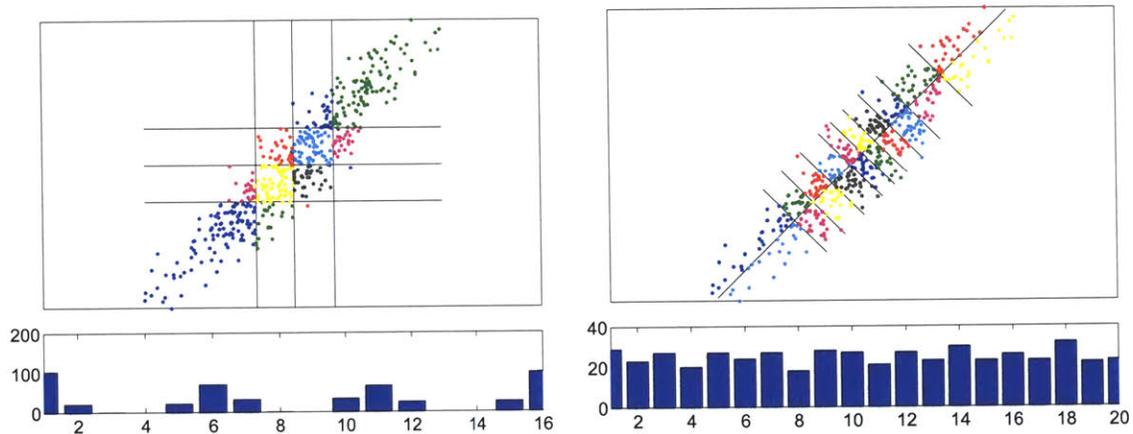


Figure 5.12: On the left, a demonstration of poor partitioning because the partition axes are mis-aligned with the data, and group membership is highly variable. On the right, we partition along the principal components of the data.

A repeated principal component transformation may seem redundant. If the data set was previously transformed to its principal components, the second transformation will have no effect. (The operation is analogous to describing a mechanical system in terms of its fundamental modes, to diagonalization of a state-space representation. Once the description is in this form, it cannot be reduced further.)

However, if we previously transformed to PC's from a synthetic data set, then the data was not transformed to its own PC's, and consequently, the second transformation will be useful. Note that we are not in any sense undoing the original transformation (to synthetic PC's), as we have discarded all but the first two or three dimensions. (Had we retained all dimensions of the original data, then the second transformation would eclipse the first.)

5.5.0.2 Partitioning weighted by eigenvalue

With synthetic data, the relative contribution of the principal components declines rapidly. In the example shown in figure 5.8, the eigenvalue of the third PC is around a quarter of the first. When

the experimental data is transformed to this synthetic basis, the standard deviation in the 3rd dimension is about 0.6 times the deviation along the 1st. That is, real data has more variability in this component than does ideal data. If our system model is accurate, this increase must be due to noise.

As described above, when we partition the data, the default behavior is to choose the number of partitions along each axis according to the standard deviation of the transformed data in that direction. (See figure 5.11.) This allows us to form groups that have approximately equal extent in each dimension. In the case of data transformed to a synthetic basis, it makes more sense to choose the number of groups according to the eigenvalue of the synthetic data: if the first eigenvalue is four times larger than the third, then we should have four times as many divisions along the first dimension as along the third. If we form too many partitions in the direction of a component with low eigenvalue, we will ‘over-fit’ the the data, and end-up grouping by noise rather than signal.

5.5.0.3 Underpopulated groups

The feature space into which we have transformed the data is a space of expected kinetics, which we would like to sample uniformly in order to group voxels with similar behavior. This is difficult to achieve with clustering algorithms, but is trivially accomplished with orthogonal partitioning (a principal advantage of the method).

Because it is minimally adaptive to the data set, one of the deficiencies of this grouping method is that some of the peripheral groups may be underpopulated, despite the enhancements discussed above. In such circumstances, we have a choice between accepting the underpopulated group (which will have noisy kinetics, and hence may generate unreliable parameter estimates), or joining these points to an adjacent group (in which case we will begin to group voxels that have significantly different kinetics).

5.6 Parameter Estimation

The third and final stage of the technique is the estimation of group parameters from the kinetics of each group.

Necessarily, the detail of this step depends on the system being studied, and thus the model underlying the kinetics. Up to this point, our description of the technique has been very general: here, however, we need to discuss issues that are specific to the analysis of tracer washout kinetics.

As discussed in section 4.7.2.2, we invariably observe multi-compartmental washout kinetics in diseased lungs. In asthmatics, bronchoconstriction is not uniform, but varies dramatically across the lung. Furthermore, it appears that this heterogeneity develops at length scales below the resolution of the PET camera used (~6mm). This conclusion is supported by the observation that even very small ROI's exhibit multi-compartmental washout kinetics (i.e. multiple ventilation rates).

Accordingly, we need to model the washout of every group using at least two exponential terms. It is not feasible, however, to attempt to extract more than two time-constants from a washout curve.

In general, the improvement in residual over a two-compartment model is slight, and we encounter increasing problems with over-fitting. A better strategy is to perform two-compartment analysis at higher resolution or large number of groups.

5.6.1 Washout model

According to the theory developed in section 4.5, the washout of a single compartment should follow a simple exponential decay, provided that no additional tracer arrives in the lung at the end of the apnea period. The apparent washout rate, $1/E'$, is the sum of the tracer resorption rate, $1/E$, and the specific ventilation, $s\dot{V}$.

Separation of these two components requires either analysis of the tracer resorption during the apnea phase, or that we predict the resorption rate from knowledge of perfusion and lung inflation [Section 4.7.2.1]. Because these quantities are not separable using washout information alone, we need only concern ourselves with estimating E' .

We have the following description of the washout of aerated lung [Equation 4.33] :

$$C_A(t) = C_A(0)e^{-t/E'} \quad (5.8)$$

Where $C_A(0)$, the concentration at the start of the washout, is proportional to perfusion. [Equation 4.23]

$$C_A(0) \sim \frac{A_{inf}\dot{Q}_A}{\dot{Q}_T V_{ROI}} e^{-t_{ap}/E} \quad (5.9)$$

For two compartment analysis we thus need to fit a curve of the form:

$$\hat{C}(t) = A_1 e^{-t/E'_1} + A_2 e^{-t/E'_2} \quad (5.10)$$

In practice, it proves more convenient to analyze in terms of washout rates rather than time constants.

$$\hat{C}(t) = A_1 e^{-\alpha_1 t} + A_2 e^{-\alpha_2 t} \quad (5.11)$$

5.6.2 4-parameter fit

There are four parameters ($A_1, \alpha_1, A_2, \alpha_2$) in equation 5.11 to be estimated.

The first strategy explored is to estimate these parameters by minimizing a least-squares cost function, E , using the Levenberg-Marquardt algorithm.

$$E = \sum (C_i - \hat{C}_i)^2 \quad (5.12)$$

The only modification we make is to add per-frame weighting to compensate for the non-uniform significance of the different frames. The justification here is similar to that for weighting the PCA [section 5.4.3].

$$E_w = \sum w_i (C_i - \hat{C}_i)^2 \quad (5.13)$$

5.6.3 Initial parameter estimates

The convergence of optimization algorithms on a global minimum is highly dependent on starting with reasonable initial estimate of the parameters.

Initial estimates of parameters A_2, α_2 are made by fitting a single exponential decay to the last two minutes of the washout kinetics. (The implicit assumption being that the fast compartment has completely washed out by this point.) The slow compartment thus estimated is subtracted from the washout kinetics, to leave a residual to be explained by the fast compartment. A single exponential is fitted to the first minute of the residual, yielding an initial estimate of A_1, α_1 .

5.6.4 3-parameter fit

One deficiency of 4-parameter analysis is that while we generate an optimal fit to washout curves, this fit is not necessarily the best approximation to the actual behavior of the tracer. We know that the washout should start with the same tracer concentration that remained at the end of the apnea period, but this is not a constraint in the 4-parameter description.

A possible improvement to this method is to make use of the apnea data, and rewrite equation 4.33 with the constraint that the concentration at the start of the apnea, $C(0)$, must equal the end-apnea level, A_{ap} . Since at $t = 0$, $\hat{C}(t) = A_1 + A_2$, this constraint is trivially enforced by requiring that $A_2 = A_{ap} - A_1$.

$$\hat{C}(t) = A_1 e^{-\alpha_1 t} + (A_{ap} - A_1) e^{-\alpha_2 t} \quad (5.14)$$

This modification also has the beneficial effect of reducing the number of parameters to be estimated by the optimization algorithm from four to three (A_1, α_1, α_2).

5.6.5 Direct amplitude estimation

Both the principal component analysis and the grouping of the data are very rapid, taking no mere seconds on a modern PC. The overwhelming computational burden of the analysis is in the group-by-group parameter estimation. The cost of this operation increases linearly with the number of groups chosen, but exponentially with the number of parameters that need to be estimated by the optimization algorithm.

The washout equation (4.33) is non-linear in terms of washout rates α_1, α_2 but linear in the amplitude terms A_1, A_2 . We can make use of this property by observing that if we fix α_1, α_2 , we can analytically calculate the amplitudes that will give the best fit in a least-squares sense.

At the error minimum, we expect that $\frac{\partial E}{\partial A_1} = 0$ and that $\frac{\partial E}{\partial A_2} = 0$. This is equivalent to:

$$\begin{bmatrix} \sum e^{-2\alpha_1 t_i} & \sum e^{-(\alpha_1 + \alpha_2) t_i} \\ \sum e^{-(\alpha_1 + \alpha_2) t_i} & \sum e^{-2\alpha_2 t_i} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} \sum C_i e^{-\alpha_1 t_i} \\ \sum C_i e^{-\alpha_2 t_i} \end{bmatrix} \quad (5.15)$$

Given α_1, α_2 , we can calculate A_1, A_2 directly by solution of this equation. This allows us to improve the performance of the parameter estimation considerably, as the optimization algorithm is

now searching for only two parameters. The time required to analyze a set of groups is reduced approximately threefold using this approach.

To minimize a weighted error norm, E_w [equation 5.13], this relation is trivially extended.

$$\begin{bmatrix} \sum w_i e^{-2\alpha_1 t_i} & \sum w_i e^{-(\alpha_1 + \alpha_2) t_i} \\ \sum w_i e^{-(\alpha_1 + \alpha_2) t_i} & \sum w_i e^{-2\alpha_2 t_i} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} \sum w_i C_i e^{-\alpha_1 t_i} \\ \sum w_i C_i e^{-\alpha_2 t_i} \end{bmatrix} \quad (5.16)$$

5.6.6 ‘Magic Point’

As mentioned above, we expect the tracer concentration at the washout start to be continuous with the end of the apnea, and should try to make use of this property in our analysis of the washout.

The 3-parameter method described in section 5.6.4 has a couple of disadvantages. Firstly, the apnea data, like the washout data, is noisy, and constraining the fitted washout curve to intersect an unreliable point degrades the quality of the analysis.

A second problem is that there is often some visible discontinuity at the start of the washout. The simplest explanation for this is that the lung is stationary in apnea, but in motion during the washout. It is easy to imagine a voxel (whose location is fixed in space) that is within the lung during apnea but only partially occupied by lung tissue during tidal breathing: the measured concentration in such a voxel will abruptly drop at the start of the washout. Such discontinuities ruin the curve-fitting, which generates an erroneous ultrafast compartment to describe the rapid change at the start of the washout. (A similar effect is illustrated in figure 5.13.)

The former problem may be resolved by adding to the start of the washout curve a special, ‘magic’ point which represents the apnea data, and then reverting to a full 4-parameter fit. We can assign weight to this point in proportion to duration of the apnea plateau and/or the amount of activity in these frames. In this way, we include the apnea information in the washout analysis, but importantly, do not constrain the fit to intersect this point. Consequently the analysis is made robust to noise in the apnea frames.

This modification alleviates, but does not fully resolve the discontinuity problem.

5.6.7 Single-Compartment Fit

Inspection of group kinetics reveals that not all groups exhibit discernible two-compartment behavior. In such groups we need to be cautious at the parameter estimation stage, as we run the risk of over-fitting our data.

An example of this is shown in figure 5.13. Here the kinetics are well described by a single, slow compartment. A small positive deviation at the start of the washout, which is almost certainly noise, is now identified as a second compartment with an ultrafast washout rate. When the group parameters are redistributed in the lung, this washout rate is clearly anomalous, being much faster than any other. Such over-fitting errors, while seemingly slight, can introduce significant errors when estimating blood gases: in this example, the ‘fast compartment’ is not only ventilated impossibly

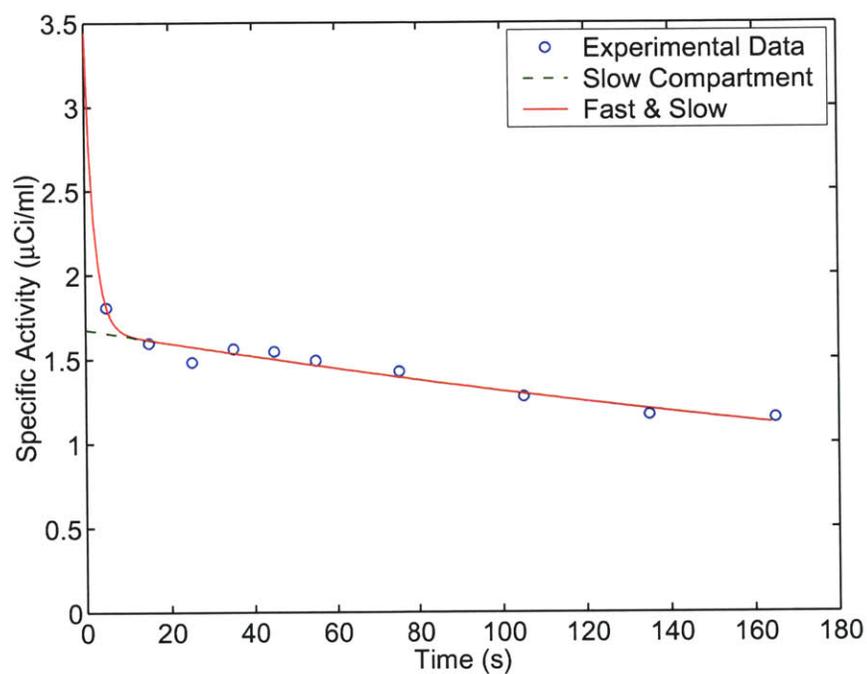


Figure 5.13: An example of the problems of ‘over-fitted’ kinetics. The data are well described by a single, slow compartment; as a consequence, the small positive deviation of the first point (which is almost certainly noise) is identified as a second compartment with very fast washout rate.

fast, but is apparently well perfused (as the amplitude assigned to this compartment is also sizable) and will thus contribute substantially to blood oxygenation.

Setting realistic limits on ventilation rates only reduces the magnitude of the problem. In the example above, this modification would merely threshold the fast washout rate to the upper limit, and decrease the corresponding amplitude.

A better approach is to examine the improvement in curve-fit achieved by transitioning from a single to a double-exponential fit. In the example above, the improvement in the residual (decrease in error) is tiny – including the fast compartment only improves the fit very slightly. Under such circumstances, we should revert to single-compartment estimation.

In our analysis, a single-compartment fit is employed if the following condition is satisfied:

$$\frac{\text{residual}(\text{single}) - \text{residual}(\text{double})}{\text{residual}(\text{mean})} < 0.05$$

In this definition, ‘residual’ is the weighted error norm used for the curve-fitting:

$$\text{residual}(\hat{y}) = \sum_i w_i (y_i - \hat{y}_i)^2$$

So that ‘residual(mean)’ thus denotes a weighted variance: $\sum_i w_i (y_i - \bar{y})^2$.

In practice, it has proved useful to employ a single-compartment fit in two other circumstances. Firstly, if the amount of tracer remaining in the last two minutes of the washout is negligible, then a single-compartment fit is used (without any comparison of residuals with a two-compartment fit). Secondly, if the two-compartment fit yields a fast compartment with a negative amplitude (i.e. kinetics which rise at the start of the washout), then the fast compartment is discarded, and only the parameters for the slow compartment are used. [But see also section 5.8 for a discussion about one factor that may be responsible for this phenomenon.]

5.7 Example

To demonstrate the implementation, let us analyze image h030-perf2, a bronchoconstricted asthmatic, using the techniques described above.

We start by generating a synthetic washout image which has a similar parameter distribution to the original data, but has no noise, so that each voxel follows a perfect double-exponential decay. Principal components are calculated from this ‘reference’ image.

The target image, h030-perf2, is then transformed to a basis of the first principal components. The useful number of PC’s for analysis of this type of data is invariably either two or three. With only a single PC (one-dimensional partitioning) we are discarding too much information to group sensibly, and will end up grouping voxels that have heterogeneous kinetics. With more than three PC’s, we are guaranteed to be grouping by noise. This can be seen by consideration of figure 5.8: for

noise-free, synthetic kinetics, the contribution of successive PC's drops very rapidly, with the fourth PC describing less than a twentieth as much information as the first.

For ease of presentation, I will follow a two-dimensional partitioning scheme here.

Generation of synthetic principal components. The first stage of the analysis is to generate a noise-free image which has a similar parameter distribution to the target image. Since at this stage a full voxel-by-voxel analysis is impractical, we make several strong assumptions.

Firstly, using the global kinetics curve (sum of all the voxels) we estimate an 'average' slow-compartment washout rate, α_2 , and assume this is uniform throughout the lung. Next, we look at the amount of activity remaining in each voxel at the end of the washout. Using the fixed slow-compartment washout rate, α_2 , we calculate a per-voxel slow-compartment amplitude A_2 , that would match the observed residual activity at the end of the washout (so implicitly assuming that the fast compartment has negligible contribution). Thirdly, we assume that the perfusion data for each voxel (mean of apnea plateau frames, $\overline{A_{ap}}$) is noise-free, and constrain all washout curves to start at this point (i.e. we require $A_1 + A_2 = \overline{A_{ap}}$), which allows to calculate a per-voxel A_1 . Finally, we use the washout integral of each voxel to calculate the remaining parameter, α_1 .

The huge amount of noise in single-voxel kinetics mean that even with this simplified analysis, our parameter estimates are not only unreliable, but often take impossible values. For example, if the apnea frames have negative mean, so that $\overline{A_{ap}}$ is negative, then we will calculate nonsensical values for A_1 . In such cases, when parameters take impossible values, we flag the voxel as 'bad', and calculate new parameter values by interpolating from neighboring voxels.

Discarding deviant data is, in general, hard to justify. We may be ignoring abnormal but real kinetics, and we are certainly biasing the estimation of remaining parameters. Keep in mind, however, that at this stage we are only trying to make a synthetic image from which to calculate principal components. In the full analysis, we shall be careful not to arbitrarily exclude such unusual behavior.

Having calculated plausible parameter sets for each voxel, we can easily generate the washout kinetics of our synthetic reference image from these parameters [equation 4.33]. We then calculate PC's and their contributions as described in section 5.4.0.2. For our example, h030-perf2, the first two principal components are shown in figure 5.14. In this case the PCA has been weighted by frame duration, as discussed in section 5.4.3.

Transformation and Grouping Having generated principal components from a reference image, we now transform our data set to the truncated basis of the first two principal components, and partition the data [Fig. 5.15], choosing the number of partitions along each axis in proportion to the corresponding eigenvalue. [See section 5.5.0.2.] In this example, this partitioning divides the data set into around 100 groups.

Parameter Estimation The mean kinetics of these groups are shown in figure 5.16. Four parameters are estimated for each group, except where a single-compartment analysis is preferable [section

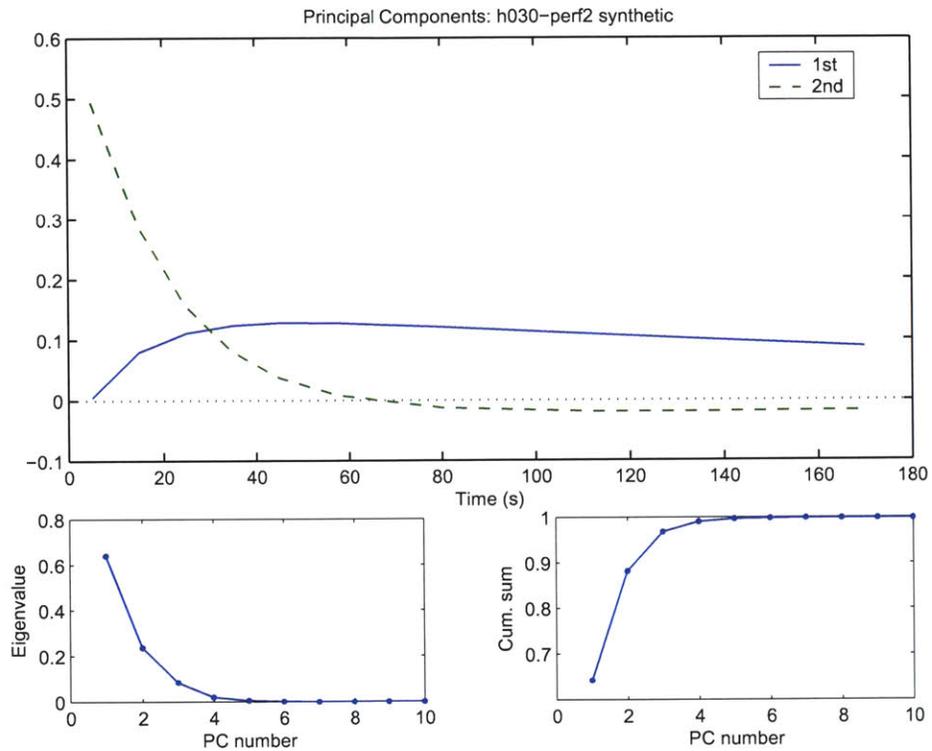


Figure 5.14: Principal components derived from a synthetic washout image. Shown are the first two principal components (top); the eigenvalues (below left), indicating the relative contribution of the components to variation in the synthetic image; and the cumulative sum (below right) of these eigenvalues.

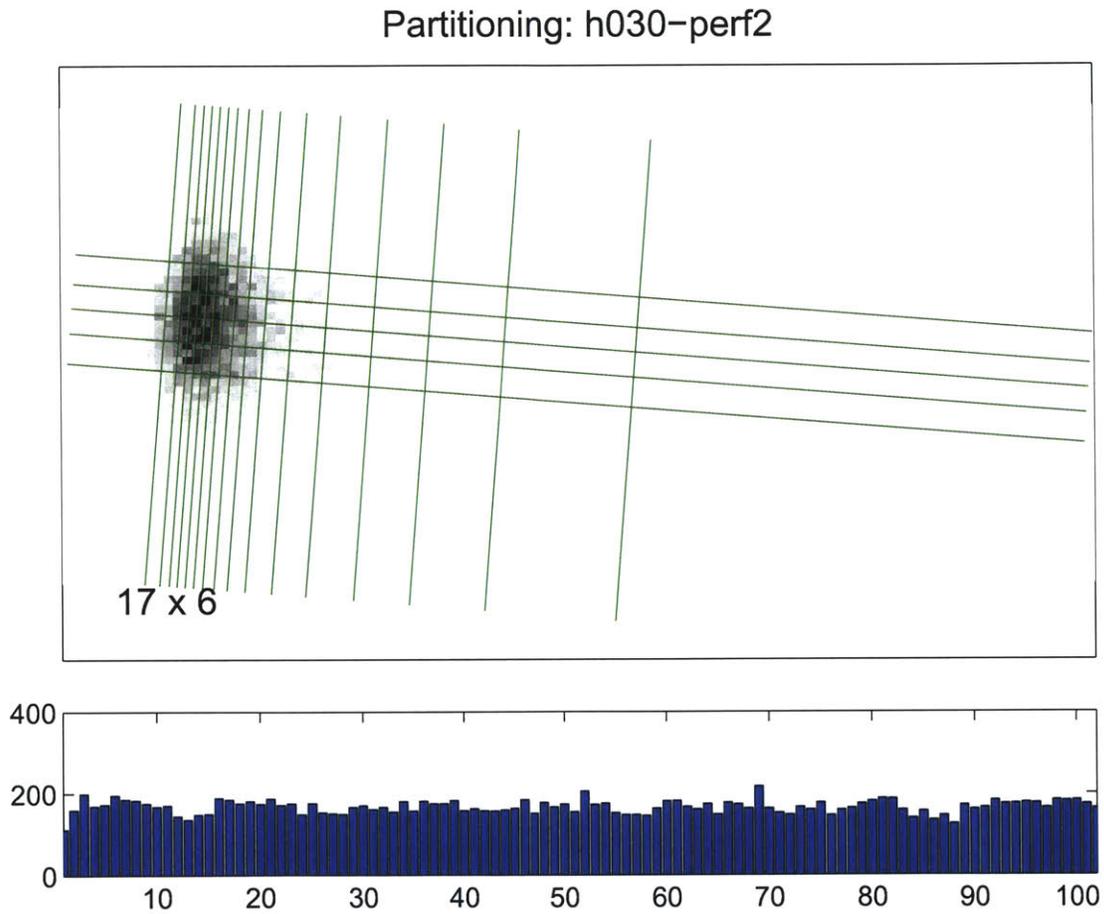


Figure 5.15: Top: Partitioning of image h030-perf2, transformed to principal components derived from a synthetic washout image. The histogram below shows the number of voxels assigned to each group.

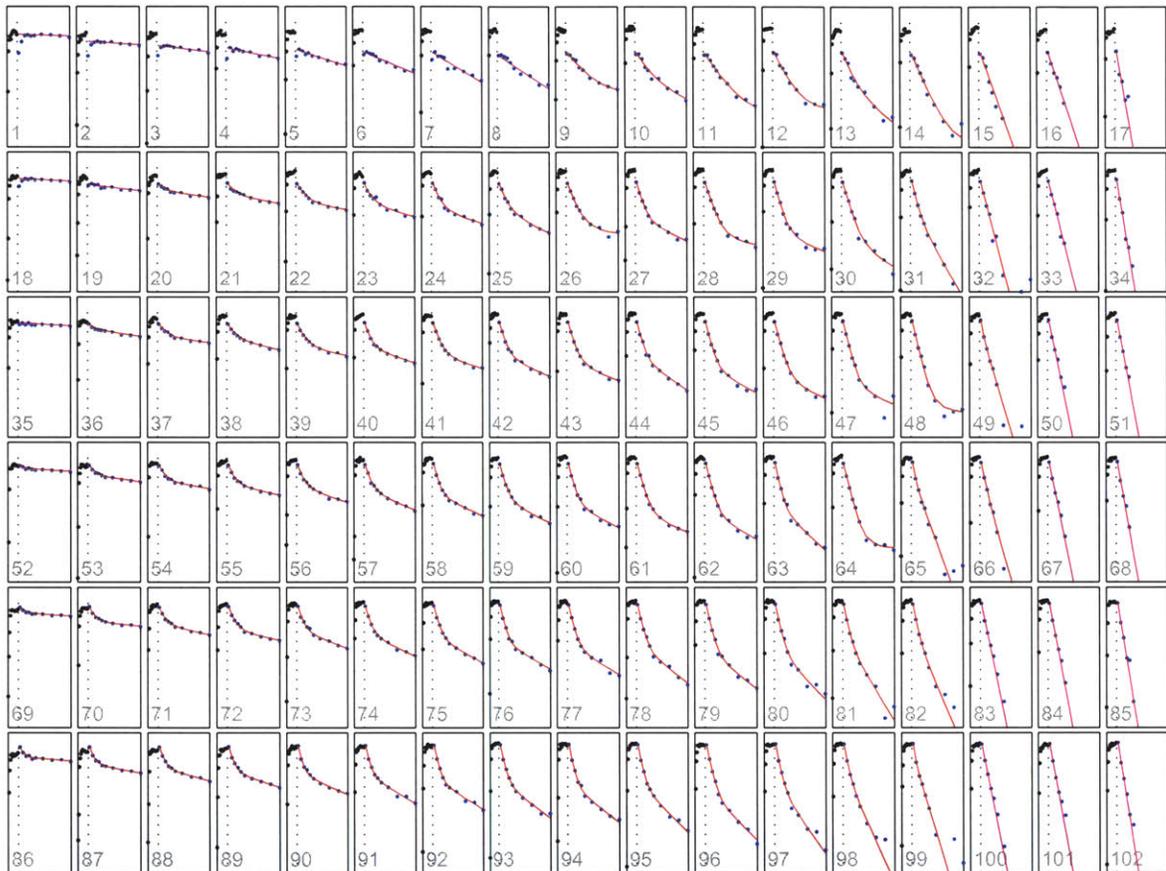


Figure 5.16: Tracer kinetics and curve-fits for the groups defined by the partitioning of h030-perf2 illustrated in figure 5.15. The dotted vertical line indicates the washout start; apnea data (not included in this analysis) are shown as black dots; washout data as blue dots; two-compartment curve-fits as red lines and single-compartment fits as magenta lines. The logarithmic vertical scale (not shown) covers three orders of magnitude from 10nCi/ml to $10\mu\text{Ci/ml}$.

5.6.7]. The corresponding curve-fits are shown on the same diagram, with double-compartment fits shown as red lines, and single-compartment fits as magenta lines. The apnea kinetics are also shown, but these data were not included in the parameter estimation, nor in the PCA (see next section).

It is evident from this graph that the method is capable of separating groups with very different kinetics, and that the model describes the group kinetics very well.

The estimated group parameters, redistributed back to their individual voxels, are displayed individually in figure 5.17. These separate parameters can be somewhat hard to comprehend, as there is considerable overlap between fast and slow compartments. A well-ventilated group which nevertheless has a bi compartmental washout may yield two ventilation rates which are both in the range which would be considered normal. In such a case, the faster rate will be assigned to α_1 and the slower to α_2 . This latter rate is much faster than most slow compartments, and consequently may appear as an anomalous ‘bright dot’ on the image of this parameter. The corresponding perfusion may also appear abnormal if it lies in a region where A_2 has been tapering to zero. These sort of effects are responsible for a good deal of speckle on the parameter images which appears to be noise.

The results of analysis of asthmatic subjects are discussed in detail in chapter 6. However, the following brief observations are illustrative of the sort of insights that can be gained from these images. In the image of parameter α_2 , we see a large, apparently homogeneous region of hypoventilation in the dependent (dorsal) regions of the lung, and while the corresponding perfusion, A_2 , is lower than normal, it is still significantly above zero. This blood is essentially shunted without oxygenation through the lung, and will result in reduced saturation in the arterial circulation. Interestingly, the fast compartment has very uniform ventilation throughout the lung, including this dependent region. Here though, the corresponding perfusion, A_1 , is diminished. This sort of bimodal ventilation pattern, with a fast compartment ventilated at normal rates and a more variably-constricted slow compartment, is typical of asthmatic lungs.

5.7.1 Single-Compartment Groups

Groups for which single-compartment analysis was selected [5.6.7] pose a problem of presentation. For these groups we have estimated not four parameters, but two, and leaving the other two parameters unassigned leaves unsightly ‘holes’ in the parametric images. The following remedial post-processing steps are proposed.

If the washout rate estimated from a single-compartment group lies within the overlap range between fast and slow compartment rates, then we can consider this group to be either a fast or slow group. In this case, we assign the identical ventilation rate to both compartments, α_1 and α_2 . Images of parameters A_1 and A_2 , with single-compartment results excluded, are then interpolated to fill values for this group. We then redistribute the single-compartment perfusion in proportion to the interpolated values for A_1 and A_2 (which are different for each voxel in the group). This operation has the effect of redistributing a single-compartment that could be considered fast or slow between the two compartments in the way that is most ‘harmonious’ with the neighboring data.

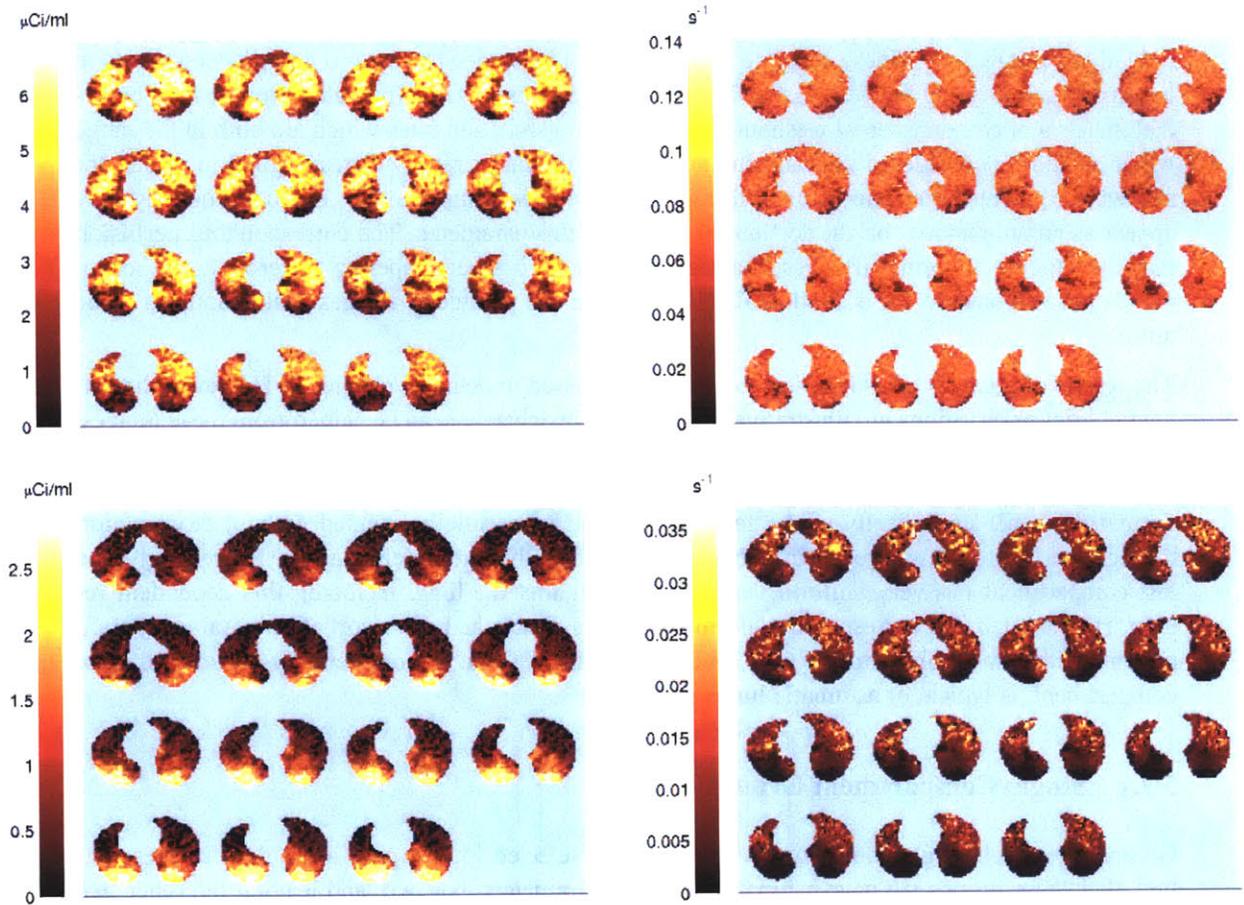


Figure 5.17: Individual washout parameters estimated from image h030-perf2. These are arranged with perfusions A_1 (top) and A_2 (bottom) on the left, and specific ventilations α_1 (top) and α_2 (bottom) on the right. Note that the large region of hypoventilation (α_2) still has significant perfusion (A_2), but that the fast-compartment ventilation (α_1) is remarkably uniform, even in this highly-constricted area.

If, on the other hand, a single-compartment group has a washout rate that can only be considered fast or slow, then the amplitude of the other, null compartment is set to zero, and the washout rate of this null compartment is interpolated from surrounding voxels. This will improve the problem of artifactual speckling in the parameter image, but not affect blood gas calculations or the equivalent ventilation image, nor any functional data calculated from the parameters, since the interpolated compartment has zero perfusion.

These post-processing operations also simplify calculations that use the estimated parameter sets, as reasonable values for all four parameters are assigned for every voxel in the image.

5.8 Grouping Criteria

An interesting idiosyncrasy of this method of analysis is the separation between grouping voxels and parameter estimation from the resultant group kinetics. The ‘obvious’ approach is to maintain consistency by using similar criteria to group the data as for the parameter estimation, so that if we include a ‘magic’ perfusion point in the curve-fitting [Section 5.6.6], we should also include this in the principal component analysis. And if we employ a weighted curve-fit, we should weight the PCA similarly.

However, there are potential advantages to using different data in the selection of groups than that used in the subsequent parameter estimation. Consider again the kinetics of the abnormal ‘pop-up’ ROI shown in figure 5.4 [page 109]. In this example, we defined an ROI by selecting voxels that had high activity in a single, short frame. Because of the very high level of noise, the kinetics of this group of voxels exhibit an anomalous spike in the frame used to select the group. But in all other frames of the image, the kinetics of this group are normal. One strategy to examine the kinetics of a region with high tracer concentration while avoiding ‘pop-up’ artifacts would be to use this ROI, but ignore the selection frame (the spike) in subsequent analysis. In doing so, we are exploiting the fact that the noise in any one frame is uncorrelated with noise in other frames.

The ‘pop-up’ effect (ROI’s biased by noise) also manifests itself with groups selected using PCA. In the example analysis presented above, the second principal component is large in the early washout frames and small thereafter [see figure 5.15]. As a consequence, selection based on this component is dominated by the tracer content in the early washout frames. In the group kinetics shown in figure 5.16, this component varies from top to bottom, so that groups 1 and 86 are extreme values of this component. We observe that the group 1 has kinetics with a small negative tail at the start of the washout, and that conversely, group 86 has a small positive tail. A full-sized plot of the kinetics of these groups is shown in figure 5.18(a). Although this could conceivably reflect genuine kinetics, it is more likely a selection artifact, particularly as the heavily-weighted frames are brief.

We can to a large extent circumvent this problem by including apnea data in the PCA and group selection, but not in the parameter estimation. Since the apnea data has high activity it also has high variance, and gets heavily weighted in the PCA. In this way, though the apnea data in the groups is indeed biased by noise, the washout kinetics have dramatically reduced bias.

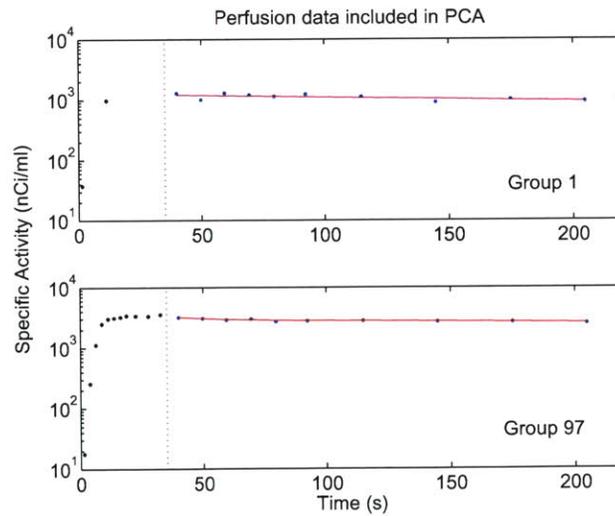
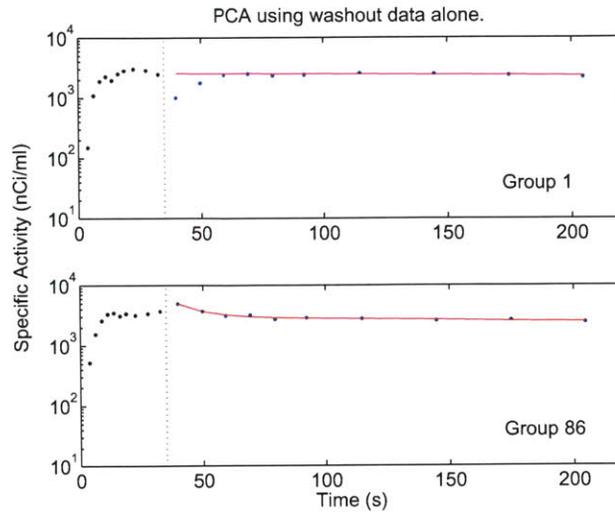


Figure 5.18: Default grouping discriminates, to a large extent, by the tracer content in early washout frames. As a consequence, extreme groups are biased by noise in these frames. Such groups, shown in the upper pair of axes, exhibit negative and positive ‘tails’ at the washout start. Inclusion of perfusion information in the PCA, but not in the subsequent parameter estimation, makes the analysis more robust to noise-biasing. The lower pair of axes show kinetics from (equivalent) groups derived by this method.

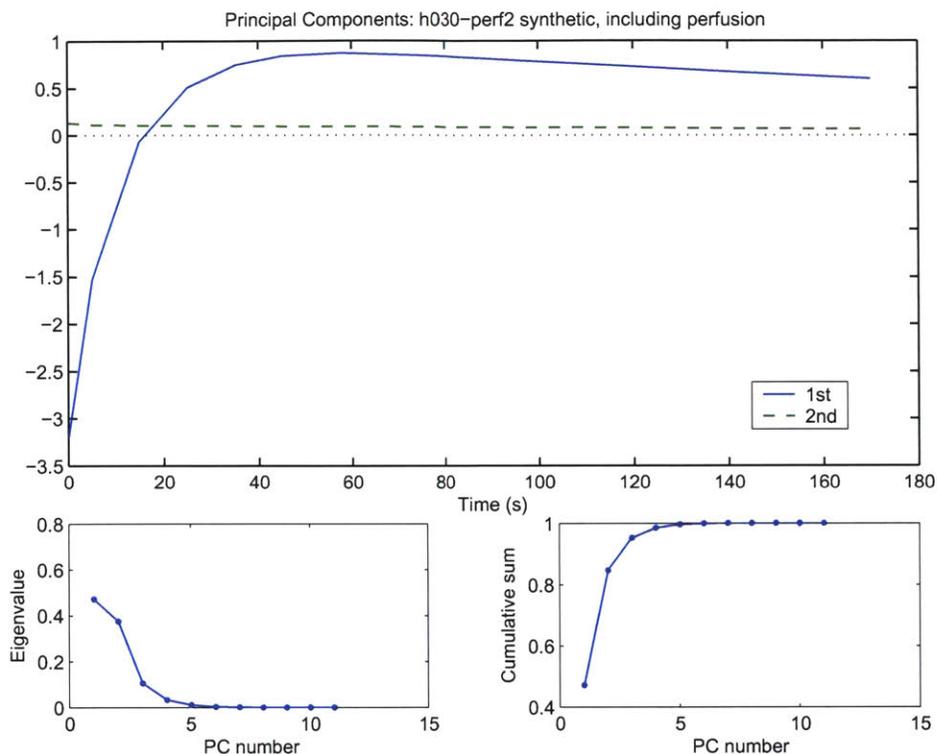


Figure 5.19: Principal components generated using a ‘magic point’ representing the perfusion plateau. The high activity in the perfusion frames, high variation in perfusion, and the weight given to this point (weighted by plateau duration) mean that it becomes the largest element of the first (most significant) principal component.

5.8.1 Revised example

Let us re-analyze the data from section 5.7 using this modified method. We add a ‘magic point’, representing the perfusion mean, at the start of the washout, and then generate principal components as before. The resulting PC’s are shown in figure 5.19: note that now the first, and thus most heavily weighted, component is greatest in its first element, corresponding to the first frame, the ‘magic point’.

Figure 5.18(b) shows the group kinetics generated by partitioning in the revised feature space. This modification has removed the artifactual ‘tails’ from previously biased groups [Figure 5.18(a)] and improved the robustness of the analysis.

5.9 Validation

Various measures were used to assess the validity of the analysis:

- Reconstruction of global kinetics.
- Quality of the model fit.
- Accuracy of analysis of artificial data.
- Perfusion prediction from washout.

Each of these is discussed below.

5.9.1 Reconstruction of global kinetics.

Using the identified parameters, we can reconstruct the tracer kinetics for each voxel. Summing the reconstructed kinetics of all the voxels results in a prediction of the global kinetics of the whole lung. If the analysis method is yielding accurate estimates, we should find that the reconstructed global kinetics matches the measured kinetics of the lung.

This comparison is sensitive to bias in our estimation: if there is a consistent error in a parameter, this trend will be amplified over thousands of voxels, and will be evident in the reconstructed global kinetics. Secondly, the consistency of the model is tested: if the model is not able to describe the behavior of the tracer, the reconstructed global kinetics will be a poor match.

The results shown in figure 5.20 are typical. The reconstructed kinetics invariably match the experimental data very well.

5.9.2 Quality of model fit

A key measure of the performance of our voxel grouping algorithms is how well the kinetics of each group matches the model behavior. A low-quality fit can have several, non-exclusive, explanations: (i) the model does not adequately describe the system; (ii) we have grouped voxels badly; (iii) a lot of noise remains in the group kinetics.

Since we employ synthetic kinetics to generate principal components [section 5.4.2], the transformation naturally promotes selection of groups whose kinetics match the model. However, any systematic departure from model behavior should be visible in the kinetics and curve-fits. In all cases studied, the model was able to fit the group kinetics well. [See, for example, figure 5.16.]

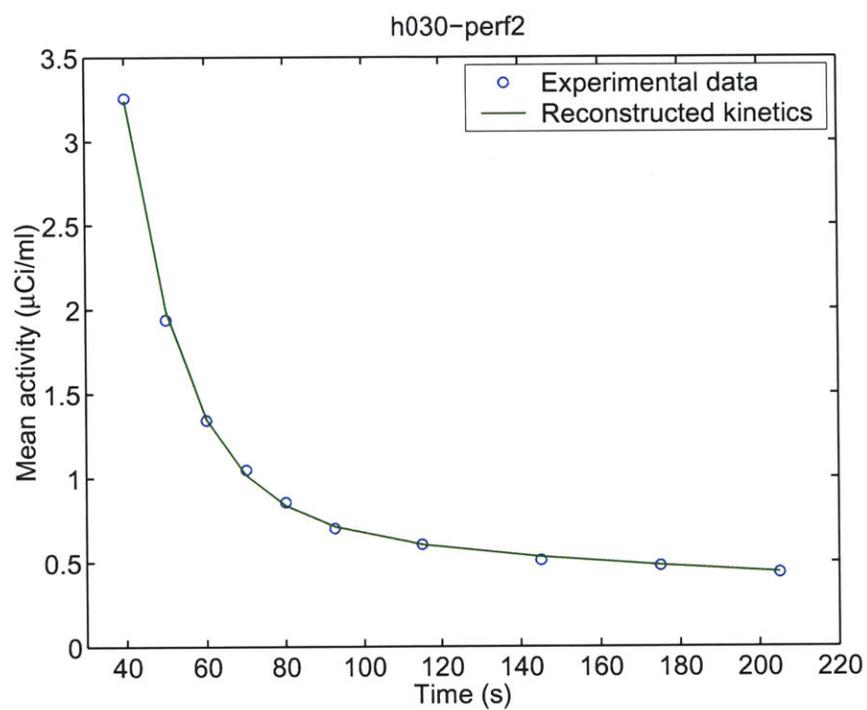


Figure 5.20: Tracer kinetics for individual voxels are reconstructed from estimated parameters, and summed to generate whole lung kinetics. Because these kinetics are summed over thousands of voxels, any bias in our estimates should be evident when these kinetics are compared to actual experimental data. As the above graph illustrates, the discrepancy is slight, and within the expected noise levels.

5.9.3 Analysis of artificial data

Since we do not know the real distribution of ventilation, perfusion, and aeration in the lungs, we do not have a ‘gold standard’ against which we can assess the accuracy of our results. One validation strategy is to analyze an artificial image to which noise has been added, and attempt to recover the parameters used to generate the artificial image.

Although this approach is powerful, it has some limitations. It will not, for example, reveal defects in our analysis due to limitations of our kinetic model. Consider for example, a region of the lung with very low perfusion but normal ventilation: because of the low perfusion, very little tracer accumulates during apnea, but tidal ventilation can convect tracer in from adjoining areas. In such a case, the tracer concentration paradoxically *increases* at the start of the ‘washout’. Such behavior is not described by our washout model.

5.9.3.1 Generation of test data

To generate plausible test data, we take parameter sets generated by analysis of experimental data using this grouped parameter estimation method. Next we (spatially) filter the parameter sets, which has the effect of removing the discretization/quantization of the parameters. Then we use the filtered parameters to generate ‘perfect’ tracer kinetics. Finally we add noise to the kinetics, varying the noise with total activity and imaging duration according to the noise model described in section 1.5.1, matching the level of noise to the original PET image.

After the test image is analyzed, we can compare the resultant parameter estimates to the the original parameter set that was used to generate the test image.

5.9.3.2 Quantifying Performance

The results of the analysis are regional estimates of the four washout parameters, and so we can immediately define the four error norms as obvious choices of optimality measures. We can also reconstruct kinetics from our parameter estimates, and compare these to the noisy input images. But what sort of error measure should we use to quantify the discrepancy? And how can we form a single unified error measure from the four parametric measures?

In the treatment below, I have elected to evaluate individual errors in each of the four parameters. Apart from simplicity and convenience, this has the advantage that we can observe parametric errors separately: the quality of estimation and dependence on number of groups may be very different for each parameter.

Error is quantified by the mean and standard deviation of the normalized error: $\text{mean}(e_i)$, $\sigma(e_i)$

$$e_i = (\hat{p}_i - p_i)/p_i$$

Where e_i is the normalized error of estimation of parameter p_i .

5.9.3.3 Optimal Number of Groups

The performance of the method depends on the number of groups into which the transformed data is partitioned. If the number is made very small, then accuracy is poor because (i) the parameters are very coarsely discretized and (ii) because we assemble groups with very heterogeneous kinetics. If, on the other hand, the number of groups is very large, then the kinetics of each group will be very noisy, and our parameter estimates will be unreliable. Somewhere in the middle, we expect to be able to strike an optimal compromise. (Susceptibility to grouping by noise is determined mostly by preprocessing and PCA settings, and is largely insensitive to the number of groups used.)

5.9.3.4 Results

Results of a typical error analysis are shown in figure 5.21. In this example, a representative parameter set was generated from image h030-perf2, and these parameters used to create an artificial emission series. Normally-distributed noise was added, with activity- and duration-dependent variance to match the original data. The noisy data was then analyzed using a varying number of groups. The performance of the method was measured by comparison of the estimated parameters to the parameters used to generate the image.

The results indicate that the analysis is, in general, fairly insensitive to the number of groups used.

Accuracy of parameter estimation (measured by the mean normalized error) was generally good, with consistently less than 5% error. The exception is parameter α_2 , in which the error is a little greater, due to the small mean value of this parameter (which amplifies absolute error), and because the tracer kinetics are relatively insensitive to small values of the parameter (the kinetics, for example, of a 0.001 s^{-1} compartment and a 0.002 s^{-1} compartment are almost identical). Such errors will have little effect on blood gas prediction.

5.9.4 Perfusion Prediction

As mentioned in section 5.6.4, we expect that the sum of amplitudes A_1, A_2 should equal the end-apnea level. We can use this property to validate the analysis method.

Since the apnea frames normally have short duration, the last apnea frame is typically very noisy. We observe that the apnea plateau is essentially flat in well-inflated lung (where the time-constant, E , is much longer than the apnea duration), and compare the amplitude sum to the apnea mean:

$$A_1 + A_2 \sim \overline{A_{ap}} \quad (5.17)$$

This comparison is valid, of course, only if we have not already used the perfusion information as a ‘magic point’ [section 5.6.6] in the parameter estimation.

As an example, grouped parameter analysis was performed on image h030-perf2. An image of the sum of parameters A_1, A_2 was generated, and filtered using a low-pass filter with cutoff wavelength

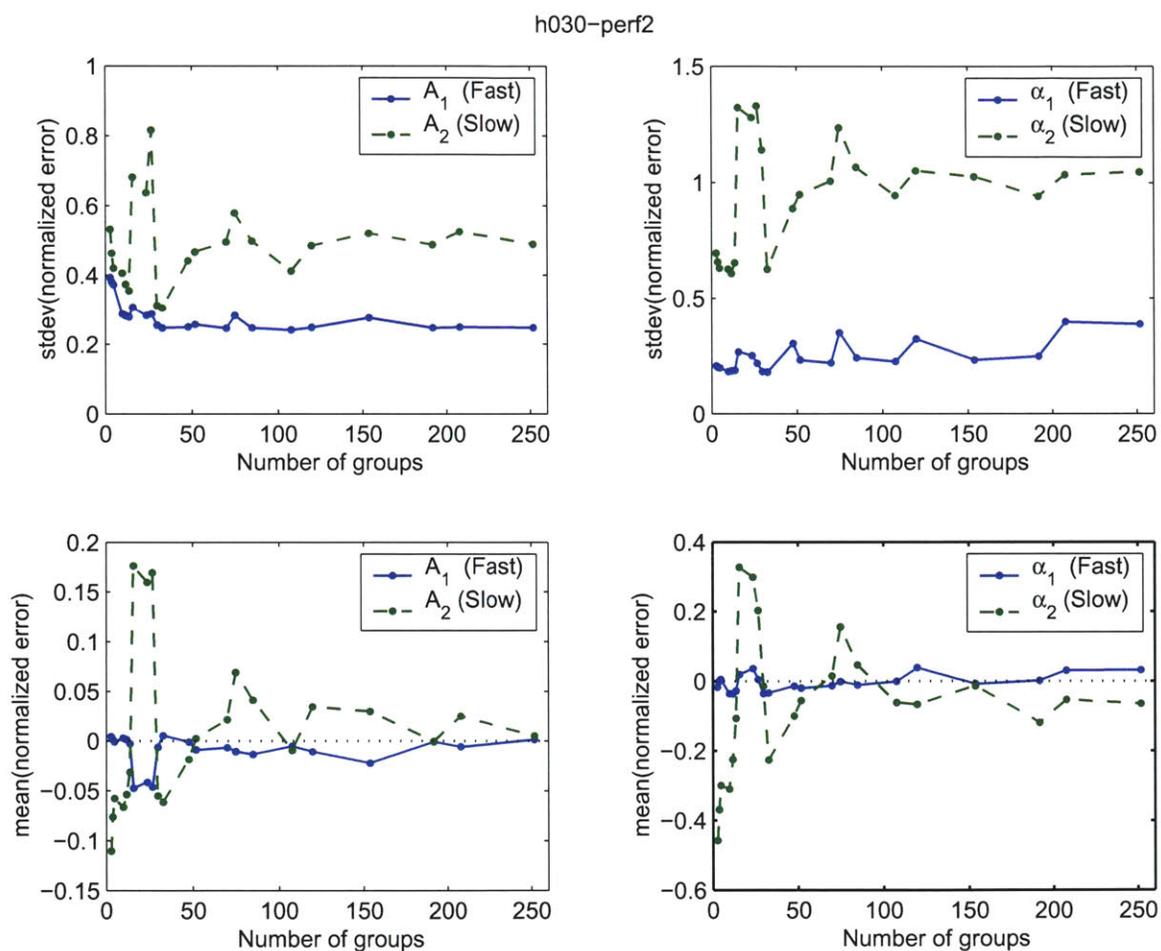


Figure 5.21: Error in estimation: dependence on number of groups. Accuracy of parameter identification was measured as the number of groups was increased. The mean error, which would indicate a systematic bias in parameter estimates, remained slight. The large standard deviation of the normalized error in α_2 is due to the small mean value of this parameter, so that normalization amplifies the absolute error.

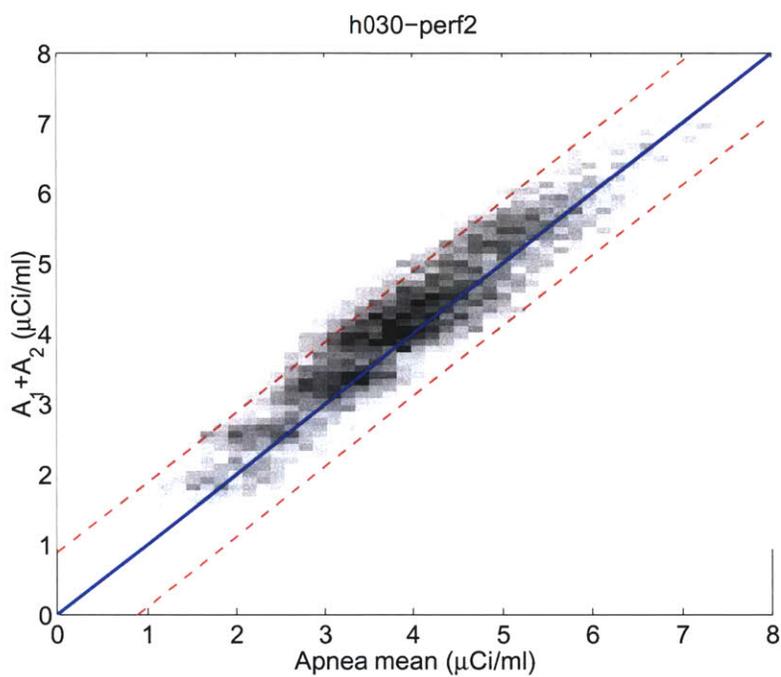


Figure 5.22: Prediction of perfusion (as assessed by mean of plateau concentration) from washout data. The tracer kinetics at the start of the washout are continuous with the end-apnea value, so that the sum of the amplitudes, $A_1 + A_2$ should equal the apnea concentration. We can use this property as a simple validation of the washout analysis. In the above graph, these quantities for each voxel are plotted against each other, showing that the washout data is an excellent predictor of perfusion.

of 10mm. The filtering was performed to remove artifacts of the discrete parameter values, making the distribution smooth rather than discontinuous. An image of the apnea plateau mean was filtered identically.

In figure 5.22, the sum, $A_1 + A_2$, for each voxel is plotted against the apnea plateau mean for that voxel, with the identity line shown in blue. The red dashed lines indicate the standard deviation of the noise in the apnea concentration.⁴ We see that the washout data is an excellent predictor of perfusion, with almost all of the data lying within the expected noise range.

5.10 Summary

The method of Kimura et al, originally applied to FDG kinetics in the brain, is applied to the analysis of nitrogen-13 washout data.

Apart from the small changes required to apply this analysis technique to a different system, the method has been enhanced in a number of ways. The most important enhancements are as follows:

- Principal components are generated from synthetic, noise-free data.
- Groups are selected by orthogonal partitioning, rather than clustering methods.
- Different data are used for grouping voxels than for parameter estimation.

This technique is demonstrated to yield accurate parameter estimates at full imaging resolution. In the following chapter, the method is used to analyze ventilation in a cohort of asthmatic subjects.

⁴The noise is estimated using the model described in section 1.5. For the filtered image, the gradient $k/V_{\text{voxel}} = 2.1$ nCi/ml²s, $A_{\text{field}} = 8.4$ mCi, and plateau duration 22.6 s. This predicts noise of standard deviation $0.9 \mu\text{Ci/ml}$.

Chapter 6

Ventilation Disruption in Asthmatics

In this chapter the methods described in chapter 5 are applied to analyze washout data from human asthmatics.

6.1 Subject Selection

To keep the amount of data presented manageable, a subset of our asthmatic subjects was selected for analysis. Specifically, studies were excluded in which the patient was obese, since in these subjects the data is much noisier [see section 1.5]. Similarly, subjects were excluded who moved during imaging, were unable to breath-hold, or had other difficulties in complying with the experimental protocol.

The selected studies were h021, h030, h035 (females) and h036, h037, h041 (males).

6.2 Experimental Protocol

Six subjects were selected for analysis from a group of PET studies of mild to moderate asthmatics. In each subject the experimental protocol was the same. A set of baseline PET scans were taken, and then an asthma attack was artificially provoked by administering a bronchoconstrictive agent (methacholine). A second set of PET scans were then acquired after the provocation.

Methacholine (MCh) was administered as an inhaled aerosol, at a dose previously determined to induce a 20% reduction in the volume of air that can be exhaled in one second (FEV_1), the so-called PC20 dose.

Other details of the protocol are summarized below.

- Subject instructed to discontinue medications 24hrs prior to study
- IV located in antecubital vein.

- Breathing monitored by impedance plethysmography (Respirace, Non-Invasive Monitoring Systems, Miami Beach, FL).
- Subject instructed to take deep breath and exhale to mean lung volume.
- Apnea duration of 30-40s, depending on subject's capacity to breath-hold.
- 30ml bolus of nitrogen-13 tracer solution infused at 5ml/s, coincident with start of apnea.
- PET imaging sequence initiated at start of apnea.
- Nebulized methacholine administered at PC20 dose (previously determined).
- Tracer infusion and imaging sequence repeated.

6.3 Tracer Kinetics

Before addressing the detail of the perfusion and ventilation, it is useful to examine briefly the overall tracer kinetics.

In figures 6.1 and 6.2 (female subjects and male subjects), tracer kinetics are shown at baseline and post-provocation. In these plots the lung was divided vertically into three ROI's, labeled non-dependent (ventral), middle, and dependent (dorsal).

Under baseline conditions, there are some differences in the washout kinetics between the three regions, but the discrepancy is slight, indicating fairly uniform ventilation across the lung.

Post-provocation, the overall kinetics were dramatically different, with substantial amounts of tracer retained in the lung at the end of the washout [Table 6.2]. Furthermore, the three regions exhibit far greater differences in washout behavior, indicating a high degree of heterogeneity in ventilation. The overall trend is for greater bronchoconstriction in dependent lung compared to non-dependent, but the amount of vertical gradient varies substantially between subjects.

6.3.1 Constancy of Ventilation

One important feature to note is the large gradient present at the start of the washout. At first sight this appears to reflect the fast time-constant of highly-ventilated lung. However, some subjects, particularly h021 in the subset shown here, exhibit washout kinetics that are mostly straight once the first 30 seconds of the washout has passed. Are we dealing with two very distinct populations of lung units? At baseline this seems a little surprising, particularly when the lung function tests are normal.

Another explanation for the observed kinetics is that the subject is breathing unevenly. In the washout description developed in section 4.5, the washout rate is given as $(s\dot{V} + \frac{1}{E})$. We have tacitly assumed, however, that the specific ventilation, $s\dot{V}$, is constant. At the end of the apnea, however, the subject is inevitably hypercapnic (elevated P_{CO_2}), and feels the urge to breathe heavily

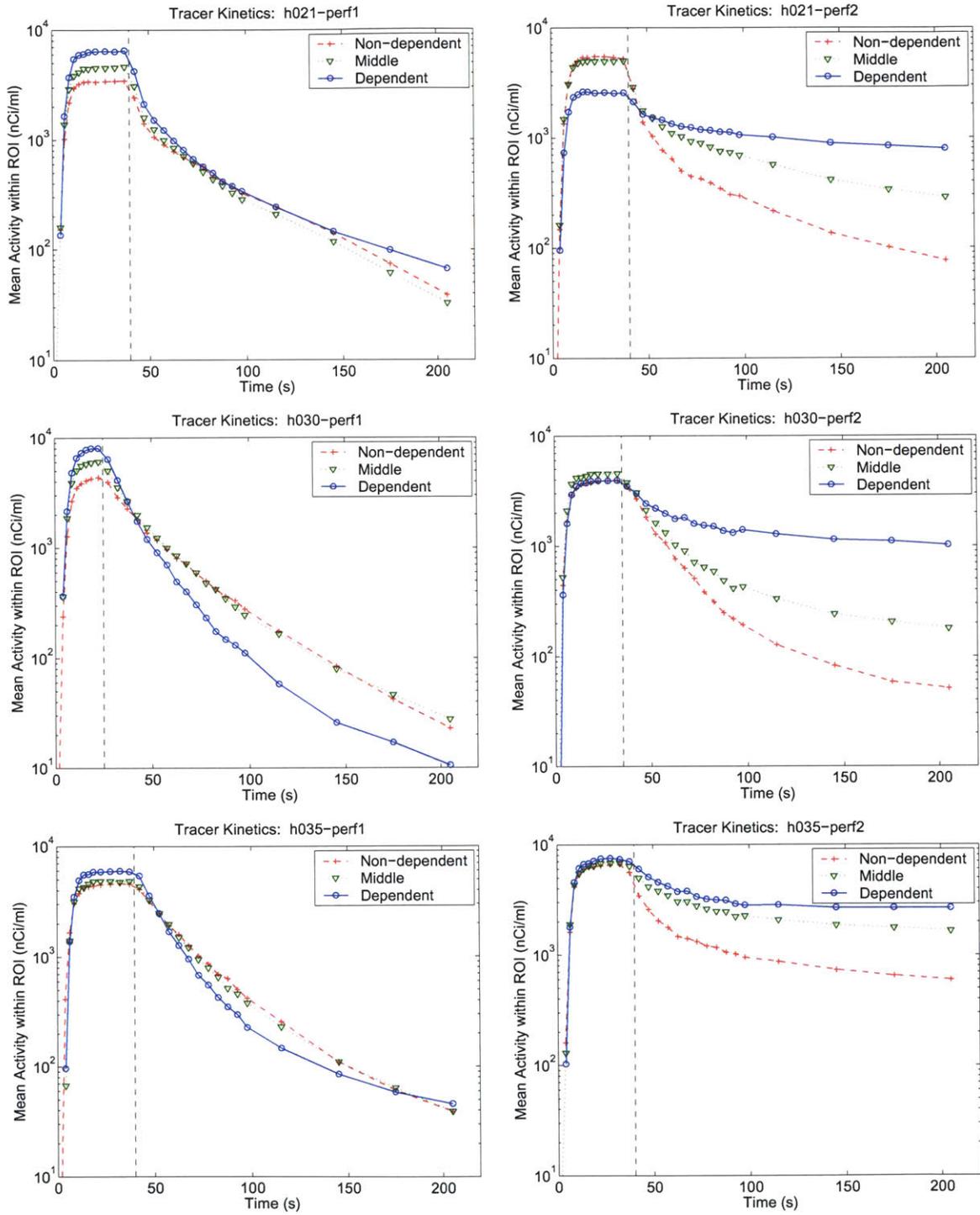


Figure 6.1: Tracer kinetics from the female subjects: h021, h030, h035. The baseline state is shown on the left, and post-provocation on the right.

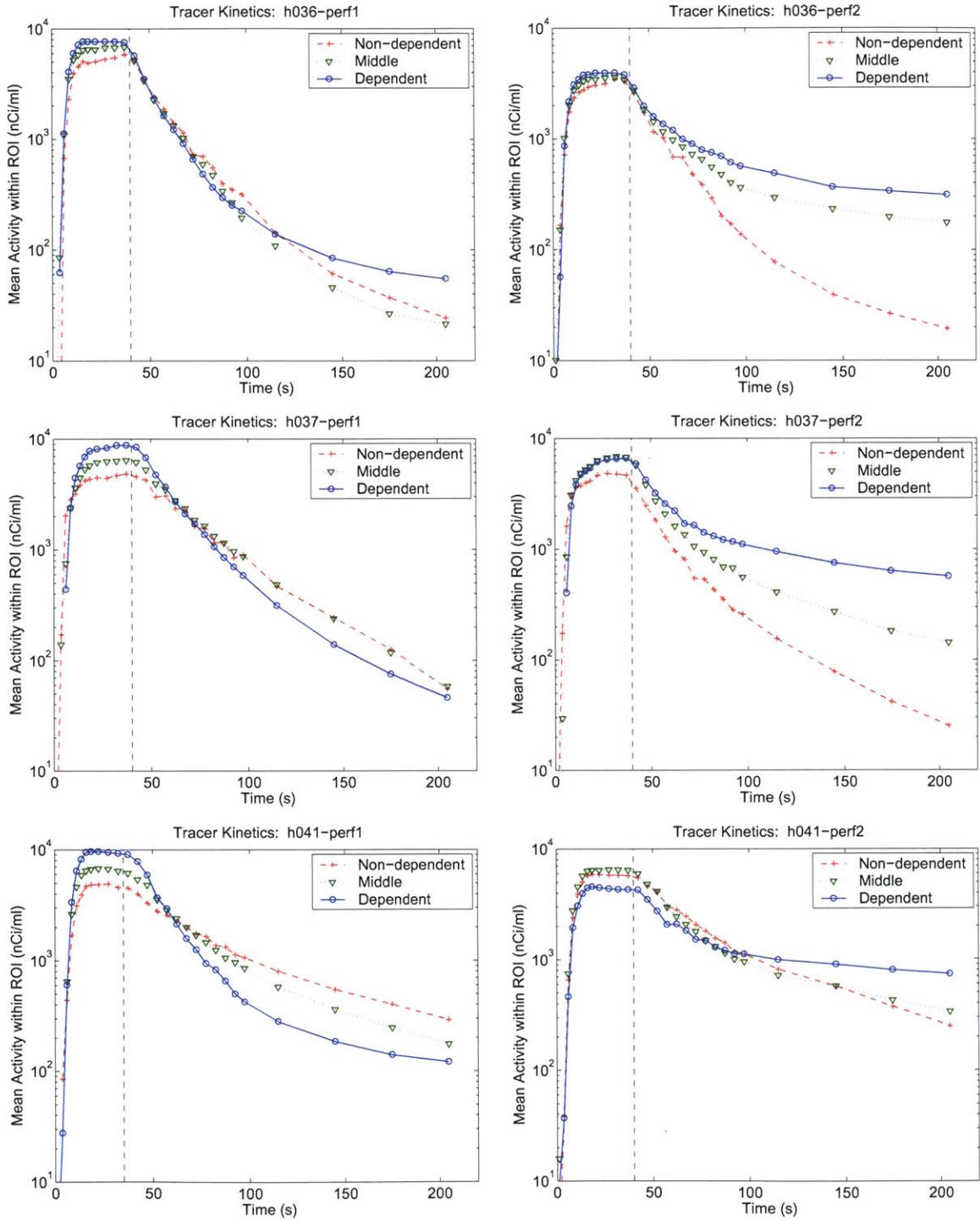


Figure 6.2: Tracer kinetics from male subjects h036, h037, h041. The baseline state is shown on the left, and post-provocation on the right.

in order to restore normocapnia.¹ Thus the steep gradient at the early washout may merely reflect transiently elevated ventilation.

Discriminating between such uneven breathing and multi-compartmental ventilation is impossible using the PET data alone. We need extra measurements of the total ventilation. In these asthmatic subjects, suitable data were acquired using inductance plethysmography. However, the analysis presented here does not attempt to compensate the kinetics or otherwise modify the analysis to remove this effect.

6.4 Perfusion

6.4.1 Method

Relative perfusion images were made simply by calculating the duration-weighted mean of the apnea plateau frames. As discussed in chapter 4, the plateau activity is proportional to perfusion provided that the lung is well-inflated, and that there is no significant non-aerated compartment.

Since these subjects had mild to moderate asthma, and were otherwise healthy, this is likely to be a reasonable assumption. Additionally, no abnormalities were visible on the apnea kinetics.

As discussed in section 4.7.2.1, since the plateau is short (≤ 30 seconds) relative to the time-constant of the lung ($E \sim 10^3$ seconds), it is not feasible to attempt to identify this parameter. Taking the plateau mean is conceptually equivalent to a single-parameter estimation of the apnea data.

The images shown here were post-processed by low-pass filtering with a 12mm cut-off wavelength (i.e. removing features smaller than 6mm).

6.4.2 Results

At baseline, the perfusion of all subjects was essentially uniform, with slight and varying amounts of vertical gradient. [Table 6.1.]

The normalized variance in plateau concentration was used as a global index of perfusion heterogeneity. The 'ideal' value of this measure was estimated as the y-intercept of the linear regression, as described in section 1.5.1. This compensates for changes in variance that are due to differing apnea durations and amount of tracer injected.

Following bronchoconstrictive provocation, heterogeneity of perfusion was observed to increase substantially. [Table 6.1] Additionally, the fraction of the heterogeneity explained by a vertical gradient decreased greatly. The exception to this is subject h021, where the bronchoconstrictive response was focused in the dependent region of the lung; the consequent redistribution of perfusion resulted in a large negative perfusion gradient. [Figure 6.3]

In several of our asthmatic patients, suspension of their prophylactic medication has resulted in significant bronchoconstriction before the methacholine provocation. However, none of the selected subgroup of subjects showed discernible bronchoconstriction at baseline.

¹Although subjects are instructed to try to regulate their breathing, the desire to breathe is overwhelming.

Study	Vertical Gradient [cm^{-1}]		Heterogeneity		Fraction explained	
	Baseline	Post-MCh	Baseline	Post-MCh	Baseline	Post-MCh
h021	7.51%	-8.23%	0.103	0.390	47.9%	22.4%
h030	6.77%	-0.47%	0.141	0.182	36.3%	0.3%
h035	3.67%	1.85%	0.077	0.117	16.6%	3.5%
h036	3.21%	1.65%	0.073	0.127	18.8%	3.4%
h037	6.51%	2.27%	0.200	0.248	22.8%	2.6%
h041	5.74%	-2.25%	0.167	0.232	36.8%	4.6%

Table 6.1: Various measures of perfusion, at baseline and post-provocation with methacholine (MCh). The vertical gradient is the mean gradient of normalized perfusion. Overall heterogeneity is measured by normalized variance (σ^2/mean^2). Also shown is the fraction of the heterogeneity in the lung explained by the vertical gradient. In almost all cases, the amount of heterogeneity explained by the vertical gradient dropped markedly, reflecting the uneven redistribution of perfusion.

Study	Tracer Retention	
	Baseline	Post-provocation
h021	1.27%	10.7%
h030	0.45%	11.2%
h035	1.03%	24.2%
h036	0.56%	5.84%
h037	1.17%	4.61%
h041	3.16%	8.68%

Table 6.2: A simple, global index of bronchoconstriction is provided by the amount of tracer retained in the lung at the end of the washout. These values are the mean concentration in the last 60s of the 3-minute washout, expressed as a fraction of the peak apnea concentration. The response to methacholine is seen to be highly variable, despite the fact that the drug was administered at the previously-determined PC20 dose.

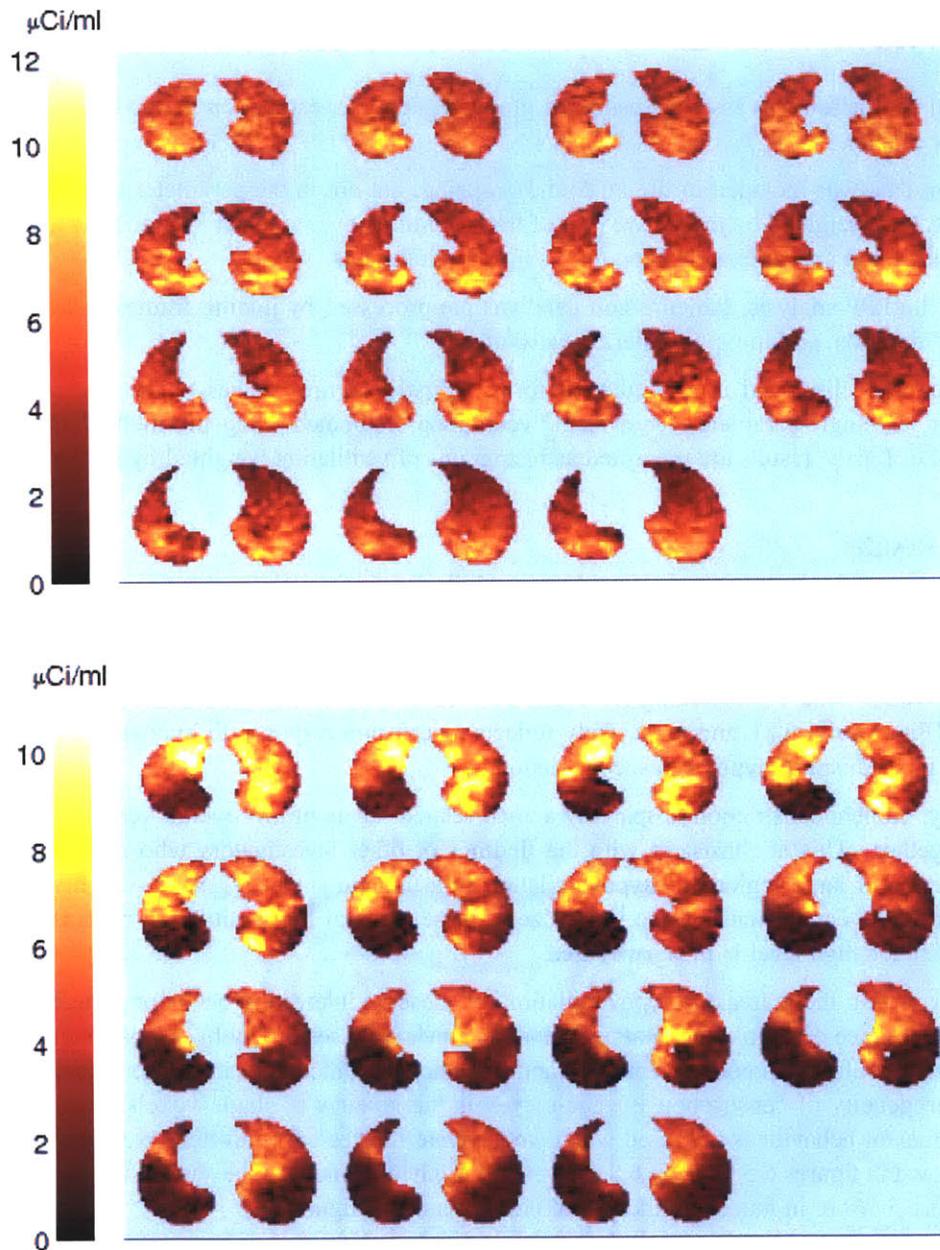


Figure 6.3: Sample perfusion images from the same subject (h021), before (top) and after (below) bronchoconstrictive provocation. Perfusion has been extensively redistributed in response to the deranged ventilation.

6.5 Ventilation

6.5.1 Method

Regional ventilation was assessed using the grouped parameter estimation method described in the previous chapter.

Perfusion data was included in the PCA and grouping, but not in the parameter estimation phase. The PCA was weighted by the square root of frame duration over total activity, and the transformed data divided into approximately fifty groups in two dimensions.

Prior to the PPP analysis, the emission data was pre-processed by joining frames of less than 10-seconds' duration, and low-pass filtering to 10mm.

The results are displayed in a number of forms. Firstly, as raw images of the four parameters. Secondly, as single-parameter 'equivalent' ventilation, estimated using the method described in section 4.6. Lastly, results are presented as histograms of ventilation weighted by perfusion.

6.5.2 Results

Two major findings are consistent across all of the studies.

Firstly, ventilation is correlated with perfusion: areas of low ventilation are areas of low perfusion, and vice-versa. This VQ-matching can easily be seen in the plots of perfusion and equivalent ventilation [Figures 6.5-6.8], and most likely reflects a perfusion response to hypoventilation, via the familiar mechanism of hypoxic vasoconstriction.

Secondly, bronchoconstriction is spatially autocorrelated: areas of low overall ventilation are clustered together. This is consistent with the findings of other investigators who have documented the presence of large regions of hypoventilation in asthmatic subjects [14, 1]. At first sight, we might draw the conclusion that the large size of these areas of hypoventilation reflects control of ventilation at a high level in the airway tree.

However, within these areas of hypoventilation, we observe interesting behavior. There is rarely a complete absence of ventilation. Rather, we almost invariably see highly bi compartmental kinetics where one population is poorly ventilated and the other ventilated at nearly normal rates. Notably, this heterogeneity of constriction is visible even in the kinetics of single voxels. Such locally bi compartmental behavior is obscured when we generate images of 'equivalent' ventilation, such as those shown in figures 6.5 through 6.8: however, this phenomenon can be seen in the group kinetics [e.g. figure 5.16] or in images of individual parameters [e.g. figure 6.4].

We can draw several conclusions from these findings. Firstly, the fact that lung compartments are either very poorly ventilated or near-normally ventilated suggests that bronchoconstriction may be bistable in nature. This possibility has previously been proposed by Anafi and Wilson [2], based on a theoretical analysis of airway mechanics. Secondly, the appearance of bi compartmental kinetics in single-voxel ROI's suggests that ventilation is modulated at length scales below the resolution of the camera ($\sim 6\text{mm}$).

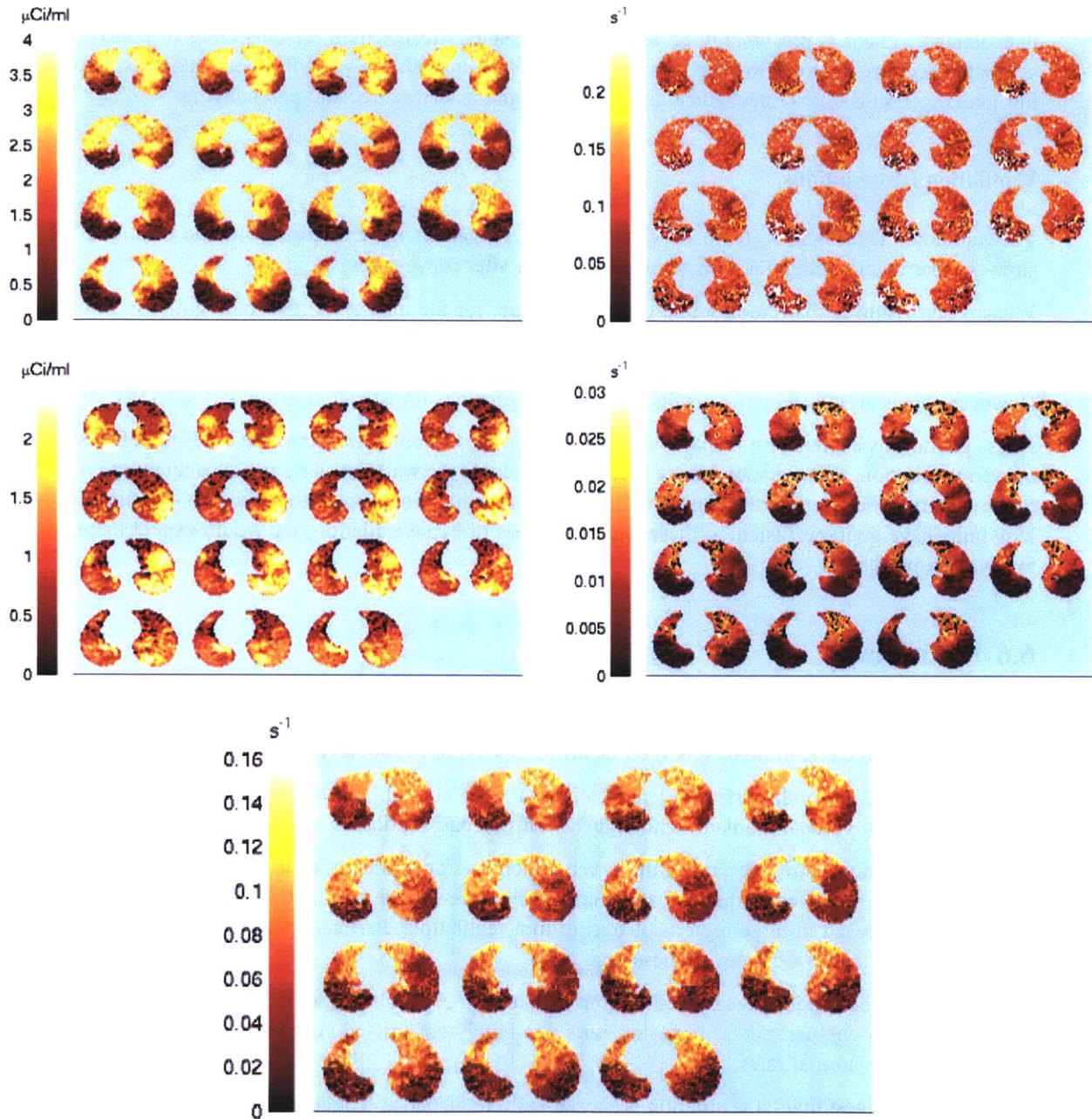


Figure 6.4: Individual parameters estimated for image h021-perf2. Perfusions A_1 (top) and A_2 (middle) are shown on the left-hand side; corresponding specific ventilations α_1 (top) and α_2 (middle) are shown on the right. Underneath is an equivalent ventilation image, which is the single-compartment ventilation that would give the same P_{O_2} . [See section 4.6.]

The link between small-scale heterogeneity and large, contiguous areas of hypoventilation remains unexplained. There is presumably an additional mechanism responsible for clustering of poorly-ventilated regions. Current work in our laboratory [Winkler et al, not yet published] suggests that this phenomenon is due to mechanical coupling through the airway tree and parenchyma.

Ventilation Distributions

The ventilation distributions [6.9,6.10] clearly show derangement of ventilation, with the amount of slow-compartment ventilation dramatically increasing after provocation.

What is surprising is that the baseline distributions are, for the most part, also bimodal (but with much faster ventilation of the slow compartments). These results could be interpreted as suggesting that, at baseline, asthmatics may have a population of vulnerable, slightly-constricted airways. However, this is at odds with the results published by previous investigators using MIGET [18].

A less glamorous explanation is that given above – that the bimodal distribution at baseline reflects uneven ventilation, that the subjects are breathing rapidly at the washout start, and subsequently return to a normal ventilation rate. This would also explain why (what appear to be) slowly-ventilated lung units have good perfusion: if these were real areas of hypoventilation, we would expect them to have low perfusion.

6.6 Summary

The techniques developed in previous chapters of this thesis were applied to study six asthmatics subjects, at baseline and immediately after an artificially-induced asthma attack.

In the baseline state, both perfusion and ventilation were normal. Both were either uniform or exhibiting a slight vertical gradient, with higher perfusion and ventilation in the dependent lung.

Following bronchoconstrictive provocation, ventilation was dramatically changed. A much larger fraction of tracer remained in the lung at the end of the three-minute washout. Bronchoconstrictive response was focused in large regions of overall low ventilation. Perfusion was also dramatically changed, decreasing, as expected in these regions.

The kinetics of these poorly ventilated regions is interesting. In almost all cases, individual voxels showed multi-compartmental behavior, with one highly-constricted compartment, and another ventilated at near-normal rates.

These results suggest that (i) ventilation is substantially regulated at length scales below the resolution of the camera (~6mm); that (ii) airway constriction may be bistable, rather than continuous, in nature; that (iii) there is an additional mechanism which caused the large-scale clustering of poorly-ventilated regions. This latter mechanism may be a mechanical coupling through the parenchyma or airway tree, but other explanations are possible.

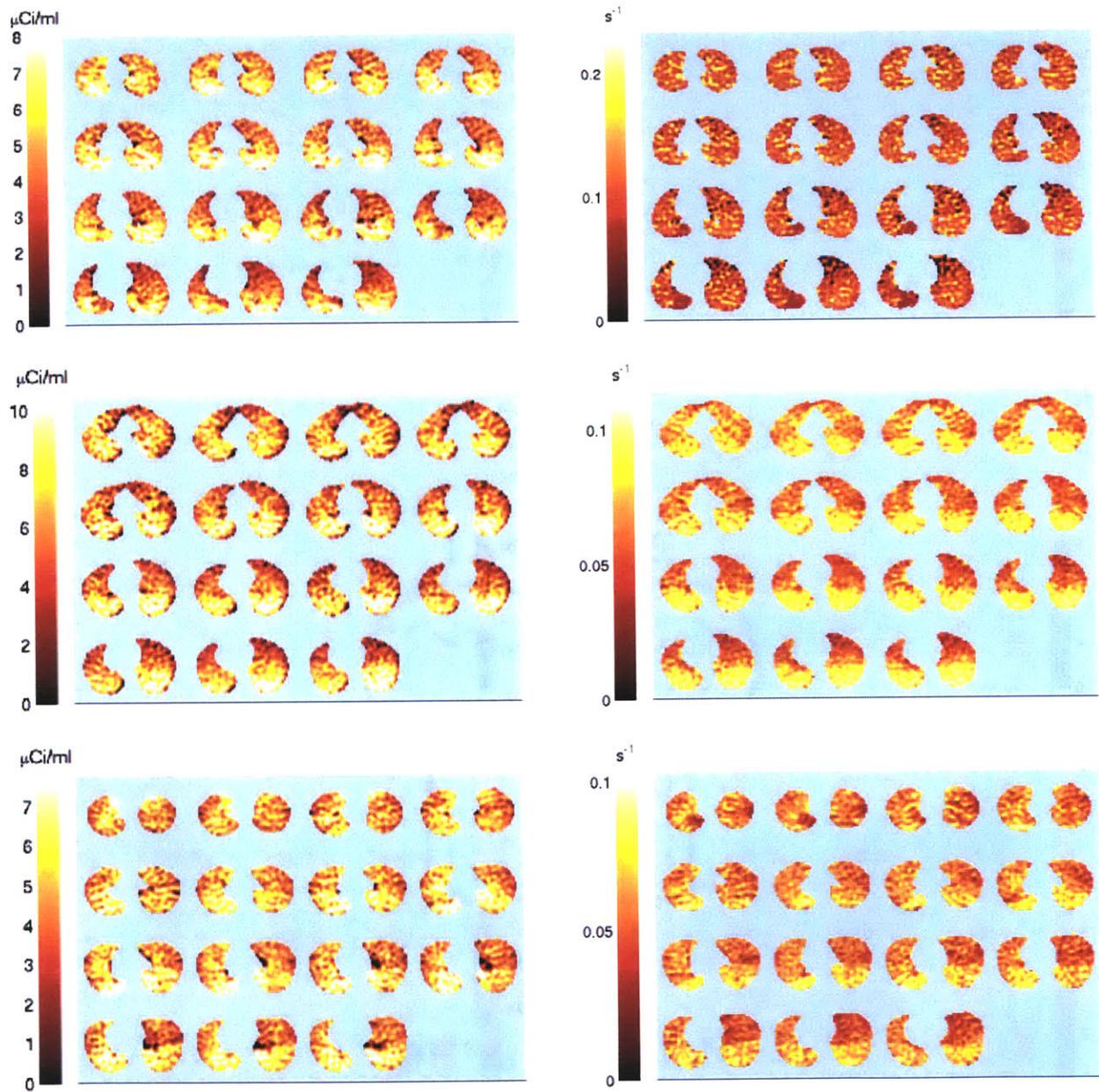


Figure 6.5: Baseline images of perfusion (left) and equivalent ventilation (right) from female subjects h021 (top), h030 (middle), h035 (bottom).

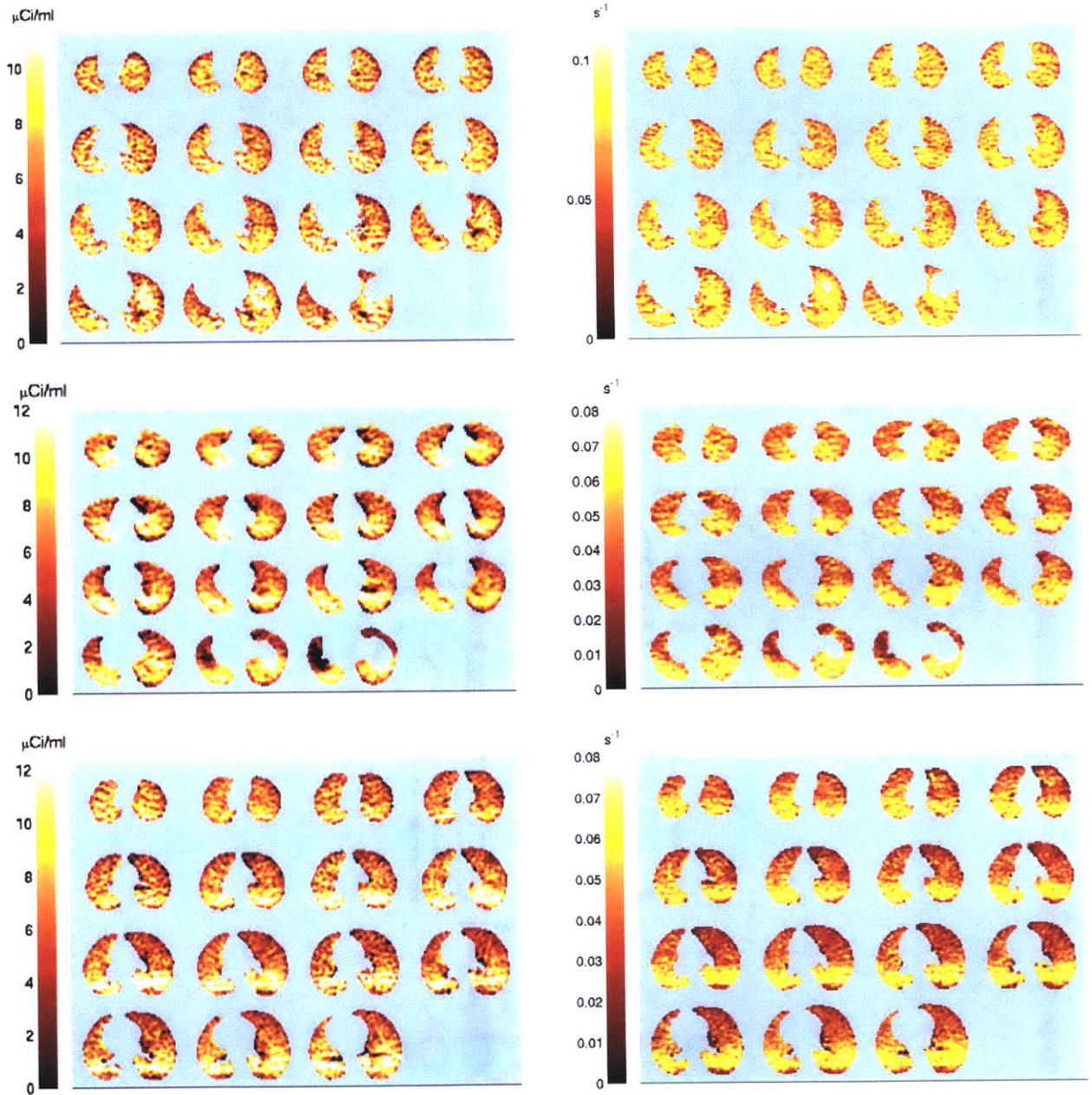


Figure 6.6: Baseline images of perfusion (left) and equivalent ventilation (right) from male subjects h036 (top), h037 (middle), h041 (bottom).

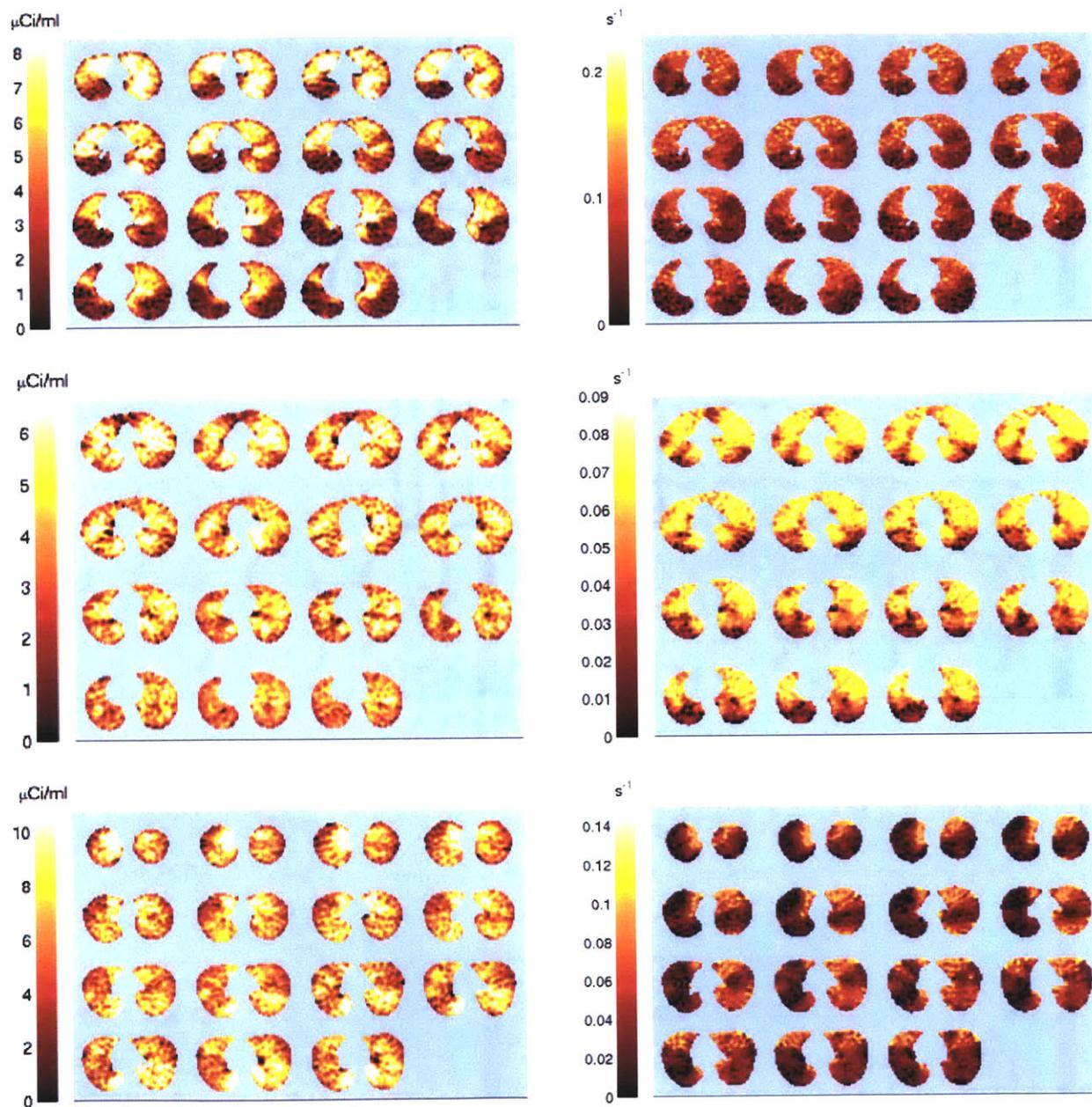


Figure 6.7: Post-provocation images of perfusion (left) and equivalent ventilation (right) from female subjects h021 (top), h030 (middle), h035 (bottom).

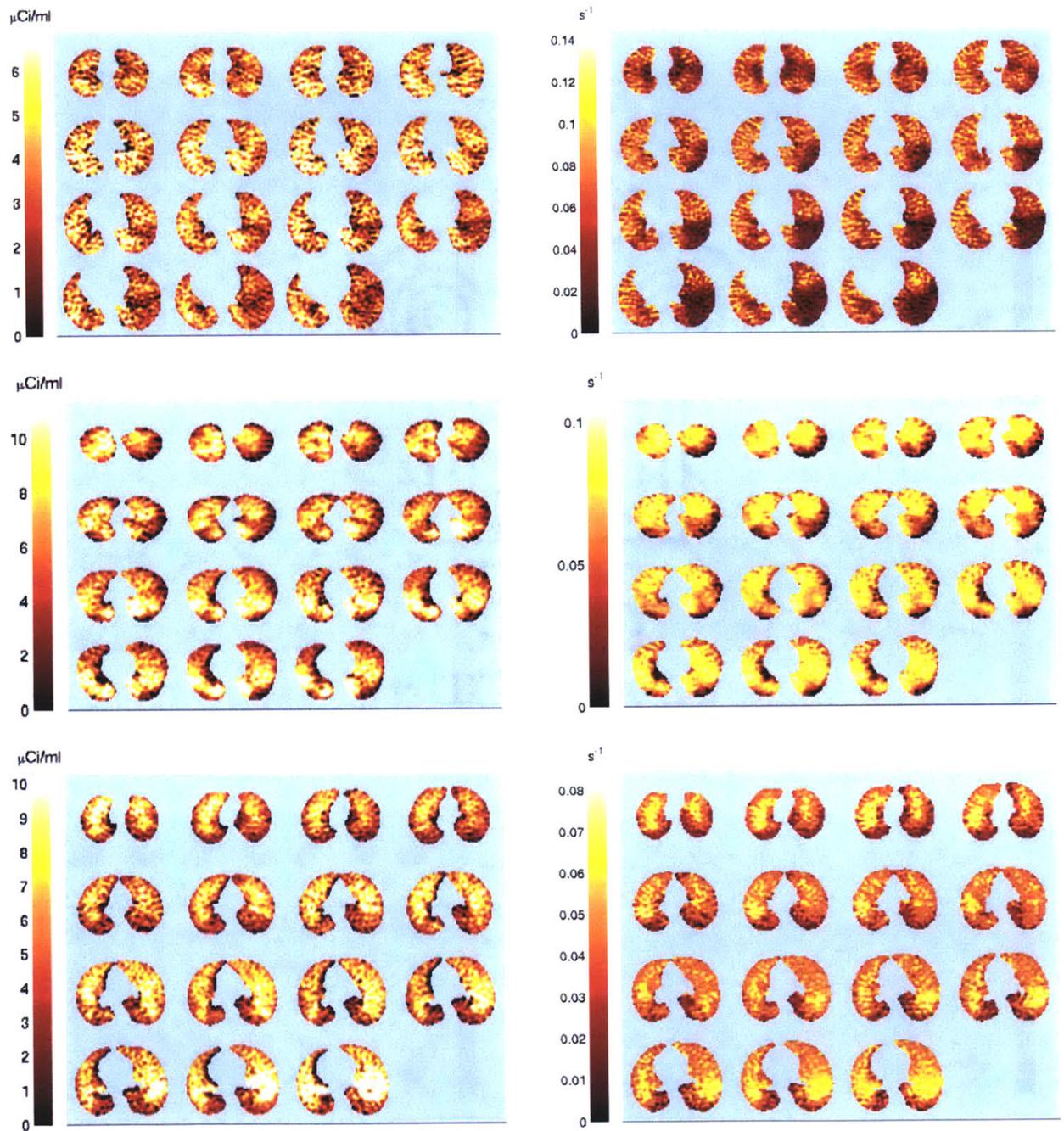


Figure 6.8: Post-provocation images of perfusion (left) and equivalent ventilation (right) from male subjects h036 (top), h037 (middle), h041 (bottom).

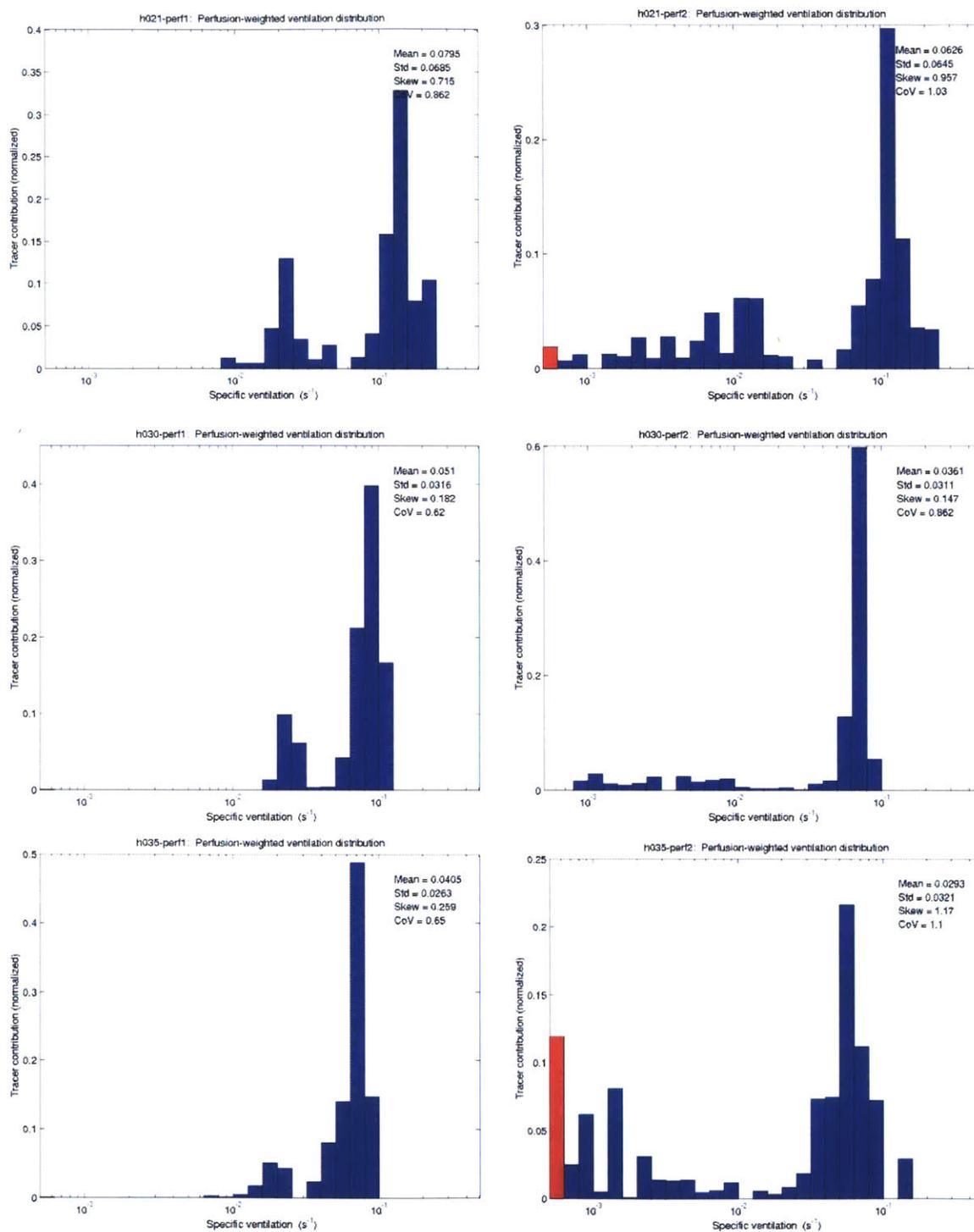


Figure 6.9: Ventilation distributions from female subjects h021, h030, h035. Baseline state is shown on left, post-provocation on right. Populations with slow and fast time constants are visible even at baseline. Post-provocation, the time constants of the slow compartment have lengthened dramatically.

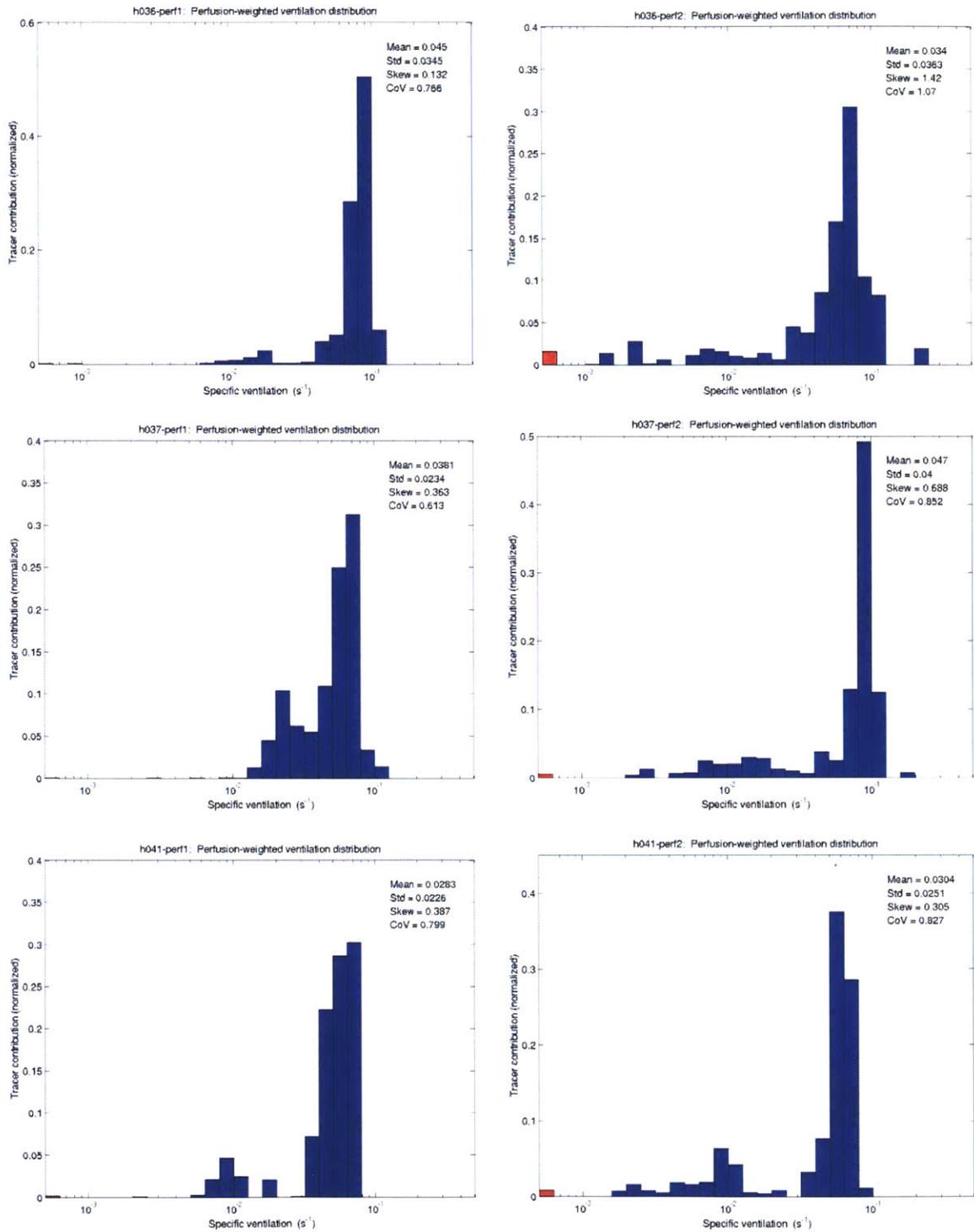


Figure 6.10: Ventilation distributions from male subjects h036, h037, h041. Baseline state is shown on left, post-provocation on right. Populations with slow and fast time constants are visible even at baseline. Post-provocation (right), the time constants of the slow compartment have lengthened dramatically, but perfusion to the slowly-ventilated lung remains slight.

Chapter 7

Conclusion

Our understanding of many lung diseases has been limited by our capability to visualize gas exchange in the pathologic lung. Positron Emission Tomography, using nitrogen-13 as a tracer, has the potential to provide clinicians with far more detailed information about the mechanism of lung disease than any existing techniques.

This thesis provides much of the fundamental work required to bring this technology out of the laboratory and into the clinical domain.

In part I of this thesis, a prototype system is described which permits the tracer solution to be prepared and infused on-site, quickly, safely, and easily. This prototype system has been in use for three years, and has been used successfully in approximately eighty animal and forty human studies, to date. The system has been continuously refined during this period, improving in reliability, performance, and ease-of-use. and is now at a point where it could be used as a reference design for a small production run. A patent for the system has been applied for, and is expected to be granted shortly.

In part II of this thesis, methods for analysis of the PET data are developed.

In chapter 4, a model of the tracer kinetics of the lung is formulated, elucidating the tacit assumptions that underlay the approximations used in previous studies with this method, permitting quantification of the error of these approximations, and clarifying the physiologic information that we can hope to extract using this imaging methods. The model provides a formal, extensible framework for the analysis of data from diseased lungs, and its use is demonstrated in a variety of conditions.

In chapter 5, a method of analysis is explored which promises to provide multiple parameter estimation robust to the very high noise levels associated with high-resolution data. The method is explored and validated, and its application demonstrated by analysis of washout data from asthmatic subjects.

Asthma Results

Some of the results presented in chapter 6 are unsurprising. We expect an asthma attack to change the overall distribution of ventilation, and similar bimodal distributions have been described by other authors [18]. Similarly, we expect that the perfusion will redistribute away from hypoventilated regions.

The new information in these data is the spatial distribution of functional parameters, particularly ventilation. The results show large, contiguous areas of hypoventilation. These areas are not uniformly hypoventilated, but are characterized by a bimodal ventilation distribution, suggesting that within these highly-constricted regions, some lung units are still ventilated at almost normal rates. Furthermore, this two-compartment behavior is visible even at the voxel level: so that heterogeneity of bronchoconstriction must occur at length scales below the resolution of the PET scanner.

These findings have not previously been reported. Furthermore, it is not feasible to obtain equivalent data with any other current imaging technique.

Appendix A

Important Experiments

A.1 Carbon Monoxide Content

This was a test to determine the typical carbon monoxide content of the gas mixture that is received by the tracer-preparation apparatus, and to confirm that the catalyst was functioning properly.

Test Setup and Results Experiment performed 9/19/02. Gas from cyclotron was collected in mylar balloons, allowed to decay, and then tested in a clinical carbon monoxide sensor. No tests for oxygen content (too small to be measurable with available equipment) or other gases were performed.

Run	Configuration	Measured CO
1	Dummy run. No bombardment.	0.006-0.007%
2	10 minute cook. No catalyst employed.	0.099%
3	10 minute cook. Catalyst inserted.	0.006-0.007%
4	10 minute cook. No catalyst.	0.051%
5	10 minute cook. Catalyst inserted.	0.006-0.007%

Discussion We were not expecting that there be any measurable CO content in the dummy run. In this case the target was filled from a carbon dioxide tank, and then vented directly to the apparatus. We believe this due to impurities in the CO₂ (listed as 10ppm THC, 10ppm H₂O) being falsely identified as CO by the sensor used.

With the catalyst in-situ, the CO content was identical to this baseline level, suggesting that the catalyst is working perfectly.

The difference in measured CO concentrations between the second and fourth runs is most likely due to the larger volume of gas that was received in the latter run, in which extra carbon dioxide (used to flush the long lines) has diluted the gas that was bombarded in the cyclotron. Since the

volumes received were not formally measured, the gas from the second run may also be slightly diluted, and so this value may slightly underestimate the real concentration.

Conclusions The carbon monoxide concentration in the cyclotron gases for a ten-minute cook is approximately 0.093%. For longer bombardments, this figure will be higher, probably scaling linearly with duration. The highest concentration we are likely to encounter is around 0.3%.

The catalyst appears to be working perfectly: recombination is complete, at least to the extent that can be measured with this apparatus. More importantly, this indicates that there is no appreciable removal of oxygen from the system, and that the mixture is stoichiometric.

A.2 Optimal NaOH concentration

In order to determine an optimal concentration for the sodium hydroxide, the following test was performed.

Test Setup A cylinder of carbon dioxide (medical grade) was connected to the Salsa Generator, and the pressure regulator set to 30psi. The absorbing chamber was filled with NaOH solution at a variety of different concentrations (1M, 2.5M, 5M, and 10M). After an initial surge as the compliant chamber distends, the system rapidly (<10s) reaches a pseudo-steady state, where the flow-rate into the chamber equals the rate at which CO₂ reacts with the NaOH. This steady flow rate was recorded, and tabulated against the total volume received.

Results During tests, we observed that the absorption rate steadily decreased. This decrease is due not just to the gradual loss of capacity to absorb, but to two faster, and larger effects. Firstly, the carbon dioxide tank used was medical grade, and of relatively low purity. Impurities in the gas progressively accumulate and the concentration of carbon dioxide in the bubble decreases. Secondly, the absorption of carbon dioxide increases the volume of the NaOH solution, so that the actual bubble volume decreases. The original absorption rate was usually restored by venting the bubble and excess NaOH volume before resuming the test.

At the 10M concentration, the absorber ultimately began to become cloudy as a white precipitate (presumably of sodium carbonate) formed.

The initial reaction rate was essentially independent of NaOH concentration, suggesting that the reaction is limited by how fast the CO₂ can diffuse into the the solution, and not by the concentration of the reactants. The rate remained high and steady for a while, and then began to fall. The point at which the reaction rate starts to decrease was roughly proportional to the initial concentration of the NaOH solution. Presumably, at this stage, the reaction rate is limited by reactant concentration.

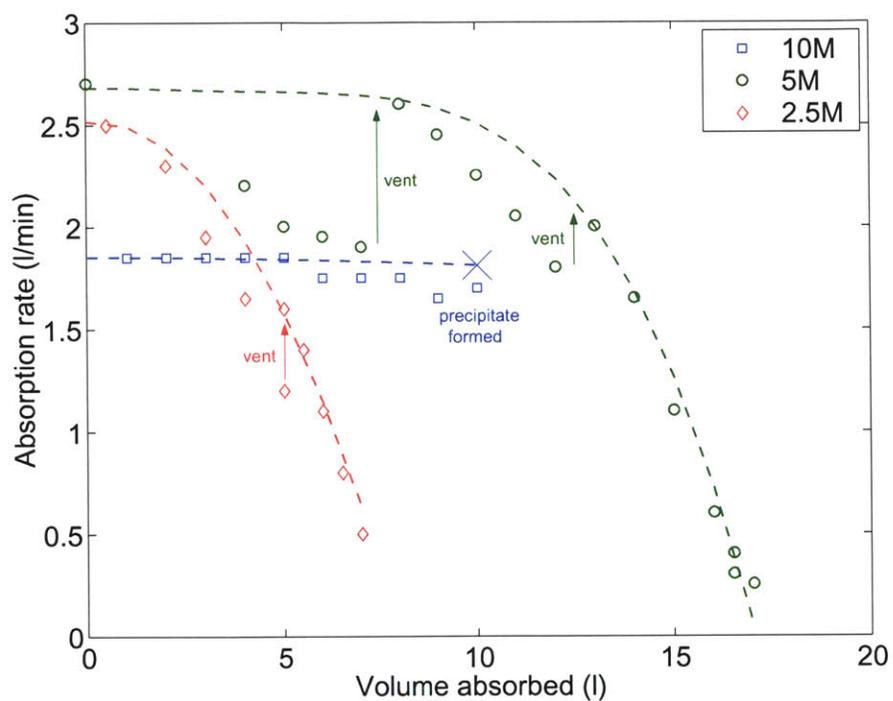


Figure A.1: Graph of absorption rate vs. volume for different NaOH concentrations. The vertical arrows indicate when the bubble was vented to restore 'normal' reaction rate. The 10M experiment was terminated prematurely due to the formation of a precipitate in the chamber.

Discussion Since the reaction is diffusion-limited, the rate is independent of concentration. Consequently an 'optimal' NaOH concentration is trivial to determine, as we are only concerned with maximizing the total capacity of the absorber.

Capacity appears to increase proportionally with concentration. Thus we should employ the maximum concentration possible. However, the precipitate observed with a 10M solution would be very dangerous in regular use, as the particulate matter would damage the valve diaphragms, and deposit in the lines, potentially occluding flow at constrictions points etc.

Conclusion A 5M solution provides plenty of capacity in excess of what we expect to use in a single study (8-10 runs at peak absorbing rate), and this concentration seems to be an excellent compromise between total capacity and the risk of formation of a precipitate.

5M NaOH is mixed by dissolving 200g [5x(23+16+1)] of anhydrous sodium hydroxide in 1000ml of water. The resulting solution has a density of approximately 1.224 kg/liter (empirical data).

A.3 Tracer loss to NaOH

One of the major factors affecting yield is the amount of nitrogen-13 that is lost due to dissolution in the absorber. During receipt of the cyclotron gases, as the carbon dioxide is being absorbed, some of the N13 in the gas bubble dissolves into the sodium hydroxide solution. The consequences of, and measures to minimize this loss are discussed in section 3.7.1.

This section summarizes a series of measurements made in attempt to quantify the magnitude of this effect. Only measurements made after the current compliant chamber was installed are reported below.

Date (m/d/y)	Specific Activity (mCi/ml)		Total Activity (mCi)		Fraction	Bubble (ml)
	Salsa	NaOH	Salsa	NaOH		
2/23/03	0.68	0.019	51.2	4.7	8.5%	>5.0
3/14/03	0.57	0.021	37.0	5.3	12.5%	=4.2
3/14/03	0.61	0.031	39.9	7.7	16.2%	=3.9

These values, while highly variable, show a tracer-loss to the absorber that is broadly in line with the estimate made from figure 3.12.

A.4 Mass Spectrography of Bubble

In this analysis a sample of the gas was taken to a custom-built mass spectrograph in the department of molecular biology at MIT.

A.4.1 Test Setup

Collection of Samples. Two samples of gas were taken. The first sample was taken from a 'dummy run'. The cyclotron target was filled with carbon dioxide, but this gas was not irradiated. The gas was then received in the Salsa Generator, and the resulting bubble manually transferred out of the absorbing chamber into a glass syringe. The second sample should be representative of a real tracer preparation: the carbon dioxide was bombarded in the cyclotron for twenty minutes, following the standard procedure for preparation of N13.

Methodology First the absorbing chamber was filled with fresh sodium hydroxide solution. This solution had been prepared with normal tap water, and allowed to stand in communication with the atmosphere for several hours. Thus it is likely that the solution had equilibrated with the atmosphere, and contained significant quantities of dissolved atmospheric gases.

The tracer preparation system was then put through a standard preparation sequence, in which the gas volume above the sodium hydroxide, and the components connecting this to the injector syringe are flushed with nitrous oxide.

A glass syringe was then connected via a three-way stopcock, and then filled and emptied repeatedly with nitrous oxide. The tubing leading out of the absorbing chamber was then disconnected, and connected to the open port on the stopcock. The lines were then flushed again, and the gas in the sampling syringe discharged through the absorbing chamber. The sodium hydroxide in the absorbing chamber was set so that its surface was at the level with the exhaust port, leaving a pocket of about 0.5ml of flush gas at atmospheric pressure.

The uncooked gas was then received from the cyclotron, in exactly the same manner as a normal radioactive gas run, except that the arrival of the target gas could not be detected with a radiation meter. Instead, the gas was admitted to the absorbing chamber after a delay of 15 seconds, which is the interval that we had previously measured for the arrival of radioactive gas. 1500ml of gas was received, and stirred until the residual bubble reached a minimum volume.

The bubble was then manually extracted to a 10ml glass syringe with an integrated valve (Precision Sampling Corp., Baton Rouge, LA), converted to a luer lock connector.

A radioactive run was then started, with a twenty minute cook time. The apparatus was prepared in the same way, and the gas was received in the same way. The crucial difference between the runs was that the irradiated sample was left in the absorbing chamber for a much longer period of time (2 hours versus 5 minutes).

In both tests, the cyclotron target was filled with pure (grade #5 or better) carbon dioxide gas. In normal use, we have found that the addition of a small amount of nitrogen (~0.3ml) to the target gas improves the product yield. However, we elected not to do this in our tests, because we wanted to be sure that all of the nitrogen detected in the gas bubble was due to other mechanisms.

Mass Spectrography It was our expectation that the only gases present in significant quantities would be carbon dioxide, nitrous oxide, and the various components of air: nitrogen, oxygen, argon, and other gases in trace amounts.

Accordingly we started the analysis by recording spectra for carbon dioxide, nitrous oxide, and room air. These spectra are reproduced below, normalized to the peak of greatest intensity. For each spectrum there are two plots, the first full-range, and the second showing a zoom on the small peaks (0-10%). The scale on the abscissa, 'BDAC', represents the magnetic field which is used; the y-axis shows the relative output of the detector as the field is increased. The relationship of the BDAC scale to the molecular mass of the species being detected is non-linear, and these values cannot simply be scaled to read masses. However, the scale in these measurements is adjusted so that BDAC 1000 corresponds exactly to a molecular mass of 44.

These spectra cover molecular masses from 11 to 46. Although the data are not reproduced here, each of the samples and reference gases were tested up to a molecular mass of 70, but none exhibited any significant peaks above mass 44.

A.4.2 Results

Methodology Firstly the spectra were encoded as feature vectors. Every location in which any peak had $> 0.5\%$ intensity was included. A table of these data are shown below.

BDAC	CO ₂	N ₂ O	Air	Sample A	Sample B
290	2.2			1.3	0.7
355		3.8	4.0	5.3	4.2
415	5.4	2.0	1.5	5.4	3.6
470		0.1	1.2	0.3	
570	0.8			0.2	0.2
697	9.5	11	100	98	47
720			0.7	0.6	0.25
740		18	0.5	8.8	13
780	0.8	1.3	25	27	12
925			2.1	1.8	1.0
1000	100	100	2.1	100	100

Next we attempt to match the spectra of the samples using the spectra of the reference gases. That is we are essentially trying to solve the following equation.

$$x_1 \vec{s}_1 + x_2 \vec{s}_2 + x_3 \vec{s}_3 = \vec{s}_m$$

Where \vec{s}_1 , \vec{s}_2 , \vec{s}_3 are the feature vectors of the reference spectra and \vec{s}_m is the feature vector of the sample of unknown composition. These vectors are each normalized by their sum, in order that the multipliers x_n should be proportional to mass fraction.

Thus we have a set of linear equations, which can be written in matrix form as follows:

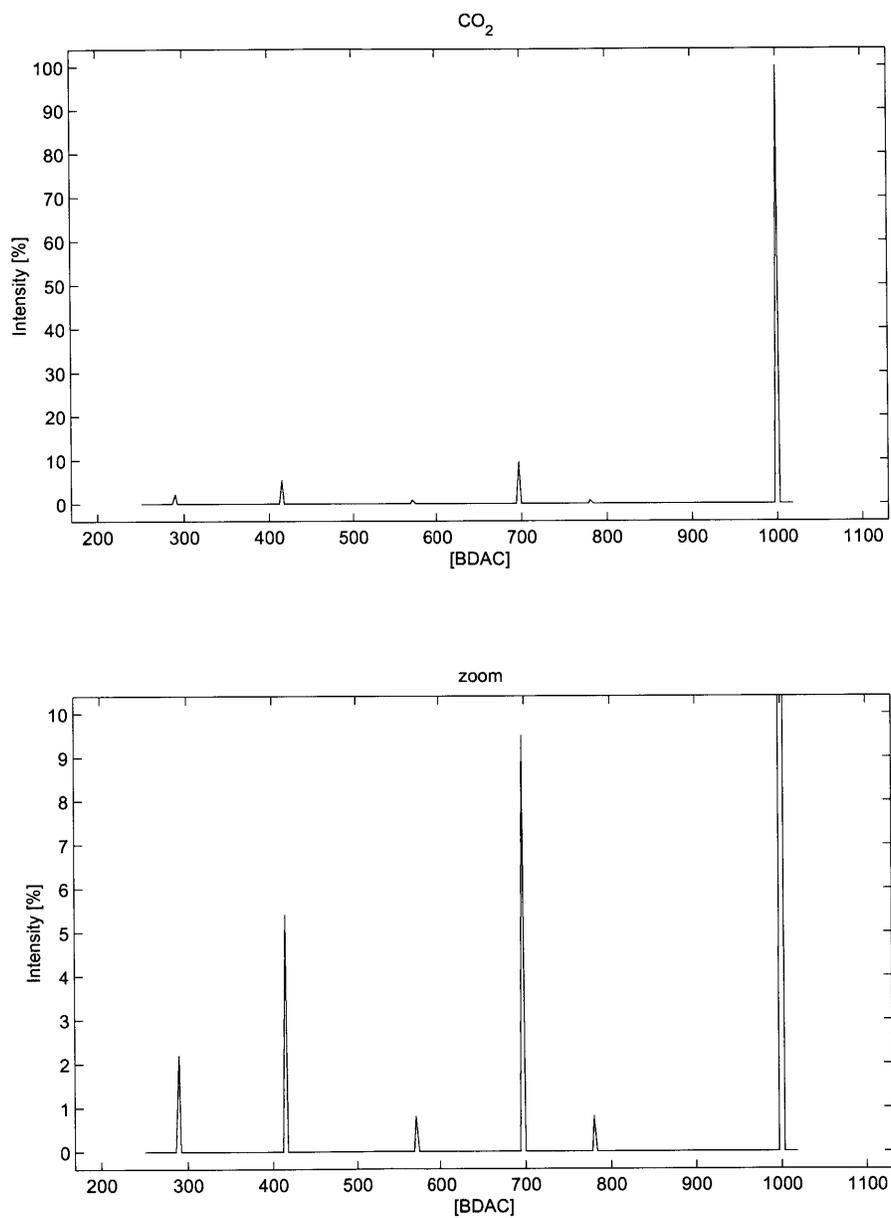


Figure A.2: Mass spectrum of Carbon Dioxide, shown on large scale (above), and in close-up (below) to show smaller peaks.

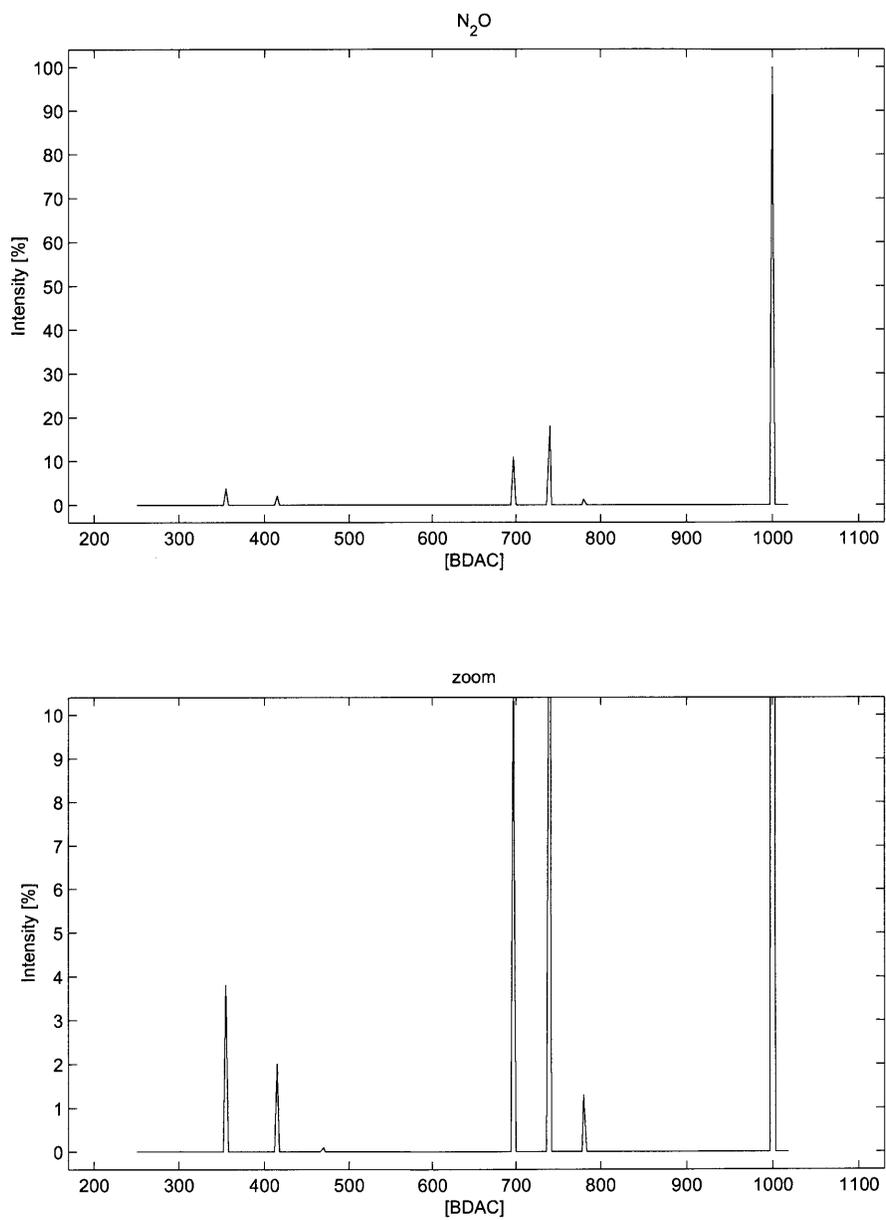


Figure A.3: Mass spectrum of Nitrous Oxide.

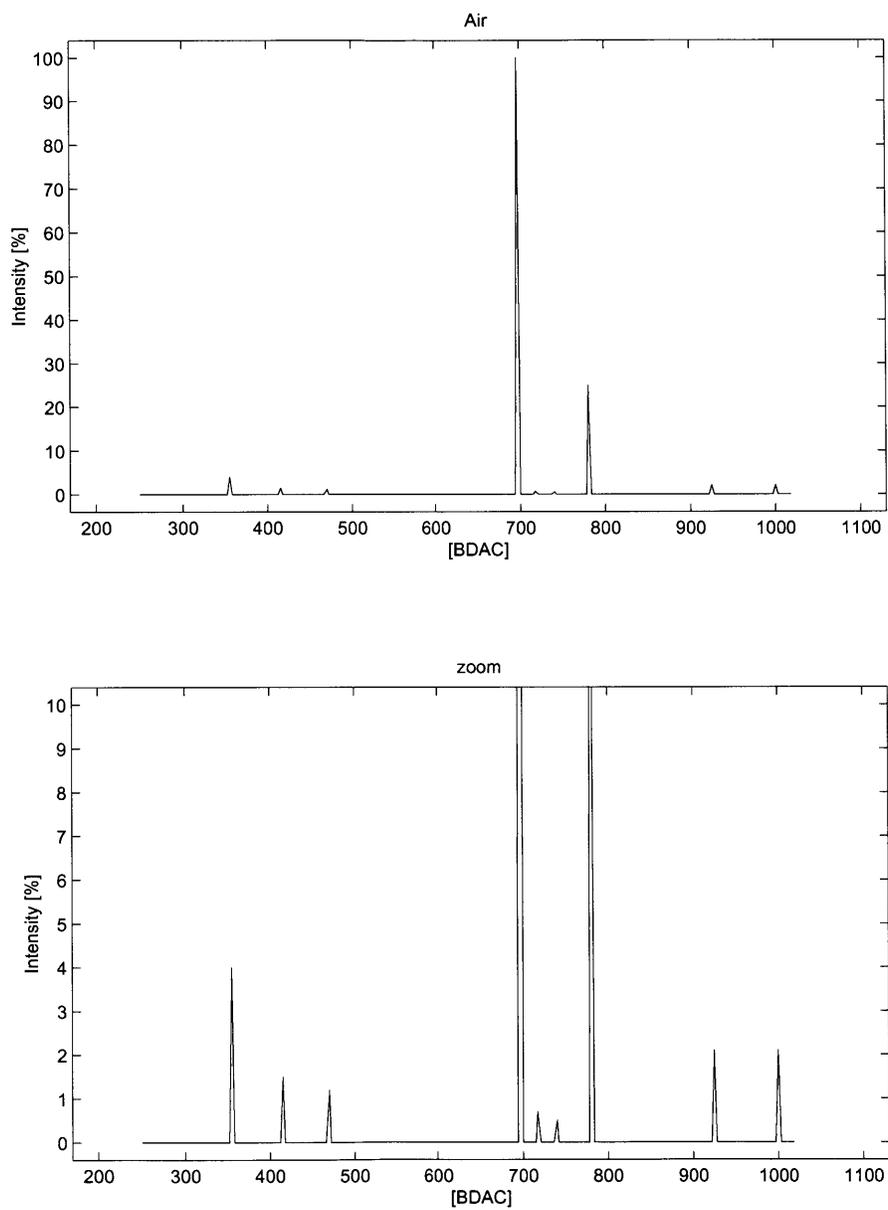


Figure A.4: Mass spectrum of Room Air

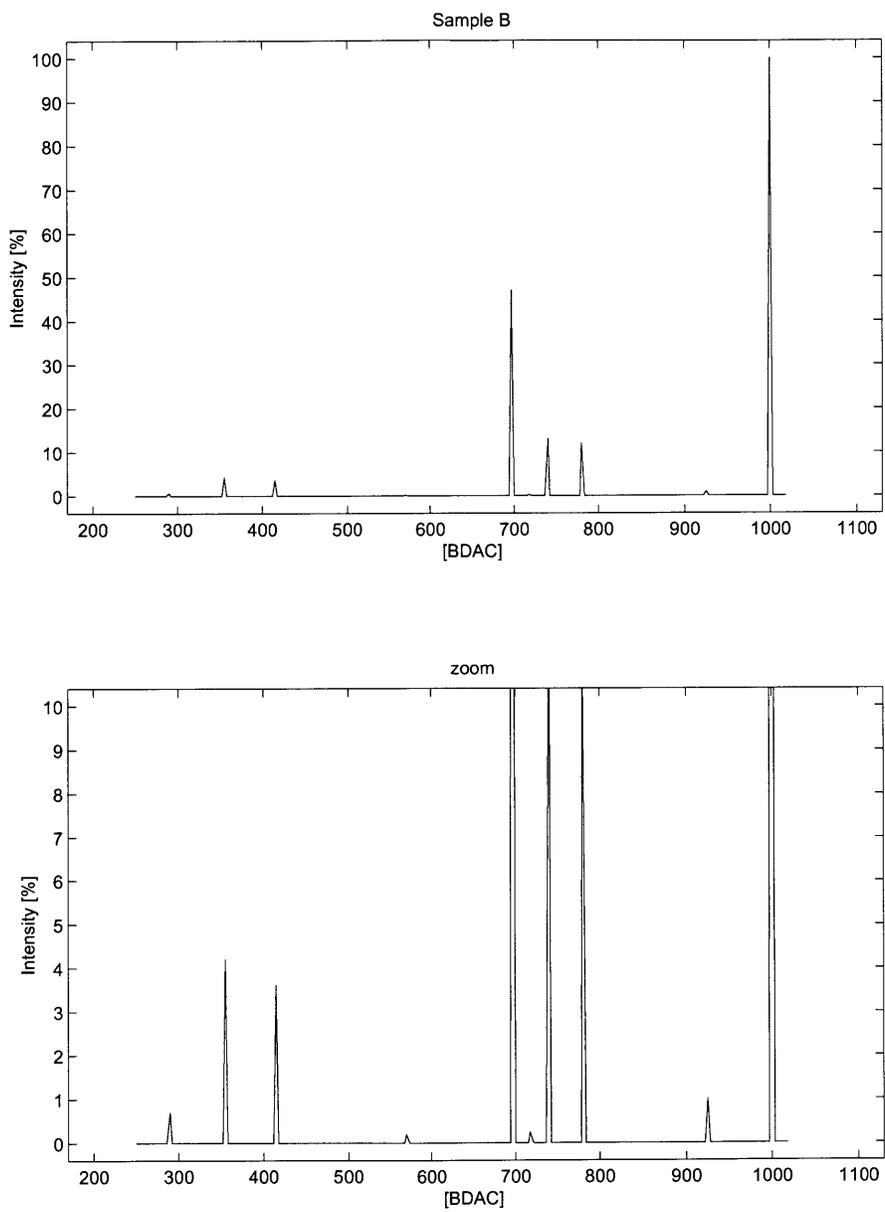


Figure A.5: Mass spectrum of Sample B (uncooked).

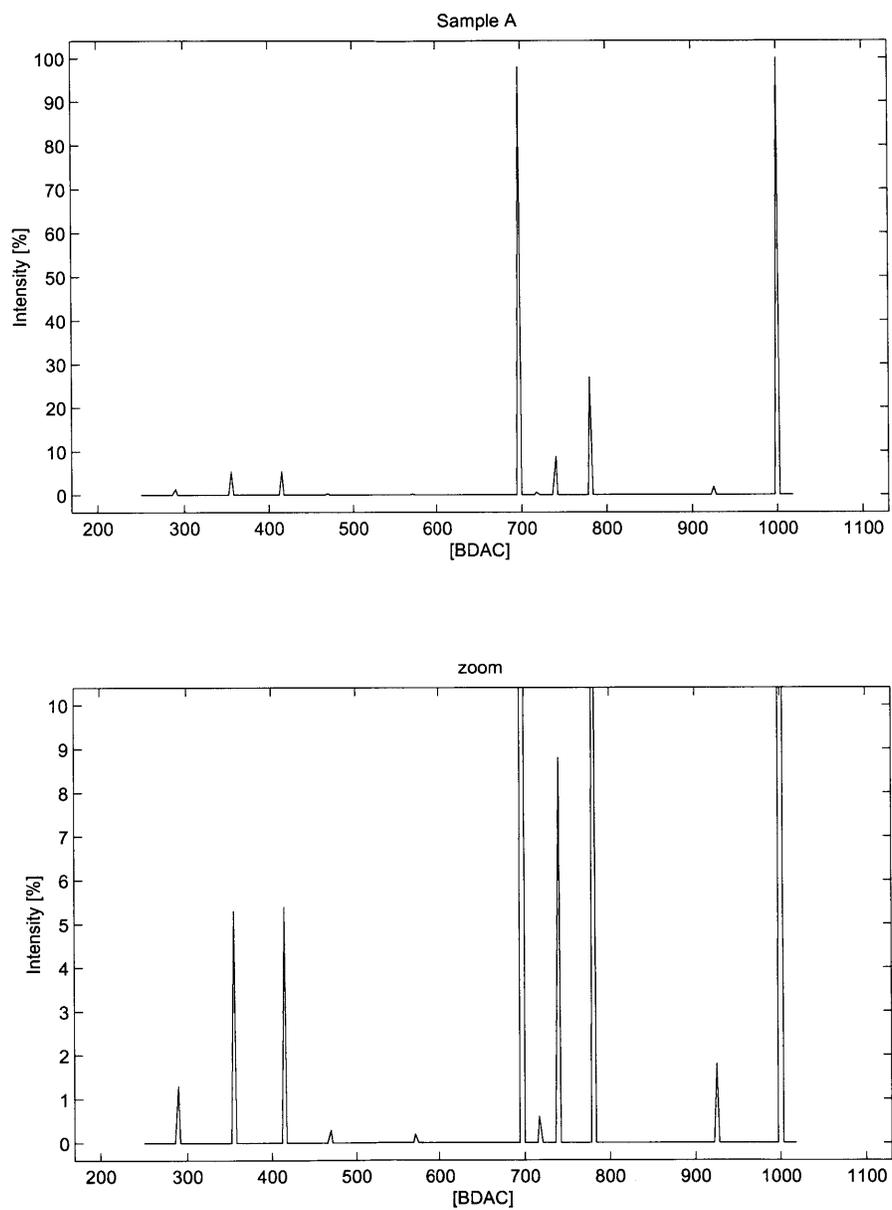


Figure A.6: Mass spectrum of Sample A (irradiated).

$$S\vec{x} = \vec{s}_m$$

This system is over-constrained, and has no solution. Instead we need to minimize some measure of the solution error. The obvious choice is simply a euclidean norm; that is, we wish to minimize $|\vec{S}\vec{x} - \vec{s}_m|_2$ subject to the constraint that $x \geq 0$

Results Using this cost function, composition of the samples was estimated as follows.

Sample	CO ₂	N ₂ O	Air	Total
A	24.81%	25.26%	49.02%	99.08%
B	18.34%	53.21%	27.82%	99.37%

Expected spectra for gas mixes of these compositions are reproduced below. These match to a high degree the real spectra of the gas samples.

A.4.2.1 Interpretation

Consider again the estimated composition of the samples:

Sample	CO ₂	N ₂ O	Air	Total	Error norm
A	24.81%	25.26%	49.02%	99.08%	0.0160
B	18.34%	53.21%	27.82%	99.37%	0.0097

Sample B is the uncooked gas. As expected, there is a high proportion of nitrous oxide, which is used as flush gas throughout. The air content is primarily due to dissolved gas in the sodium hydroxide solution equilibrating with the bubble, but may also be due to residual air in the cyclotron target, and piping. What is slightly surprising is the high level of carbon dioxide in the sample. I would have expected a greater proportion to have reacted with the sodium hydroxide.

Sample A has both a higher air content and carbon dioxide content. The former is probably an artifact of the fact that the bubble was left in the absorbing chamber for a longer period, while the radioisotope decayed. The CO₂ increase presumably reflects the progressive exhaustion of the absorber. Finally, the decreased level of nitrous oxide is likely to do continued dissolution of the gas into the sodium hydroxide.

Note that the differences between these samples appear to be due simply to the length of time that the gas bubbles were left in communication with the sodium hydroxide absorber. This is in keeping with our expectations, as the nitrogen-13 radioisotope is produced in such minute quantities (10^{-7} ml) that we would not expect its presence to be directly measurable. The dominant effect of

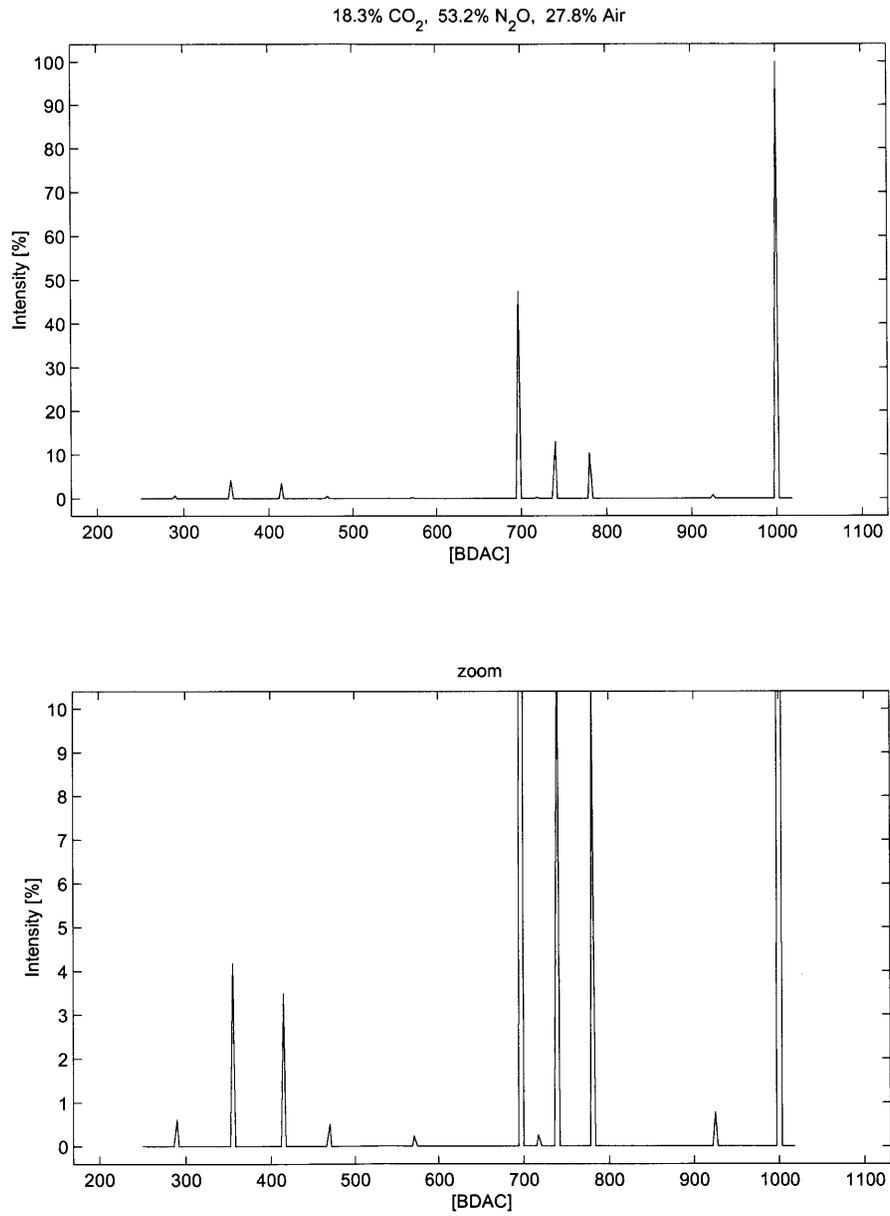


Figure A.7: Expected spectrum of gas mix: 20.3% CO₂, 52.3% N₂O, 27.4% Air

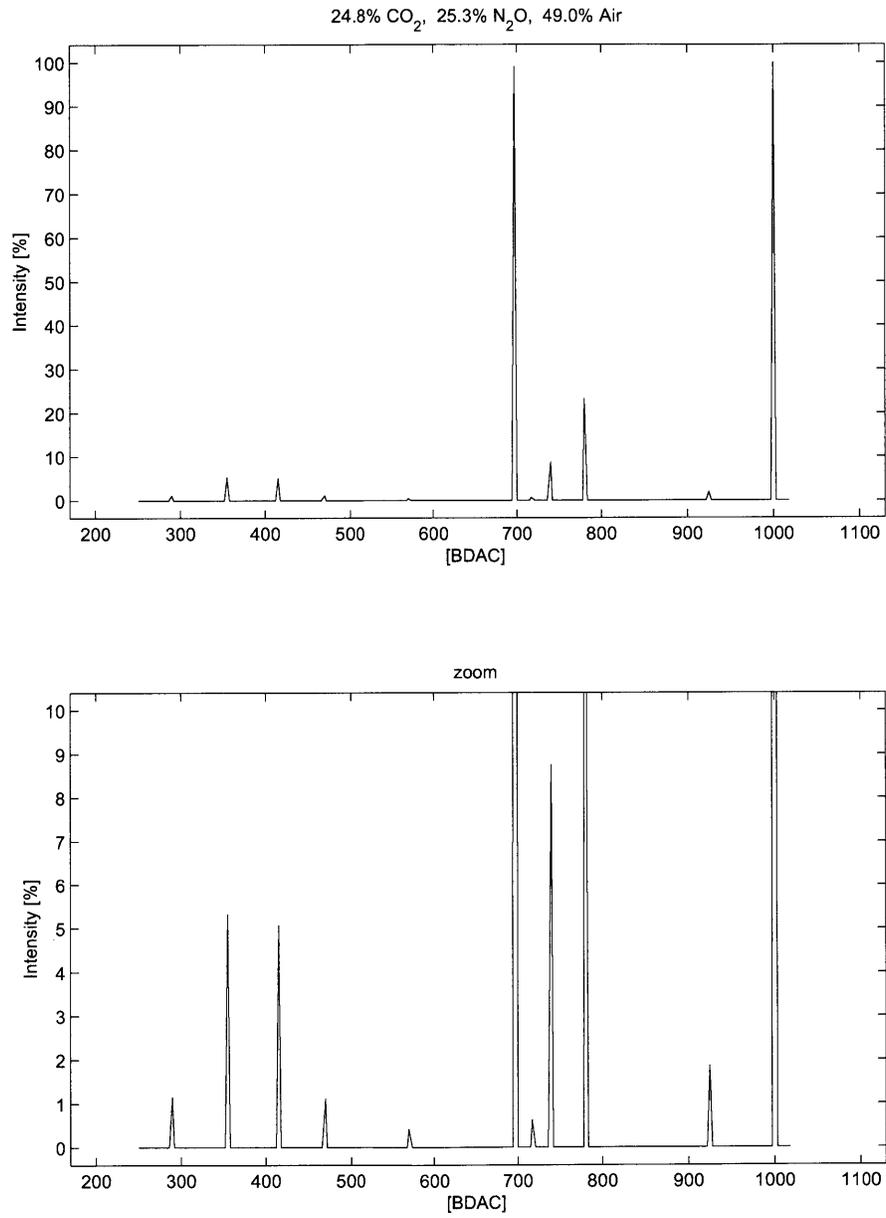


Figure A.8: Expected spectrum of gas mix: 27.3% CO₂, 24.7% N₂O, 48.0% Air

the cyclotron bombardment is to cause dissociation of carbon dioxide into carbon monoxide and oxygen. However we use a custom catalyst upstream of the absorbing chamber which promotes re-formation of the carbon dioxide. In separate experiments, we have measured the peak carbon monoxide concentration in the irradiated gases to be 0.3% by volume. After a similar gas mixture is passed through the catalyst, recombination of carbon dioxide was as complete as could be measured by the instrument we used (accurate to 0.001%), even at high flow-rates.

A.4.2.2 Validity

The composition of the samples, as estimated here, is dependent on the cost function which is minimized in the calculation. The results given above are based on a euclidean error norm:

$$\text{Cost} = \|\vec{\epsilon}\|_2 = \sqrt{\sum \epsilon_i^2}$$

If a different cost function is used, we get slightly different results. Using, for example, a straight error sum:

$$\text{Cost} = \|\vec{\epsilon}\|_1 = \sum \text{abs}(\epsilon_i)$$

In this case, the estimated compositions are as follows:

Sample	CO ₂	N ₂ O	Air	Total	Error norm
A	24.68%	25.44%	48.85%	98.61%	0.0225
B	18.28%	53.30%	27.51%	99.09%	0.0150

In all cases, the total error is slight, of the order of 1-2%. These data imply that we should have a high degree of confidence in our results. Any other gases can only be present in minute quantities. Additionally, the magnitude of total error is so small that slight variation in a reference gas could account for the change. For example, the air sample is relatively high in carbon dioxide, which may reflect the fact that it was sampled indoors, in a poorly-ventilated area.

It is crucial to note that none of the samples exhibits a peak in a location that the reference gases do not, and all were tested up to mass 70. This means that any other species present must, at the minimum, have both fragments of coincident mass, and fundamentals that are either coincident with these peaks, or have molecular mass greater than 70.

A.5 Target Priming

As discussed in 3.10.2, we observed a sudden drop in yield when the cyclotron switched to a higher-purity tank of carbon dioxide. This led us to investigate how adding contaminants to the the target increased the amount of radioactive material produced.

Our early tests suggested that only (non-radioactive) nitrogen, when added to the target, made a significant difference to the tracer yield. Obtaining data of reasonable quality was confounded, however, by the very high variability in the yield, even when the amount of contaminant was not changed.

Accordingly a series of experiments were performed using a highly simplified apparatus designed to remove as many sources of variability as possible.

A.5.1 Test setup

A.5.1.1 Target priming apparatus

In order to reproducibly add small volumes of nitrogen, a six-port, dual-position valve (VICI Chem-inert model C22-6186) was used to switch calibrated sample loops in and out of the line from the carbon dioxide tank to the cyclotron target.

The valve is configured as shown in figure A.9. Before the target is filled, nitrogen is allowed to flow through the sample loop at a high flow-rate for several seconds, so that the internal volume of the loop is completely filled with nitrogen. Flow is stopped and the loop allowed to decompress to atmospheric pressure. The valve is then switched so that the loop is now inserted into the line to the target, and the target is then filled with carbon dioxide, flushing the small volume of nitrogen through the line into the target volume.

A.5.1.2 Measuring radioactive yield

A line was connected from the cyclotron exit valve straight into a glass vessel located in a well-counter. A second line leads out of the vessel to carry exhaust gas well away from the counter. The measuring apparatus is contained in a fume hood.

After bombardment, the valve is opened, and the cyclotron target allowed to depressurize through the glass vessel. Within a minute (time constant ~ 10 s), the target pressure drops from ~ 100 psi to atmospheric. At this point, since the volume of the vessel is small relative to the (decompressed) target volume, we can reasonably assume that the gas remaining in the glass vessel is a representative sample of the target gas. Furthermore, with pressure equilibration complete, convection does not occur, and diffusive effects down the long, narrow lines we assume to be minimal. Thus the amount of radioactive material in the glass vessel should be stable, and decrease only by spontaneous decay.

For each test run, a varying volume of nitrogen was introduced into the target, and bombarded for a fixed duration of 10 minutes at a constant beam current of $30 \mu\text{A}$. Immediately following bombardment, the exit valve was opened, and a series of measurements of activity within the well-counter were manually recorded.

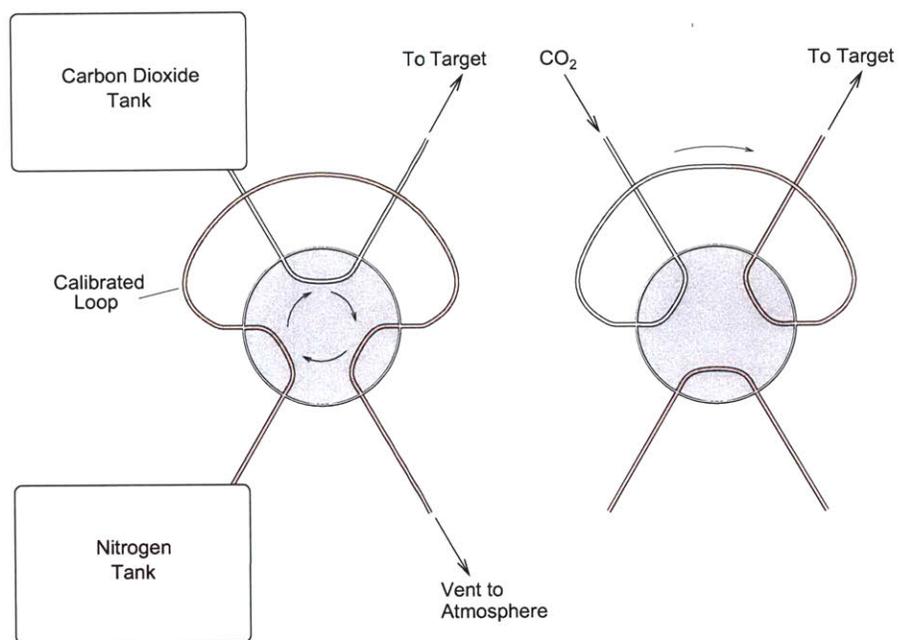


Figure A.9: Equipment used to add small volumes of nitrogen to the target gas. A calibrated sample loop is filled with nitrogen (left) and switched into the line to the target.

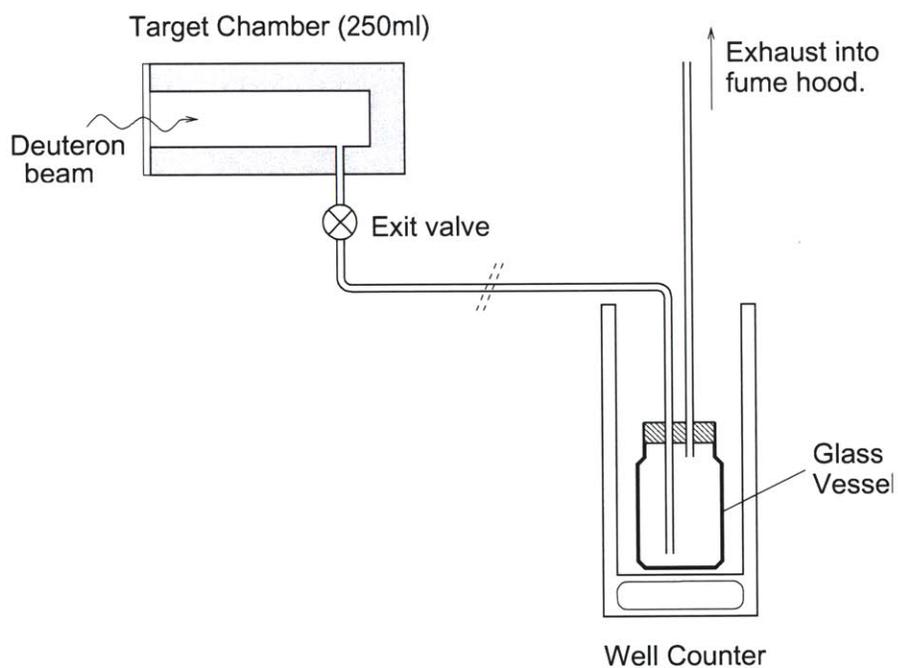


Figure A.10: Apparatus for measuring effect of adding nitrogen to the cyclotron target. Following bombardment, the target is allowed to depressurize through a small vessel located in a well-counter. Once the system has equilibrated, a sample of the target gas remains in the vessel, and its radioactivity is measured with the well-counter.

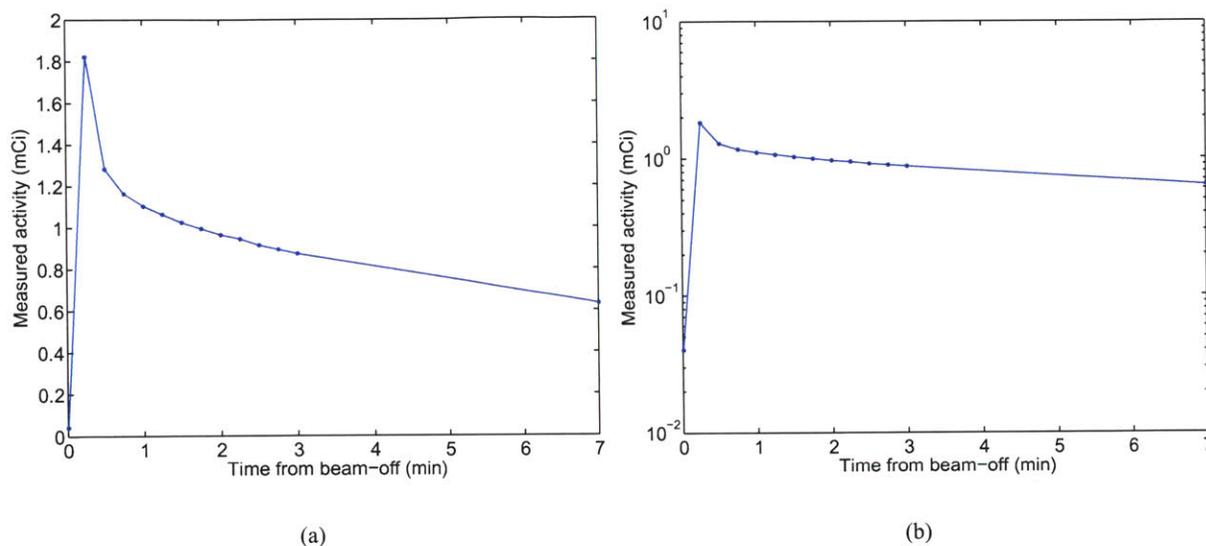


Figure A.11: Example measurements from a target priming run, on a linear scale (left) and logarithmic scale (right).

A.5.2 Results

Examples of the measurements from a single run are shown in figure A.11. As the target is vented through the vessel, the pressure is transiently elevated, the mass of radioactive material within the vessel is increased due to compression of the gas. This results in a brief spike in the measured activity, which rapidly falls to a steady exponential decay, once pressure equilibration is complete.

A single measure of activity generated is calculated by fitting a straight line through the log of the later part of the measured activity data, and evaluating the resultant exponential at a fixed reference time ($t=3\text{min}$). Results processed in this way are shown in figure A.12.

It is evident from these results that the addition of nitrogen to the target has a large effect on the radioactive yield. The data, however, are surprisingly variable, considering that every effort was made to remove sources of noise.

A.5.3 Existence of Second Species?

About half-way through these tests, it became clear that the well-counter reading was not returning to zero between run and the next. A set of data taken from such a run is shown in figure A.13. Starting at $t = 3\text{min}$, the target and test vessel are continuously flushed with carbon dioxide. Initially, the amount of activity drops very rapidly, but soon reaches a stable level where the measured activity

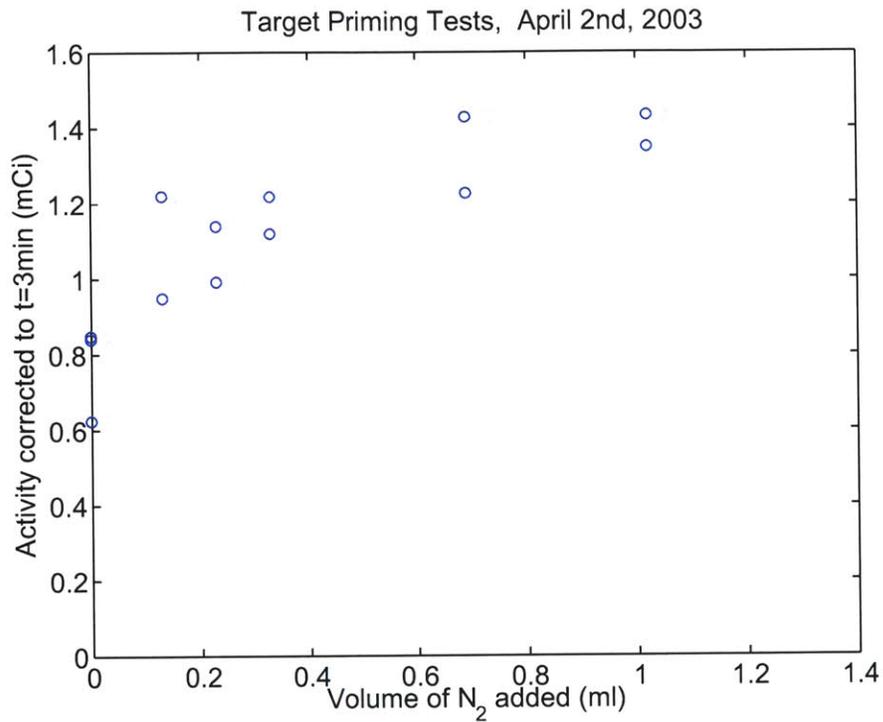


Figure A.12: Amount of radioactivity generated in the cyclotron target, plotted against the amount of nitrogen added. Despite a high degree of variability, it is apparent that addition of nitrogen to the target gas can increase the radioactive yield by 50-100%. The numbers adjacent to each point indicate the sequence in which runs were performed.

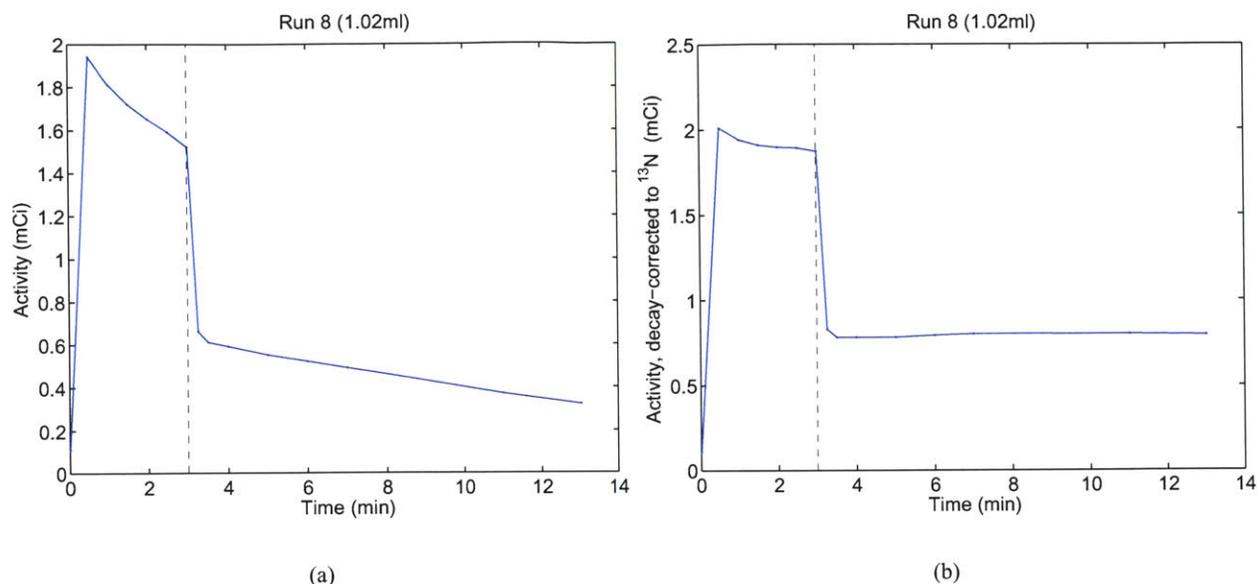


Figure A.13: Measurements taken from a run in which activity remained, apparently in the test vessel, even during continuous flushing. On the left is the raw data; on the right, the same data are shown corrected for the decay of nitrogen-13.

drops slowly. If we decay-correct these data for the half-life of 9.97 minutes, this slow decrease becomes horizontal, indicating that this is merely radioactive decay of the N13.

It appears that a significant fraction of the nitrogen-13 is being somehow retained in the test vessel. Further insight into this phenomenon can be gained by plotting all of the measurements on a single, continuous timescale [A.14].

We see now that the activity that remains after the 9th run (starting at 17:02), which we might assume originated in this run, is clearly the same activity that remained from the previous run, which has been gradually decaying. The salient feature of this run is that no nitrogen was added to the target, whereas 1.02ml was added in the previous run. This suggests that radioactivity is only retained in the test vessel when nitrogen is added to the target before bombardment. Sure enough, in the next run, when 0.69ml was added, the amount of residual activity increases again.

We can now process the data firstly to remove from the measurement of total nitrogen-13 generated the offset due to 'sticky' nitrogen-13 retained from previous runs; secondly we can calculate how much the amount of 'sticky' nitrogen-13 increased with each run.

The reprocessed data are shown in figure A.15. This graph suggests that the amount of 'sticky' nitrogen-13 generated in each run is approximately proportional to the volume of nitrogen-14 added to prime the target. This trend is evident, at least, from runs 8 onward. In runs 1-5 (no data recorded for the 6th run), however, the amount of retained N13 is negligible. Since no changes were made

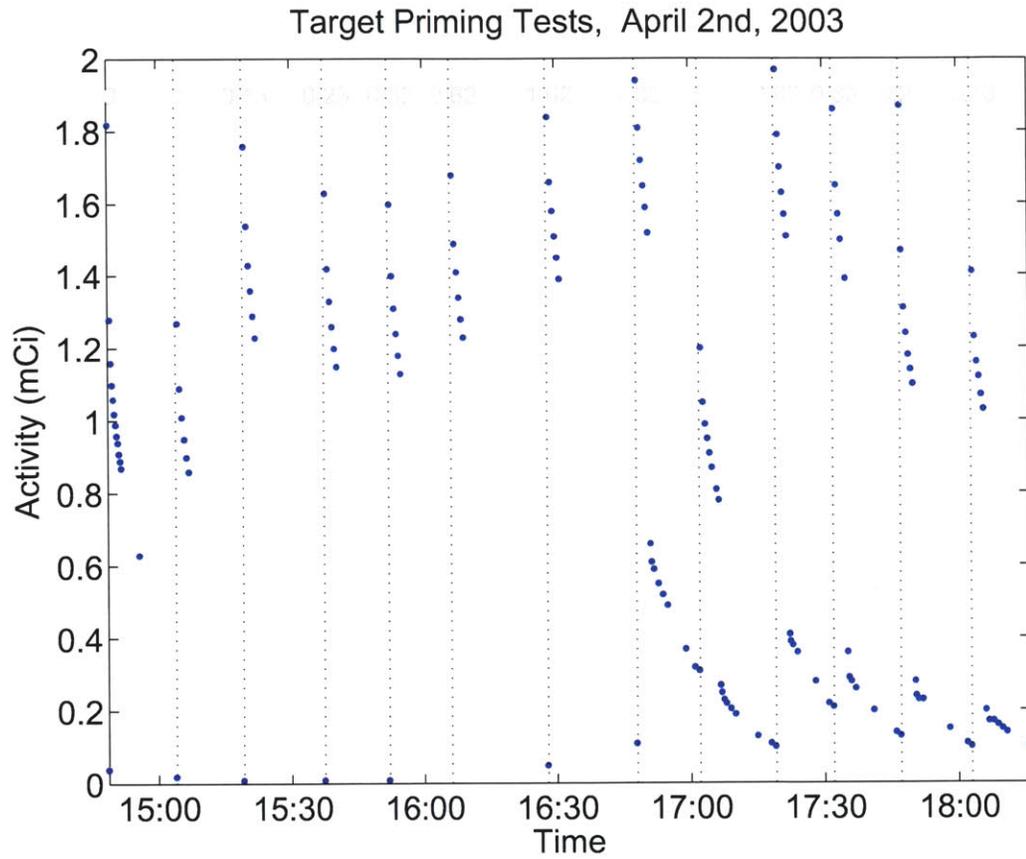


Figure A.14: Data from all of the runs assembled onto a single time-scale. This presentation enables us to observe that the residual activity following run 9 (0ml) is, in fact, the same activity generated by run 8 (1.02ml), as it continues to decay at the N13 half-life.

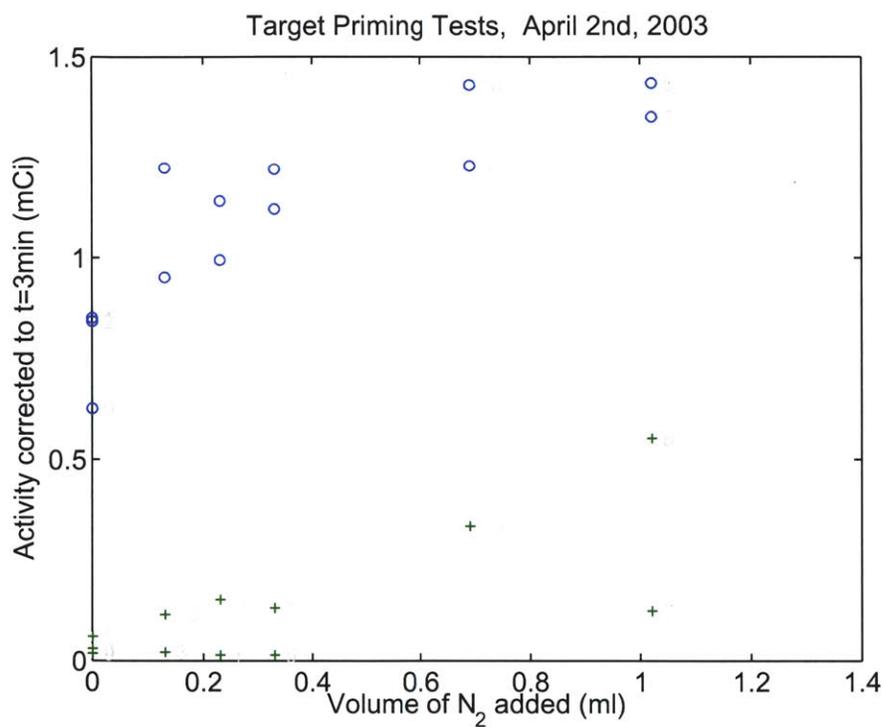


Figure A.15: Reprocessed presentation of yield data, showing, for each run, the total amount of nitrogen-13 generated *de novo* (blue circles), and the increase in the 'sticky' form (green crosses).

during the experiment to the measuring apparatus, to the cyclotron target, or in the bombardment, except that sample loops of differing volumes were switched in and out, it is difficult to understand the discontinuity in the results.

What is going on here? The evidence suggests that two distinct forms of nitrogen-13 are being generated. One appears to be gaseous and non-reactive, as we would expect. The other form, which is only observed when nitrogen-14 is added to the target gas, seems to be reacting with the test vessel. Since the test vessel is glass, it seems likely that the reaction is with the polyethylene lid (or potentially the lines or connectors). The evidence points to a chemical reaction, rather than passive diffusion, as stopping and restarting the flush had no discernible effect on the observed decay rate. Additionally, if a diffusive mechanism were at work, we would expect that the nitrogen-13 that had diffused into the bulk of the container would gradually diffuse out, and be convected away, so that the apparent decay rate would always be faster than the spontaneous radioactive decay rate.

It must be stressed that this speculation is based on data that have not been reproduced and which are of questionable quality. Although more investigation is warranted, the matter is not pursued further in this thesis.

A.5.4 Data from Earlier Tests

The data shown above are the results of a third set of similar experiments; tests were performed on two earlier occasions, with results that were less reliable. On the first occasion, nitrogen was not added to the target using a sample-loop, but was manually injected into the line using a syringe. On the second occasion, the sample loop system was employed, but the test vessel had a large leak, so that the results were compromised by a ‘cloud’ of radioactive gas that filled the well-counter. Nevertheless, the results, when processed, also reflect the trend shown in figure A.12 above, and a summary is reproduced here [Figure A.16] for completeness and to affirm the validity of the finding described above.

Note also that due to changes in the equipment used (and small differences in the data processing) the vertical scales of each of the graphs in figures A.16 and A.12 are not comparable.

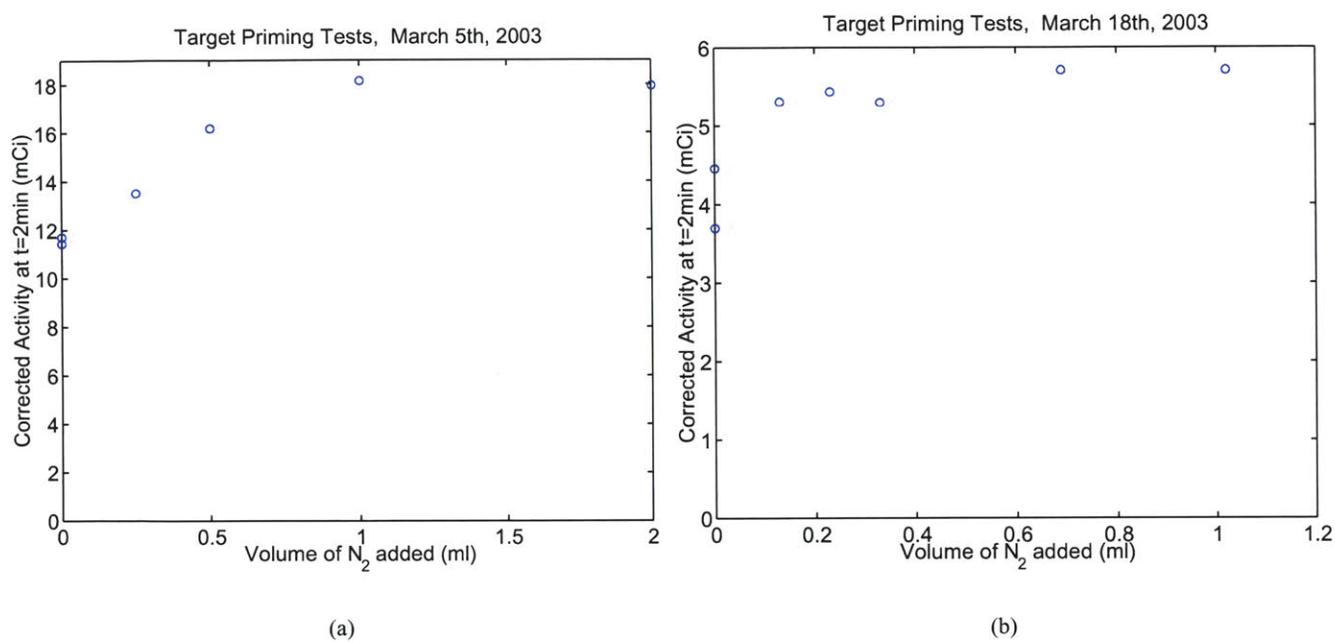


Figure A.16: Summary of results from earlier tests: these graphs show similar information to figure A.12, but due to experimental flaws, have uncertain validity.

Bibliography

- [1] T. A. Altes, P. L. Powers, J. Knight-Scott, G. Rakes, T. A. Platts-Mills, E. E. de Lange., B. A. Alford, J. P. r. Mugler, and J. R. Brookeman. Hyperpolarized ^3He MR lung ventilation imaging in asthmatics: preliminary findings. *J Magn Reson Imaging*, 13(3):378–384, Mar 2001.
- [2] R. C. Anafi and T. A. Wilson. Airway stability and heterogeneity in the constricted lung. *Journal of Applied Physiology*, 91:1185–1192, 2001.
- [3] P. Atkins. *Physical Chemistry*. W.H.Freeman and Company, New York, 6th edition, 1998.
- [4] L. H. Brudin, C. G. Rhodes, S. O. Valind, P. D. Buckingham, T. Jones, and J. M. Hughes. Regional structure-function correlations in chronic obstructive lung disease measured with positron emission tomography. *Thorax*, 47(11):914–921, Nov 1992.
- [5] J. C. Clark and P. D. Buckingham. *Short-lived Radioactive Gases for Clinical Use*. Butterworths, London and Boston, 1975.
- [6] R. S. Harris, D.-B. Willey-Courand, C. A. Head, G. G. Galletti, D. M. Call, and J. G. Venegas. Regional VA, Q, and VA/Q during PLV: effects of nitroprusside and inhaled nitric oxide. *J Appl Physiol*, 92(1):297–312, Jan 2002.
- [7] S. C. Jones, W. M. Bucelewicz, R. A. Brissette, R. Subramanyam, and B. J. Hoop. Production of ^{13}N -molecular nitrogen for pulmonary position scintigraphy. *Int J Appl Radiat Isot*, 28(1-2):25–28, Jan 1977.
- [8] Y. Kimura, H. Hsu, H. Toyama, M. Senda, and N. M. Alpert. Improved signal-to-noise ratio in parametric images by cluster analysis. *Neuroimage*, 9(5):554–561, May 1999.
- [9] Y. Kimura, M. Senda, and N. M. Alpert. Fast formation of statistically reliable FDG parametric images based on clustering and principal components. *Phys Med Biol*, 47(3):455–468, Feb 2002.
- [10] T. Lango, T. Morland, and A. O. Brubakk. Diffusion coefficients and solubility coefficients for gases in biological fluids and tissues: a review. *Undersea Hyperb Med*, 23(4):247–272, Dec 1996.

- [11] C. S. Levin and E. J. Hoffman. Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. *Phys Med Biol*, 44(3):781–799, Mar 1999.
- [12] D. R. Lide, editor. *CRC Handbook of Chemistry and Physics*. CRC Press.
- [13] S. M. Mijailovich, S. Treppo, and J. G. Venegas. Effects of lung motion and tracer kinetics corrections on PET imaging of pulmonary function. *J Appl Physiol*, 82(4):1154–1162, Apr 1997.
- [14] T. Munkner and A. Bundgaard. Regional V/Q changes in asthmatics after exercise. *Eur J Respir Dis Suppl*, 143:62–66, 1986.
- [15] D. Y. Rosenzweig, J. M. Hughes, and T. Jones. Uneven ventilation within and between regions of the normal measured with nitrogen-13. *Respir Physiol*, 8(1):86–97, Dec 1969.
- [16] J. G. Venegas and G. G. Galletti. Low-pass filtering, a new method of fractal analysis: application to PET images of pulmonary blood flow. *J Appl Physiol*, 88(4):1365–1373, Apr 2000.
- [17] M. F. Vidal Melo., R. S. Harris, D. Layfield, G. Musch, and J. G. Venegas. Changes in regional ventilation after autologous blood clot pulmonary embolism. *Anesthesiology*, 97(3):671–681, Sep 2002.
- [18] P. D. Wagner, D. R. Dantzker, V. E. Iacovoni, W. C. Tomlin, and J. B. West. Ventilation-perfusion inequality in asymptomatic asthma. *Am Rev Respir Dis*, 118(3):511–24, Sep 1978.
- [19] D. B. Willey-Courand, R. S. Harris, G. G. Galletti, C. A. Hales, A. Fischman, and J. G. Venegas. Alterations in regional ventilation, perfusion, and shunt after smoke inhalation measured by PET. *J Appl Physiol*, 93(3):1115–22, Sep 2002.