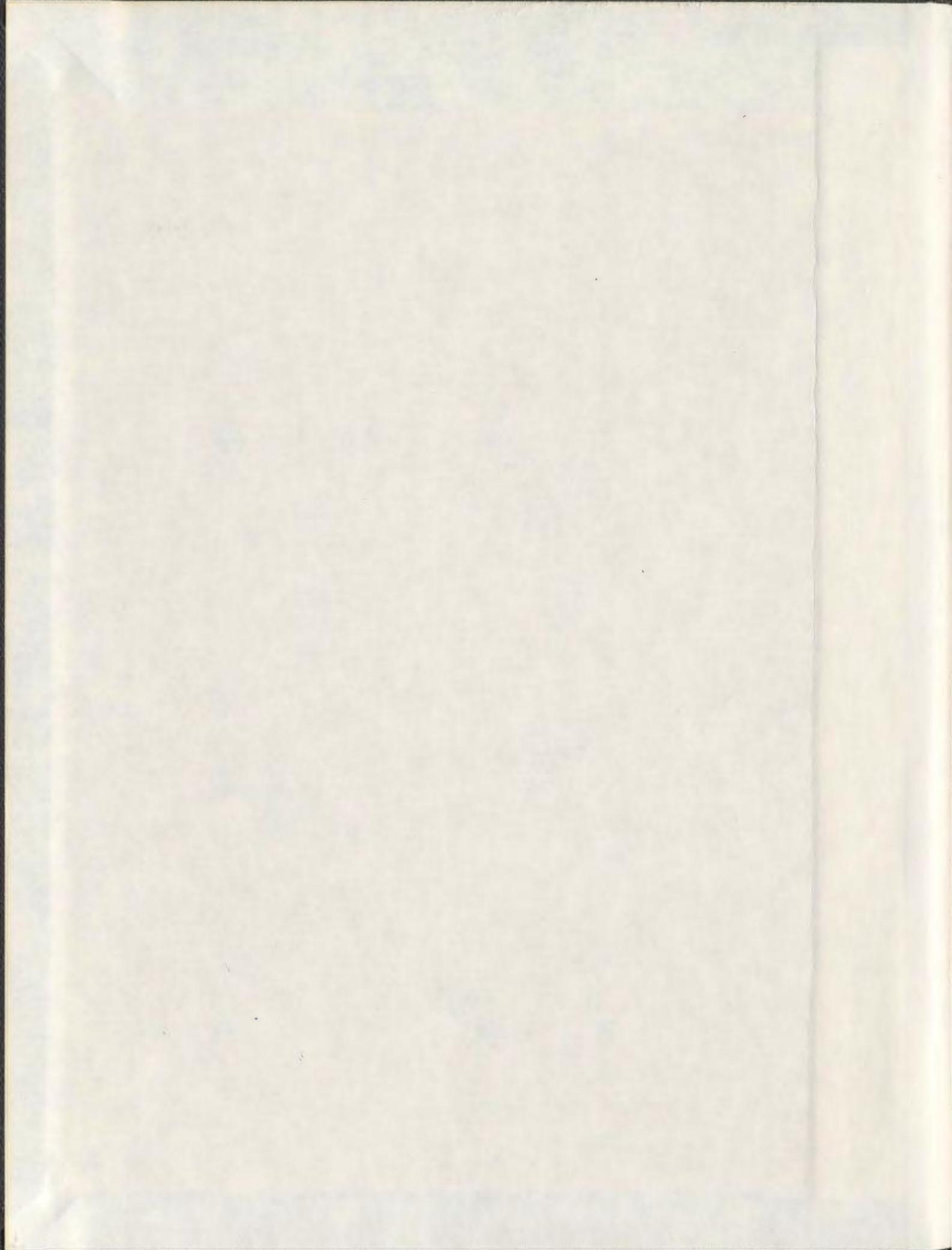


APPLICATION OF PHYSICS AND MATHEMATICS
IN CLINICAL NUCLEAR MEDICINE

LIN LING



APPLICATION OF PHYSICS AND MATHEMATICS IN CLINICAL
NUCLEAR MEDICINE

by

© Lin Ling

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy.

Division of Biomedical Sciences
Memorial University of Newfoundland

September 2013

ST. JOHN'S

NEWFOUNDLAND AND LABRADOR

Abstract

This PhD research project consisted of two parts. In the first part a new pharmacokinetic model was introduced to improve the accuracy of kidney function estimation based on a sampling schedule of 2 and 4 blood sample measurements. Previous models such *SET1* and *SET2* have been shown to be unreliable with respect to the choice of sampling schedule and in some cases provide physiologically impossible values of the glomerular filtration rate (*GFR*). The new model called Tk-GV uses a Tikhonov regularized gamma variate function to fit the plasma clearance data. Based on a group of 46 patients, a comparison of four pharmacokinetics models (*SET1*, *SET2*, OLS-GV, and Tk-GV) revealed that the Tk-GV model was the most robust with respect to sample size and sampling schedule and provided no physiologically impossible values of *GFR*. When compared to constant infusion results in the literatures, the Tk-GV model was shown to eliminate the proportional overestimation of *GFR* produced by the *SET1* model given by Chantler's correction and the constant overestimation produced by the *SET2* model.

Application of the Tk-GV model for estimating a patient's *GFR* requires the collection of 4 blood samples. In cases where it is not possible to collect four blood samples, two blood samples are collected and the patient's *GFR* is calculated using the *SET1* model. A criterion called the renal sufficient index (*RSI*) was developed by others using a group of children to correct the overestimation of *GFR* as obtained

from the *SET1* model. In this work the *RSI* was applied to a mixed group of 26 patients (adults, children, males, and females) and shown, based on additional medical information, to be able to accurately distinguish between patients with normal and abnormal *GFR*.

Bone mineral density (BMD) is important for evaluating bone health, especially for elderly people such as post-menopause women and men with prostate disease. However, analyzing a patient's BMD change is difficult because of naturally occurring short-term and long-term fluctuations in a patient's BMD, and also the short-term and long-term errors provided by the equipment used to measure BMD. The World Health Organization has provided a least significant change (LSC) criterion for distinguishing between normal and abnormal changes in a patient's BMD value. Unfortunately the LSC criterion only accounts for short-term machine error. In the second part of this thesis a new criterion called the total detectable difference (TDD) is introduced for analyzing changes in BMD in consideration of naturally occurring changes in the patient's BMD, and instrumentation error occurring on both short and long time scales. Based on the analysis of a group of 8,800 patients, a TDD value of approximately 0.045 g/cm² is obtained for the hip and 0.060 g/cm² for the lumbar spine. Based on the study of 9,379 patients, it is shown that the LSC criterion has a potential of over-diagnosing BMD change by approximately 30% compared to the TDD criterion.

Acknowledgments

I would like to thank Rick Scanlan from the Department of Nuclear Medicine for helping me with the laboratory equipment and radiation safety training. I also would like to thank Stan, Kelly, Tara, Maureen, Nina, and Tammy from the Department of Nuclear Medicine, who have taught me a great deal about clinical practices.

I greatly appreciate the support given to me by the Faculty of Medicine at Memorial University and the department of Nuclear Medicine at Eastern Health.

I would like to thank my M.Sc. supervisor Dr. Jolanta Lagowski from the Department of Physics and Physical Oceanography at Memorial University for her help in providing me with the necessary skills for computer modeling and for conducting basic research.

I would like to thank Dr. Christopher Kovacs for helping me access the CaMos data and the support from the CaMos Research Group.

Finally I would like to thank my husband Luc Beaulieu, my parents Ling Chuanming and Liu Jing in China and my beautiful baby boy Patrick for all their love and support.

Structure

This thesis includes two research projects. 60% of this thesis focuses on the plasma clearance model which includes an introduction, literature review, data and analysis methods, results and discussion and possible applications. The remaining 40% of this thesis introduces a new criterion for analyzing bone mineral density change as described above. The outline of the thesis is listed as follows:

1. Chapter 1 gives a general introduction of Nuclear Medicine. The topics include radioisotopes, radiation detection, instrumentations, and tracer kinetics.
2. Chapters 2 to 7 present the first project of this thesis. Chapters 8 to 10 present the second project of this thesis.
3. Chapter 2 provides a theoretical background and motivation of the plasma clearance project. This chapter provides a brief overview of the clinical assessment of kidney function and a literature review of the major pharmacokinetic models along with a description of their performance.
4. Chapter 3 outlines the data and analysis methods used to investigate the proposed model. The uncertainty of the measurement and numerical methods such as the Nelder Mead minimization method are discussed.

5. Chapter 4 presents the results of the calculations applied to 46 patients using the *SET1*, *SET2*, OLS-GV and Tk-GV models using the full samples and subsets.
6. Chapter 5 compares the performance of the *SET1*, *SET2*, OLS-GV and Tk-GV models including the estimated glomerular filtration rate (*GFR*), the effects of subset, case study and a cross comparison with other plasma clearance methods.
7. Chapter 6 applies the Tk-GV model to 24 h data.
8. Chapter 7 presents a validation study of a kidney function correction method to compensate for the inaccuracy of the two-sample *SET1* method.
9. Chapter 8 introduces the concept of bone mineral density (BMD). The proposed method of analyzing time-based BMD results is introduced after a literature review. The motivation and the theory of the new criterion, total detectable difference (TDD), are provided.
10. Chapter 9 lists the data and analysis methods used for investigating the new TDD criterion.
11. Chapter 10 presents the results and discussion of the TDD criterion based on 8,800 patients dataset. Details about the patient selection, data analysis and clinical impact are listed.
12. Chapter 11 summarizes the major results and implications of both research projects and describes future research projects.
13. Appendix A shows the clinical protocol implemented for collecting 4 blood samples at the General Hospital in St. John's, NL and the mathematical formulae for the theory of the Tk-GV model.

Contents

Abstract	ii
Acknowledgments	iv
Structure	v
List of Tables	xvi
List of Figures	xxiii
Abbreviations	xxiv
1 Introduction	1
1.1 Nuclear Medicine	1
1.2 Radioisotopes	2
1.2.1 Decay and Half-life	3
1.2.2 Interactions with Matter	5
1.3 Radiopharmaceuticals	6
1.3.1 Chelation	9
1.4 Gamma Counter	10
1.5 Theory of SPECT	13

1.6	Tracer Kinetic Modeling	16
1.7	Bone Densitometer	17
1.8	My Contributions	23
2	Glomerular Filtration Rate (<i>GFR</i>) Determination	27
2.1	Glomerular Filtration Rate	27
2.1.1	Impact of Kidney Function	30
2.2	<i>GFR</i> Measurements	32
2.2.1	Biomarkers	32
2.2.2	Estimated <i>GFR</i> (<i>eGFR</i>)	36
2.2.3	Renography	38
2.3	Plasma Clearance	39
2.3.1	Pharmacokinetic Models	41
2.3.2	One-compartment Model	45
2.3.3	Two-compartment Model	48
2.4	Limitations of the Compartment Models	50
2.5	Motivation	55
2.6	Theory	55
2.6.1	Gamma Variate (<i>GV</i>) Model	55
2.6.2	Rate of Exchange, <i>GFR</i> and <i>Vol</i>	58
2.6.3	Ill-posed Problem	60
2.6.4	Tikhonov Regularization	61
3	Data and Analysis Methods	68
3.1	Data	68
3.2	Measurement Uncertainty	70
3.3	Numerical Methods	71

3.4	Regression Methods	72
4	Results	74
4.1	<i>SET1</i> Results	74
4.2	<i>SET2</i> Results	83
4.3	OLS-GV Results	88
4.4	Tk-GV Results	92
5	Discussion	96
5.1	<i>GFR</i> Estimation	96
5.1.1	Non-physical Results	96
5.1.2	Variations within the Model	98
5.2	Effect of Subsets	101
5.2.1	4-sample Subsets	101
5.2.2	Hump Subsets	109
5.3	Case Study	110
5.4	Cross Comparison	113
5.5	Best Sampling Schedule	115
5.6	Clinical Impact	117
5.7	Summary	120
6	Applications to 24 h Data	121
6.1	Data	121
6.2	<i>SET1</i> versus Tk-GV	123
6.2.1	<i>GFR</i> Results	123
6.2.2	Curve Fit	124
6.2.3	4 h versus 24 h	127

6.3	Effect of Subsets	129
6.4	Exploring the <i>GV</i> Function	135
6.5	Summary	137
7	<i>GFR</i> Correction	138
7.1	Introduction	138
7.2	Data and Analysis Methods	140
7.3	Results and Discussion	141
7.4	Conclusion	145
8	Bone Mineral Density	146
8.1	Introduction	146
8.2	Literature Review	149
8.3	Motivation	152
8.4	Total Detectable Difference	153
9	Data and Analysis Methods	156
9.1	Data	156
9.2	Analysis Methods	157
9.2.1	Half-normal Distribution	158
9.2.2	Small Sample Correction	161
9.3	TDD Calculation	161
10	Results and Discussion	164
10.1	Results	165
10.2	Gender	166
10.3	Age	170
10.4	Examination Time	172

10.5 GE and Hologic	174
10.6 CaMos data	177
10.7 Impact on Diagnosis	179
10.8 Conclusion	180
11 Conclusions	182
11.1 Plasma Clearance	182
11.2 Bone Mineral Density	183
11.3 Future Directions	184
Appendix A	188
Bibliography	188
A	208
A.1 Clinical Protocol for Measuring <i>GFR</i>	208
A.2 Mathematical Formulae	210

List of Tables

2.1	The stages of chronic kidney disease (CKD). All GFR values are in units of ml/min/1.73 m ² (1.73 m ² represents the standard body surface area (BSA) of a healthy young adult).	32
4.1	GFR (ml/min) and Vol (L) values calculated from the $SET1$ model for the 46 patients using the full samples (denoted as “full”) and subsets: (60, 180)min and (10, 180)min.	75
4.2	The effects of the 1,148 2-sample subsets on the values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min) and \overline{Vol} (L) using the $SET1$ model. \overline{GFR} (ml/min) is the mean GFR , and SD_{GFR} (ml/min) is the standard deviation of the GFR values. \overline{Vol} (L) is the mean Vol , and SD_{Vol} (L) is the standard deviation of the Vol values.	78
4.3	The effects of the 3-sample subsets (2,296 for all 41 patients) on the values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min), and \overline{Vol} (L) using $SET1$. \overline{GFR} (ml/min) is the mean GFR and SD_{GFR} (ml/min) is the standard deviation of the GFR values. \overline{Vol} (L) is the mean Vol and SD_{Vol} (L) is the standard deviation of the Vol values.	79

4.4	The effects of the 4-sample subsets (2,870 for all 41 patients) on values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min) and \overline{Vol} (L) using <i>SET1</i> . \overline{GFR} (ml/min) is the mean <i>GFR</i> and SD_{GFR} (ml/min) is the standard deviation of the <i>GFR</i> values. \overline{Vol} (L) is the mean <i>Vol</i> and SD_{Vol} (L) is the standard deviation of the <i>Vol</i> values.	80
4.5	The <i>GFR</i> results of the <i>SET1</i> model for the 41 patients using the full 8 samples and four subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.	80
4.6	Estimated <i>GFR</i> (ml/min) and <i>Vol</i> (L) values of the <i>SET2</i> model for the 46 patients using the full samples and the (10, 20, 60, 180)min subset.	84
4.7	Estimated <i>GFR</i> (ml/min) calculated with the <i>SET2</i> model using the 4-sample subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min from the 41 patients. Q stands for quartile.	85
4.8	The effects of 4-sample subsets (2,870 subsets) using the <i>SET2</i> model on values of $\bar{\lambda}_1$ and $\bar{\lambda}_2$, \overline{GFR} (ml/min) and \overline{Vol} (L) from the 41 patients. SD_{GFR} (ml/min) is the standard deviation of the <i>GFR</i> values. SD_{Vol} (L) is the standard deviation of the <i>Vol</i> values.	86
4.9	Quartiles of the mean estimated parameters \bar{K} , $\bar{\alpha}$, and $\bar{\beta}$, and the values of \overline{GFR} (ml/min), and \overline{Vol} (L) for the OLS-GV model from the 41 patients using the 2,870 4-sample subsets. \overline{GFR} is the mean <i>GFR</i> and SD_{GFR} is the standard deviation of the <i>GFR</i> values. \overline{Vol} (L) is the mean <i>Vol</i> and SD_{Vol} (L) is the standard deviation of the <i>Vol</i> values.	89

4.10	The <i>GFR</i> and <i>Vol</i> results of the OLS-GV model for the 41 patients using the full 8 samples and four chosen subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.	89
4.11	The <i>GFR</i> and <i>Vol</i> results of the Tk-GV model for the 41 patients using the full 8 samples and four different subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.	92
4.12	Estimated values of $\bar{\lambda}$, $\overline{\ln K}$, $\bar{\alpha}$, $\bar{\beta}$, \overline{GFR} , \overline{Vol} and the standard deviations SD_{GFR} (ml/min) and SD_{Vol} (L) values using the 2,870 4-sample subsets from the 41 patients using the Tk-GV model.	93
5.1	The slope and R^2 values obtained from the correlation between the results from the <i>SET1</i> , <i>SET2</i> , OLS-GV and Tk-GV models using the 4-sample subsets (shown from Figures 5.1 to Figure 5.4) and the full samples from OLS linear regression.	107
5.2	Estimated <i>GFR</i> values for Pt15 and Pt19 using the 4-sample subsets (70 subsets) using the <i>SET1</i> , <i>SET2</i> , OLS-GV and Tk-GV models. The GFR_8 (in ml/min) and Vol_8 (in L) are the estimated results using 8 samples. The \overline{GFR}_4 and \overline{Vol}_4 are the mean estimated results of <i>GFR</i> and <i>Vol</i> using 4 samples.	112
6.1	Estimated <i>GFR</i> (ml/min) values from the 10 patients with 4 h, 12 h and 24 h sampling schedules, using the <i>SET1</i> and Tk-GV models.	123

6.2	The values of the slope and R^2 from the linear regression of the estimated GFR values obtained from the $SET1$, $SET2$ and Tk-GV models using 4 h, 6 h, 8 h and 12 h sampling schedules to those under the same models using the 24 h sampling schedule for the 10 patients.	134
9.1	The probability density function (PDF) of the folded-normal distribution (FND), half-normal distribution (HND), and normal distribution (ND), where μ is the mean value and σ is the standard deviation. . .	160
10.1	TDD results (g/cm^2) of approximately 5,500 patients analyzed using the GE Lunar Prodigy densitometer at the General Hospital using the HND and FND. The R^{2*} (see the definition below) value is for both the HND and FND. Area indicates the folded area of the FND from the ND.	165
10.2	The TDD (g/cm^2) values of the female and male groups, for HND and FND, using a GE Lunar Prodigy densitometer.	169
10.3	The TDD_{HND} results (g/cm^2) of four age groups from the 5,500 patients, using a GE Lunar Prodigy densitometer.	170
10.4	The TDD results (g/cm^2) using the HND for 10 examination time intervals from the 5,500 patient data, using a GE Lunar Prodigy densitometer.	172
10.5	The TDD values (g/cm^2) from the GE and Hologic densitometers in St.John's, NL.	175
10.6	TDD results (g/cm^2) from NL and the CaMos database using the GE Lunar densitometers, with number of patients in the parentheses. . .	177

10.7 TDD results (g/cm^2) from NL and the CaMos database using the Ho- logic QDR 4500 densitometers, with number of patients in the paren- theses.	177
--	-----

List of Figures

1.1	Schematic representation of the transition from ^{99}Mo to ^{99}Ru	7
1.2	The single-well gamma counter and thyroid uptake camera (manufactured by Laboratory Technologies, Inc.) used at the General Hospital in St.John's, Newfoundland and Labrador.	12
1.3	Diagram of a typical gamma camera. This diagram was adapted with permission from the Journal of RadioGraphics [23].	14
1.4	(a) Demonstration of a BMD image of the lumbar spine and (b) the hip, obtained using a GE Lunar Prodigy Dual-energy X-ray absorptiometry densitometer.	21
2.1	Schematic representation of the anatomy of a human kidney.	28
2.2	Schematic representation (lightly shaded area) of the Area Under the Curve (AUC) used for calculating the elimination rate of a tracer. The plasma concentration of the tracer is plotted as a function of the time (with time=0 set at the finishing point of the injection) after the administration. The three phases in the renograph are labeled as I, II and III.	43

2.3	Schematic representation of the one-compartment model used to model the plasma clearance. The circle represents the plasma as the compartment. The two arrows indicate the flow of the radiopharmaceuticals in and out of the compartment.	46
2.4	The two-compartment model used to describe plasma clearance. The two circles C_1 and C_2 represent compartments for the plasma and secondary volume of distribution of the tracer respectively. The flow in and out of the two compartments is indicated with arrows showing the rate of exchange K_{in} , K_{out} , K_{12} and K_{21}	49
4.1	Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the <i>SET1</i> model from the 41 patients. The hump subsets are <i>1 to 3</i> , <i>1 to 4</i> , . . . , <i>5 to 8</i> and <i>6 to 8</i>	82
4.2	Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the <i>SET2</i> model from the 41 patients. The hump subsets are <i>1 to 5</i> , <i>1 to 6</i> , <i>1 to 7</i> , <i>2 to 8</i> , <i>3 to 8</i> and <i>4 to 8</i>	87
4.3	Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the OLS-GV model from the 41 patients. The hump subsets are <i>1 to 5</i> , <i>1 to 6</i> , <i>1 to 7</i> , <i>2 to 8</i> , <i>3 to 8</i> and <i>4 to 8</i>	91
4.4	Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the Tk-GV model from the 41 patients. The hump subsets are <i>1 to 5</i> , <i>1 to 6</i> , <i>1 to 7</i> , <i>2 to 8</i> , <i>3 to 8</i> and <i>4 to 8</i>	95

5.1 Comparison of the GFR_{SET1} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$ 103

5.2 Comparison of the GFR_{SET2} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$ 104

5.3 Comparison of the GFR_{OLS-GV} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$ 105

5.4 Comparison of the GFR_{Tk-GV} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$ 106

5.5 Comparison of performance of the concentration time (with time=0 set at the finishing point of the injection) curve fitting using the $SET1$ (fine dashed line), $SET2$ (solid line), OLS-GV (dash-dot-dash line) and Tk-GV (dashed line) models for Pt1 using the full samples. The original 8 samples are represented with dark circles. 111

5.6	Comparison of the 46 estimated <i>GFR</i> values using <i>SET1</i> (dark circles) and <i>SET2</i> (gray rhombus) compared with Tk-GV, using the full samples. The dashed line is the identity line $Y = X$	114
5.7	Comparison of (a) $(GFR_{SET1} - GFR_{Tk-GV})/GFR_{Tk-GV}$ from <i>SET1</i> with Tk-GV and (b) $(GFR_{SET2} - GFR_{Tk-GV})/GFR_{Tk-GV}$ from <i>SET2</i> with the Tk-GV model for the 46 patients. The vertical dashed lines indicate the CKD stages from CKD1 to CKD5.	118
6.1	Logarithm of the concentration versus time (with time=0 set at the finishing point of the injection) using the <i>SET1</i> and Tk-GV models for Pt3. The original data is shown with hollow circles, the Tk-GV fit is indicated by a solid line while the <i>SET1</i> fit is indicated by a dashed line.	125
6.2	Logarithm of the concentration versus time (with time=0 set at the finishing point of the injection) using the <i>SET1</i> model for Pt3. The original data are shown as crosses (+). The straight lines indicate the fits obtained using the <i>SET1</i> model with different sampling times as shown in the legend. The insert on the bottom left shows how the <i>SET1</i> model overestimates the concentration data for earlier time.	126
6.3	(a) Estimated <i>GFR</i> results (ml/min) for the 10 patients with a 24 h sampling schedule using the Tk-GV model (black bars) and the <i>SET1</i> model (white bars) using the 4 h sampling schedule. (b) The relative difference in <i>GFR</i> between the two models $\frac{GFR_{SET1} - GFR_{Tk-GV}}{GFR_{SET1}}$	128

6.4	Correlation between the estimated GFR values using the $SET1$ model and the Tk-GV model from a) the 4 h and b) the 24 h samples. The dashed line indicates the identity line $Y = X$ while the solid line is the fit.	130
6.5	Comparison of the 10 estimated GFR values from the Tk-GV model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend. . . .	131
6.6	Comparison of the 10 estimated GFR values from the $SET1$ model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend. . . .	132
6.7	Comparison of the 10 estimated GFR values from the $SET2$ model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend. . . .	133
6.8	Percentage of the power function (hollow circles) and exponential function (solid circles) in the whole GV model as a function of sampling time t (with time=0 set at the finishing point of the injection) in min for a) Pt3 and b) Pt13.	136
7.1	a) Estimated GFR_{SET1} results using the 2-sample $SET1$ model (ml/min). b) Estimated results of $GFR_{Haycock}$ (ml/min/1.73 m ²) (dark bars) and $RSI_{Haycock}$ (white bars). The solid line indicates the Haycock criterion. c) Estimated results of $GFR_{f(Vol,W)}$ (ml/min) (dark bars) and $RSI_{f(Vol,W)}$ (white bars). b) and c) The dashed line represent the 85.89% RSI threshold.	142

7.2	Correlation of RSI_{Haycock} and $RSI_{f(Vol,W)}$ results using the linear regression. The solid line indicates the best fit using a second order polynomial from the 26 patients, and the dashed line indicates the fit from the initial study by Wesolowski et al. based on study of 133 children.	144
8.1	Demonstration of the composition of total BMD changes including short-term and long-term machine errors, as well as patients' BMD variations. The short-term and long-term machine errors were estimated from local study while the short-term and long-term patients' BMD variations were estimated from the reference [155].	155
9.1	Demonstration of the generation of the folded-normal distribution (FND) from the normal distribution (ND).	159
9.2	A flow chart indicating how the TDD value is calculated from the initial patient's dataset.	162
10.1	The HND fitting for patients with three sequential BMD studies at the examination site of (a) the lumbar spine L1-L4 and (b) the lumbar spine L2-L4, using a GE densitometer.	167
10.2	The HND fitting for patients with three sequential BMD studies at the examination site of (a) the femoral neck and (b) the total hip, using a GE densitometer.	168
10.3	The TDD values (g/cm^2) for the four examination sites of the four age groups (<50 , $(50, 59)$, $(60, 69)$ and $70<$). The color scheme indicating each examination site is shown in the legend.	171

10.4	The TDD values from the 10 time intervals for the four examination sites: lumbar spine L2-L4, lumbar spine L1-L4, total hip and femoral neck. The color scheme indicating each examination site is shown in the legend.	173
10.5	a) Comparison of the TDD values (g/cm^2) from the GE (white bar) and Hologic QDR 4500 densitometer (gray bar) for the lumbar spine L2-L4, L1-L4, femoral neck and total hip. b) Relative difference $(TDD_{\text{GE}} - TDD_{\text{Hologic}})/TDD_{\text{GE}}$ of the TDD values between the GE and Hologic densitometer.	176
10.6	Percentages of the detectable BMD differences using the LSC (dark bars) and TDD (white bars) from the local GE Lunar densitometer.	179

Abbreviations

- **A** Age
- **AUC** Area Under the Curve
- **BMD** Bone Mineral Density
- **BSA** Body Surface Area
- ***CL*** Clearance
- **CT** Computed Tomography
- **CV** Coefficient of Variance
- **DEXA** Dual Energy X-ray Absorptiometry (DXA)
- **DMSA** Dimercaptosuccinic Acid
- **DTPA** Diethylene Triamine Pentaacetic Acid
- ***ERG*** Relative Error of the Estimated GFR
- ***ERV*** Relative Error of the Estimated Vol
- **EDTA** Ethylenediaminetetraacetic Acid
- **FND** Folded-normal Distribution

- ***GFR*** Glomerular Filtration Rate
- ***GV*** Gamma Variate
- ***H*** Height
- ***HND*** Half-normal Distribution
- ***LSC*** Least Significant Change
- ***MRI*** Magnetic Resonance Imaging
- ***NM*** Nuclear Medicine
- ***ND*** Normal Distribution
- ***OLS*** Ordinary Least Square
- ***PE*** Precision Error
- ***Q*** Quartile
- ***QCT*** Quantitative Computed Tomography
- ***ROI*** Region of Interest
- **R^2** Coefficient of Determination
- **R^{2*}** Adjusted Coefficient of Determination
- ***RSI*** Renal Sufficient Index
- ***SD*** Standard Deviation
- ***SE*** Standard Error

- *SET1* Sum of 1 Exponential Term
- *SET2* Sum of 2 Exponential Terms
- *SETn* Sum of n Exponential Terms
- *SD* Standard Deviation
- *Tc* Technetium
- *TDD* Total Detectable Difference
- *Tk* Tikhonov (Regularization)
- *US* Ultrasonography
- *Vol* Volume of Distribution
- *W* Weight

Chapter 1

Introduction

This chapter presents a brief introduction of Nuclear Medicine, basic nuclear physics, radiopharmaceuticals, and radiation detectors.

1.1 Nuclear Medicine

Nuclear medicine is a branch of medicine and medical imaging that uses the nuclear properties of matter in diagnosis and therapy. Nuclear medicine is unique (different from other medical imaging modalities) in that it provides both anatomical and functional information. Clinical information can be derived from observing the radiation emitted from radiopharmaceuticals (or radionuclide alone) administered orally, or intravenously. Radiopharmaceuticals are radioactive pharmaceuticals made up of a radionuclide and a carrier molecule. The carrier molecule delivers the radionuclide to the specific physiological area to be examined or treated [1]. Measurements in nuclear medicine can be performed either *in vitro* or *in vivo*. *In vitro* measurements are based on samples (e.g. blood, urine) taken from patients after administering the radiopharmaceuticals. The radioactivity in these samples can help determine the

physiological functions of the body or organs. *In vivo* measurements are performed on patients directly with an external detector (e.g. gamma camera) for measuring the radiation emitted from the radiopharmaceuticals inside the body. In general, *in vivo* measurements are more commonly used than *in vitro* measurements.

Although natural radioisotopes were discovered before 1900 [2], the first nuclear medicine study was not performed until the late 1940s using radioactive iodine for thyroid cancer treatment. With the development of pharmacology and biochemistry, more radiopharmaceuticals were designed and used for detecting various diseases. By the 1970s, most organs could be visualized using nuclear medicine techniques. Concurrently, radiation detection technology improved the quality of the nuclear medicine images. The first gamma camera for recording radiation counts and producing planar images was invented by Hal Anger in the 1950s [3]. In 1963, the first Single Photon Emission Computed Tomography (SPECT) study was performed by Kuhl and Edwards [2]. SPECT is a tomographic technique of nuclear medicine which uses gamma rays to scan the body at different angles and combines the acquired two-dimensional (2D) images to reconstruct a three-dimensional (3D) representation of the body [4]. By the 1990s, Positron Emission Tomography (PET) was invented [5]. PET has an advantage over other imaging modalities such as Magnetic Resonance Imaging (MRI) [6] or ultrasonography [7] in that it can detect metabolic abnormalities. Recent developments in imaging techniques consist of the fusion of several existing imaging techniques such as SPECT/CT, PET/CT, PET/MRI, etc.

1.2 Radioisotopes

A radioisotope is an atom with an unstable nucleus that can release energy when it undergoes radioactive decay. Commonly used radioisotopes in nuclear medicine

are fluorine 18 (^{18}F)[†], Phosphorus 32 (^{32}P), Gallium 67 (^{67}Ga), Rubidium 82 (^{82}Rb), Strontium 89 (^{89}Sr), Technetium 99 (^{99}Tc), Indium 111 (^{111}In), Iodine 123 (^{123}I), Iodine 131 (^{131}I), Thallium 201 (^{201}Tl), etc [8]. Besides their use in medical diagnosis, radioisotopes are also used for treating certain diseases by delivering high radiation dosage to regions of interest. For example, hyperthyroidism and thyroid cancer can be treated using ^{131}I , and palliative bone pain can be managed using ^{89}Sr [9].

1.2.1 Decay and Half-life

Most radioisotopes decay by one or several of the following ways: 1) alpha decay, 2) beta-minus decay, 3) beta-plus decay, or 4) gamma decay. The following list shows some examples of these decay processes [10]:

$$\alpha \text{ decay: } \quad {}^A_Z X \rightarrow {}^{A-4}_{Z-2} Y + {}^4_2 \text{He}$$

$$\beta^- \text{ decay: } \quad n \rightarrow p + e^- + \bar{\nu}_e$$

$$\beta^+ \text{ decay: } \quad \text{energy} + p \rightarrow n + e^+ + \nu_e$$

$$\gamma \text{ decay: } \quad {}^A_Z X \rightarrow {}^A_Z X + \gamma.$$

Alpha (α) decay occurs when an atomic nucleus (e.g. $^{238}_{92}\text{U}$) emits an alpha (α) particle (a ${}^4_2\text{He}$ helium nucleus) yielding another nucleus (e.g. $^{234}_{90}\text{Th}$) in the process. Because of their mass, alpha particles lose most of their energy within a relatively short distance (short penetration depth) making it difficult to detect them with an external detection camera [10]. Therefore, alpha decay is not commonly used in nuclear medicine.

There are two types of beta (β) decay: beta-minus decay and beta-plus decay [10]. Beta-minus (β^-) decay occurs with the conversion of a neutron (n) into a

[†]The full notation is ${}^{18}_9\text{F}$ where 18 is the mass number and 9 is the atomic number.

proton (p) along with the ejection of a negatively charged beta particle (e^- , electron) and an antineutrino ($\bar{\nu}_e$). On the other hand, beta-plus (β^+) decay results in the conversion of a proton (p) into a neutron (n) accompanied by the ejection of a positron (e^+) as well as an electron neutrino (ν_e). There is a relevant decay mode called electron capture, also known as inverse beta decay, which occurs when the nucleus captures an orbital (e.g. K -shell or L -shell) electron, with the conversion of a proton (p) into a neutron (n) accompanied by the ejection of a neutrino (ν_e) described as $p + e^- \rightarrow n + \nu_e$.

Gamma (γ) decay occurs when an atomic nucleus decays to a lower energy state from an excited state. The energy difference between the two states is released by the emission of a gamma ray photon (γ) [10]. Gamma rays are very important in nuclear medicine because of their relatively long penetration range that allows information (e.g. location and amount) from the administered radiopharmaceuticals to be collected.

Other useful transition processes include: isomeric transition where the decay process yields gamma radiation without emitting or capturing a particle from the nucleus, and electron/positron annihilation (the basis of PET scanning) where an electron (e^-) and a positron (e^+) collide (annihilate) resulting in a photon pair (gamma rays, γ) $e^- + e^+ \rightarrow \gamma + \gamma$, etc.

Radioactive decay is a random process where the probability of a single event (such as whether a specific atom will decay or not) cannot be determined precisely. However, probability theory enables the radioactive decay process to be described using the decay equation:

$$A(t) = A_0 e^{-\lambda t}, \quad (1.1)$$

where $A(t)$ is the radioactivity at time t ($t \geq 0$), A_0 is the initial radioactivity at

time $t = 0$, and λ is the decay constant given by $\lambda = (\ln 2)/T_{1/2} = 0.693/T_{1/2}$ where $T_{1/2}$ is the half-life. Half-life is defined as the time for a radioactive sample to decay to half of its initial radioactivity [1]. The half-lives of the radiopharmaceuticals used in nuclear medicine range from seconds to years. Those radiopharmaceuticals with relatively short half-lives ensure that medical scanning can be performed in a timely fashion and that patients are not exposed to prolonged radiation.

1.2.2 Interactions with Matter

Gamma radiation ionizes matter through three major processes: the photoelectric effect, Compton scattering, and pair production [10].

The photoelectric effect transfers an incident photon's energy to an electron causing it to be ejected from the atom as a photoelectron. The photoelectron's kinetic energy is the difference between the incident photon's energy $h\nu$ (where h is the Planck constant and ν is the frequency of the photon) and the electron's binding energy. This photoelectron usually travels in a different direction from the incident gamma photon [10].

Compton scattering occurs when an incident photon (with energy $h\nu$) interacts with matter, causing the ejection of an orbital electron (known as a Compton electron) and the scattering of the photon at a reduced energy $h\nu'$. The difference between the photoelectric effect and Compton scattering is that in the latter case the electron is ejected via a scattering process while in the former case the electron is emitted after absorbing energy from the gamma photon.

Pair production can occur when the energy of a gamma photon exceeds 1.022 MeV [10]. A gamma photon (with energy of $h\nu$) can be converted into an electron-positron pair (e^+ and e^-) by interacting with a nucleus. Since each of the particles

has a rest mass energy of 0.511 MeV, the remaining gamma photon energy (greater than 1.022 MeV) appears as the kinetic energy of the electron-positron pair. Other pairs can also be produced including a tau and anti-tau or muon and anti-muon pair.

When photons pass through matter, the gradual loss in the intensity of the photons is called attenuation. Attenuation reduces the intensity of the incident beam due to absorption or scattering. The intensity, I , of a beam of photons going through an absorber is given by:

$$I = I_0 e^{-\mu_l l}, \quad (1.2)$$

where I_0 is the original intensity of the beam, l is the distance traveled in the absorber material and μ_l is the attenuation coefficient (also called the linear attenuation coefficient) [10]. The minus sign ($-$) in the argument of the exponential indicates that the intensity decreases with increasing distance l . The quantity μ_l can be expressed as $\mu_l = \mu_m \rho$ where ρ is the density of the absorber and μ_m is the mass attenuation coefficient of the absorber. μ_m depends on the atomic number of the absorber and the photon energy. The mass attenuation coefficient μ_m includes three parameters as follows $\mu_m = \tau + \sigma + \kappa$ where τ is the part due to the photoelectric effect, σ is the part due to Compton scattering and κ is the part due to pair production.

1.3 Radiopharmaceuticals

Our study of kidney function (Chapters 2 to 7) involved the use of the metastable (m) isotope technetium 99 (^{99m}Tc) as a tracer for measuring plasma clearance. ^{99m}Tc is the metastable isomer of ^{99}Tc . Technetium 99 has an atomic number of 43 and a mass number of 99. Its half-life is approximately 2.12×10^5 years. On the other hand, ^{99m}Tc has a half-life of approximately 6.02 hours which is suitable for medical

imaging and blood sample collection techniques. Moreover, the relatively short half-life of ^{99m}Tc limits the radiation exposure to patients. ^{99m}Tc has many advantages for kidney function tests such as: 1) Its half-life (6.02 hours) is comparable to the duration of many diagnostic studies; 2) It emits gamma rays within the detectable range of diagnostic equipment; 3) It is chemically suitable for chelating with certain pharmaceuticals [11]. In clinical nuclear medicine applications, ^{99m}Tc -labeled compounds occupy approximately 85% of all radiopharmaceutical usage.

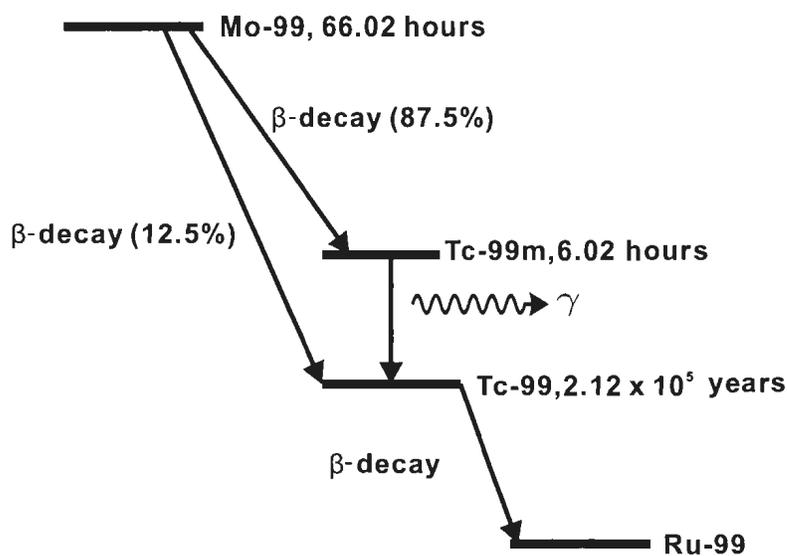


Figure 1.1: Schematic representation of the transition from ^{99}Mo to ^{99}Ru .

Molybdenum 99 (^{99}Mo) is a radiopharmaceutical with a half-life of 66.02 hours. The beta-minus decay from ^{99}Mo to ^{99}Tc has a probability of 12.5% as shown in Figure 1.1. The second beta-minus decay (with a probability of 87.5%) results in the formation of ^{99m}Tc through $^{99}\text{Mo} \rightarrow ^{99m}\text{Tc} + e^- + \bar{\nu}_e$ where $\bar{\nu}_e$ is the emitted antineutrino. ^{99m}Tc undergoes an isomeric transition to ^{99}Tc through the emission of a gamma photon. It is the gamma photon from this transition that is detected

during measurements.

A generator is a system which holds a parent/daughter (herein $^{99}\text{Mo}/^{99m}\text{Tc}$) mixture in a transient equilibrium. The generator is used for transporting radioisotopes whose parents' half-lives are relatively long compared with the transportation time. Because of their relatively short half-lives, certain radioisotopes cannot be shipped by commercial carriers and hence must be made locally or on-site. ^{99}Mo can easily be stored and transported between medical institutions.

A ^{99}Mo to ^{99m}Tc generator is based on an ion-exchange column [1]. The working process of a generator can be described by the following. Firstly, ^{99}Mo exists as ^{99}Mo sodium molybdate (Na_2MoO_4) and is initially bounded to an alumina (Al_2O_3) column. When ^{99}Mo in the molybdate ion decays, it is transformed into ^{99m}Tc as a pertechnetate ion $^{99m}\text{TcO}_4^-$. Passing a saline eluant solution through the alumina column removes the $^{99m}\text{TcO}_4^-$ and leaves $^{99}\text{MoO}_4^{2-}$ on the alumina column. The $^{99m}\text{TcO}_4^-$ reacts with the saline solution forming sodium pertechnetate ($\text{Na}^{99m}\text{TcO}_4$). After the ^{99m}Tc isotope is extracted from the sodium pertechnetate in one elution, the ^{99m}Tc 's activity in the generator will regenerate from zero. The total amount of ^{99m}Tc extracted depends on the time interval between the elutions, the quantity of ^{99}Mo and the efficiency of the elution.

There are several types of chemical impurities produced by a $^{99}\text{Mo}/^{99m}\text{Tc}$ generator eluate such as: 1) ^{99}Mo impurities contained within the product ^{99m}Tc , 2) the radiochemical impurity introduced by hydrolyzed technetium, and 3) the chemical impurity of free Al^{3+} ions [1]. Therefore, quality control is required.

1.3.1 Chelation

Chelation is the formation of chemical bonds between a ligand and a single atom [12]. The ligands used in radiopharmaceuticals are usually organic compounds. Chelants react with metal atoms therefore preventing the latter from reacting with other elements. Chelants are used in chelation therapy for heavy metal detoxification.

One of the most commonly used chelants of ^{99m}Tc is diethylenetriamine pentaacetate (DTPA). ^{99m}Tc -DTPA is used for kidney function tests because it is mostly eliminated from the kidneys after an intravenous administration [13]. ^{99m}Tc -DTPA was the radiopharmaceutical used in this study. Other commonly used chelants for ^{99m}Tc for performing kidney function tests are Mercapto Acetyl Tri Glycine (MAG3) and Dimercaptosuccinic acid (DMSA) [14]. ^{99m}Tc -MAG3 is used to detect scarring or necrosis of the renal cortex, as well as pyelonephritis (an infection in the kidneys) [15]. ^{99m}Tc -DMSA is usually used for renal cortical imaging.

Besides ^{99m}Tc , there are several other radioisotopes used for kidney function tests. One of the alternative radiopharmaceuticals is ^{51}Cr -ethylenediaminetetraacetic acid (^{51}Cr -EDTA) which is freely filtrated by the glomerulus (see definition in Chapter 2) in the kidneys [16]. However, ^{51}Cr -EDTA is less often used than ^{99m}Tc -DTPA in Canada. Another radiopharmaceutical is ^{123}I ortho-iodohippurate (^{123}I -OIH) which is cleared by tubular secretion in the kidneys.

One thing to be considered while using ^{99m}Tc -labeled radiopharmaceuticals is the possibility of protein binding of the radiopharmaceuticals with the plasma after the administration. Protein binding affects kidney function test results because it is the unbounded fraction of ^{99m}Tc -DTPA in the body that produces the pharmacological effect. The normal range of protein-binding of the ^{99m}Tc -DTPA is between 3.7% and 13.5% [17]. Dual injections (inject the same tracer twice over a certain time period)

of ^{99m}Tc -DTPA introduces a larger protein binding rate compared to using a single injection [18]. To reduce the inaccuracy of the kidney function test due to protein binding, one can use an ultrafiltrated solution of plasma (centrifuge the plasma for an additional 10 to 20 minutes) rather than a standard filtrated plasma (centrifuge the blood for 10 to 15 minutes until the plasma is separated from the blood) and use chromatography (a laboratory technique for separating mixtures) to determine the binding fraction [12].

1.4 Gamma Counter

The conventional instrument for counting gamma rays is the Geiger Counter [19]. The Geiger counter is composed of a tube filled with an inert gas and two electrodes (anode and cathode). When radiation enters the tube, the gas is ionized resulting in the formation of positively charged ions (moving towards the cathode) and electrons (moving towards the anode). During this process, additional ion pairs are also generated. The net effect is the generation of a measurable electrical current. A more precise device for counting radioactivity is the scintillation detector.

Scintillation is a physical phenomenon whereby certain materials can emit light (luminescence) when struck by radiation [10]. Radiation emitted from the body enters the device through a window and travels to the scintillator crystals. In the scintillator crystal, the incident radiation produces photoelectrons which move around the crystal exciting orbital electrons. The excitation of these orbital electrons results in orbital vacancies, which are then filled by higher energy electrons. The excess energy of the decaying electrons is released as photons. Some of the photons strike the photocathode and produces electrons. This process is repeated in a photomultiplier tube (PMT) where the signals are magnified to a measurable electric signal.

The inorganic scintillator, thallium-activated sodium iodine (NaI(Tl)), is the most widely used scintillator in nuclear medicine. NaI(Tl) has several advantages such as: 1) Its dense structure ($\rho = 3.67 \text{ g/cm}^3$) which makes it a good absorber, 2) NaI is transparent to light, 3) The output signal is approximately linear over a wide range of energy (for a relatively thick crystal, e.g. 25 mm) which is suitable for counting, and 4) It has a high light output because its emission spectrum matches well with the sensitivity of photomultiplier tubes. A maximum efficiency ($>90\%$) can be achieved for incident wavelengths between 400 and 450 nanometers (nm); and 5) The cost is relatively low [1].

The NaI(Tl) counters used in this work were commercial well counters where the radioactive materials were placed inside a container consisting of wells of various shape such as 1.6 cm diameter \times 3.8 cm deep or 13 cm diameter \times 25 cm deep. NaI(Tl) counters are very suitable for counting samples with relatively small radioactivity (on the order of microcuries) due to their high geometric efficiency.

Figure 1.2 shows the single-well gamma counter used in the department of nuclear medicine at the General Hospital in St. John's, NL. The cylinder at the bottom is the single well, surrounded by a shielding lid. That is where the test sample, plasma containing a radiopharmaceutical, was located when measuring the activity. The single-well gamma counter is connected to a computer for displaying and analyzing the data. The accuracy of the single-well counter can be affected by factors such as dead time correction, detector efficiency or geometric efficiency [1, 20]. A routine quality control for ensuring the performance of the single-well counter should include steps such as daily tests for the calibration of peaking and sensitivity, monthly test using the Chi-square tests, occasionally (quarterly) test for the energy resolution and an annual test for the detector efficiency [21]. There is also an adjustable probe for

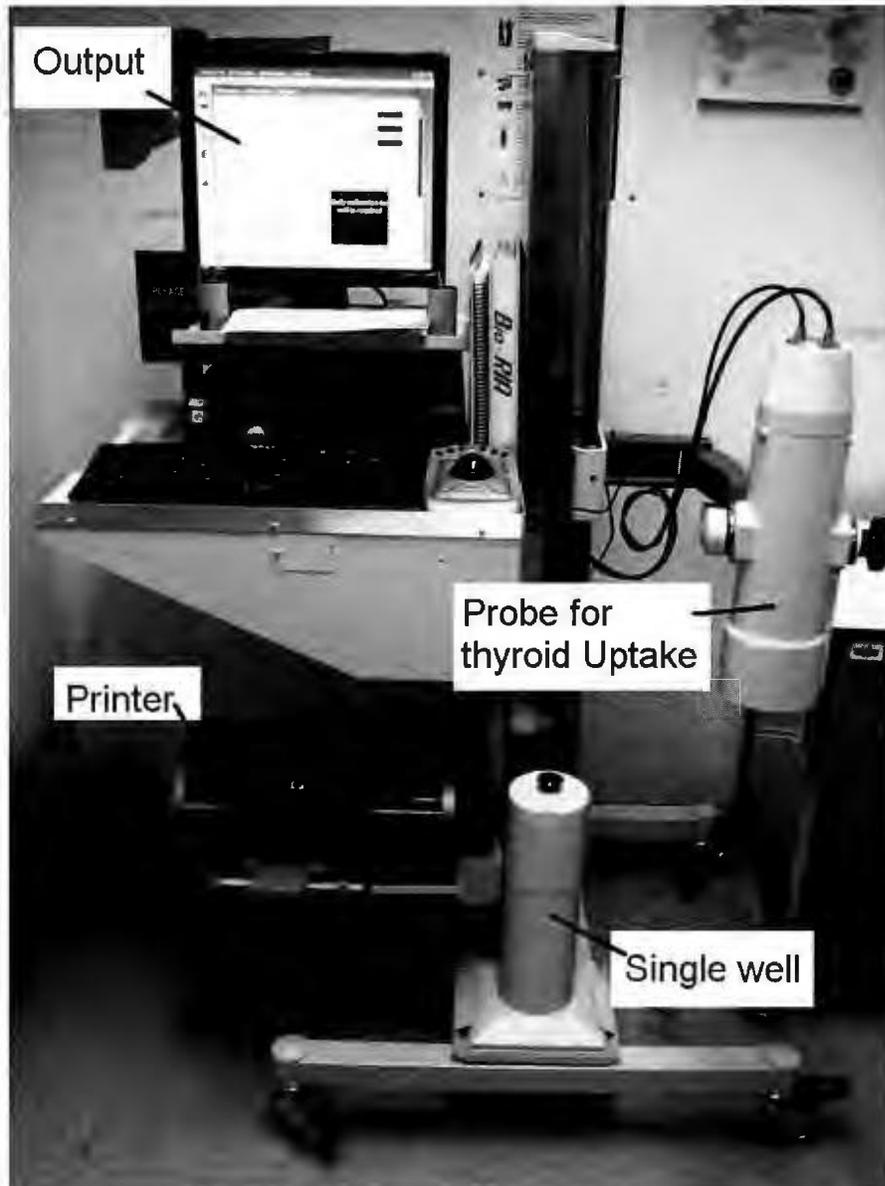


Figure 1.2: The single-well gamma counter and thyroid uptake camera (manufactured by Laboratory Technologies, Inc.) used at the General Hospital in St. John's, Newfoundland and Labrador.

performing a thyroid uptake test using ^{123}I or ^{131}I . To perform such measurements, the patient sits in a chair facing the stationary probe positioned over their thyroid gland in the neck.

1.5 Theory of SPECT

The vast majority of nuclear medicine investigations are based on the usage of SPECT imaging. SPECT imaging is used to provide a three-dimensional (3D) image of regions of interest in the body. In order to construct the image the patient is injected with a solution containing a radioactive tracer that moves throughout the patient through the bloodstream. Gamma cameras are used to measure the gamma rays emitted from the tracer to form two-dimensional (2D) images. Using a computer assisted tomographic technique, the two-dimensional images are combined to form a three-dimensional representation of the regions of interest under study. SPECT imaging is commonly used for performing kidney function tests. By providing dynamic images during a live scan, SPECT images can show how the kidneys are clearing the tracer out in real time. At the heart of a SPECT machine are gamma cameras. The most commonly used gamma cameras are scintillation cameras which were first developed by Hal O. Anger in 1950s [22].

Figure 1.3 shows a schematic representation of a gamma camera and how it is used to convert captured gamma rays into an image [23]. Gamma rays emitted from the radiopharmaceutical in the patient exit the body (labeled as subject shown in Figure 1.3) and reach the collimators. The collimators which are usually made of tungsten or lead are used for refining the direction of the gamma rays by blocking or absorbing gamma rays which arrive at an unwanted angle. Once the photons pass the collimators, they strike the NaI(Tl) scintillation crystal which converts the

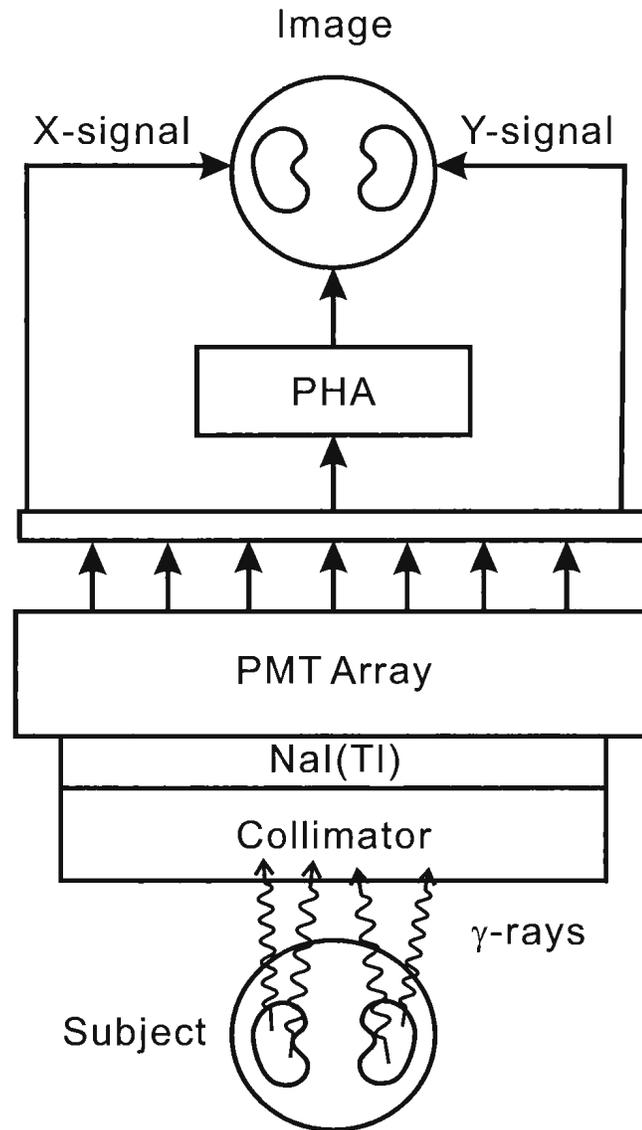


Figure 1.3: Diagram of a typical gamma camera. This diagram was adapted with permission from the Journal of RadioGraphics [23].

energy of an incident gamma-ray photon to a lower energy photon (light) usually in the visual range. The light enters a photomultiplier tubes (PMTs) which converts the incident photon into an electron via the photoelectric effect which is then amplified into a measurable signal. This signal is split to produce information regarding the X and Y positions of the initial gamma ray along with a signal evaluated by a pulse height analyzer (PHA) which selectively counts pulses from the PMTs that fall within certain voltage amplitude intervals, known as channels. The combined information is used to produce a 2D image as shown on top of Figure 1.3. The 2D images can be acquired either by a single-head rotating gamma camera (such as an arc with 180 degree rotation or a multi-head gamma camera where multiple gamma cameras rotate at the same time capturing gamma rays). The multi-head SPECT is more commonly used at the present because it takes less examination time and less radiation dose.

The 2D image, also known as a planar image, has no information on the depth and structure at different depths. In order to obtain a 3D image using a gamma camera, multiple 2D projections are required taken from different directions, which is the principle behind SPECT. Using sophisticated tomographic reconstruction algorithms such as an analytical algorithm or iterative algorithm [24], 2D images taken from multiple angles are then reconstructed to 3D images of the patient. Image reconstruction methods used in SPECT also help improve the quality of the 3D images by reducing the effect of noise factors.

When analyzing the quality of the images, there are, in general, three aspects to be considered. The first one is the unsharpness[†], seen as blurring or fuzziness. Factors such as the geometric distance between the detector and subject, the subject's indistinguishable boundaries, involuntary or voluntary motions and receptor limitations

[†]Unsharpness is a term used in medical imaging processing, which refers to the loss of edge detail of the geometric properties of the object or image.

in displaying the images, can contribute to the unsharpness. The second aspect is the contrast of an image which allows subtle features to be distinguished. The distinction can be a difference in physical density or chemical composition (atomic number). To improve this, some substances can be used as contrast agents to enhance the intrinsic contrast. This technique is usually used in x-ray imaging, medical ultrasonography and MRI imaging. The third aspect is image noise. Any irrelevant information in the image can be defined as noise. These could be structure noise (e.g. unimportant structures), radiation noise (e.g. nonuniform intensity of the beam or scattered radiation), receptor noise (e.g. unbalance detectors in the gamma camera) and quantum noise (e.g. fluctuations of electric power supply or the thermal noise due to Brownian motion[†])

1.6 Tracer Kinetic Modeling

Nuclear medicine diagnosis based on gamma counter measurements can be performed by pharmacokinetic analysis. One of the most important theories in pharmacokinetics is tracer kinetic modeling.

The object of the tracer kinetic model is called the tracer (herein a radioisotope or radiopharmaceutical) which is the substance that follows a physiological or biochemical process [1]. The kinetic process of how the body handles the tracer can be mathematically described using parameters such as: volume of distribution Vol (e.g. in ml or liter), transit time T (e.g. in min or sec), and clearance CL (e.g. in ml/min).

Vol is calculated at transient equilibrium by $Vol = D/C$ where D is the amount of tracer administered (e.g. in mg or Bq (for describing radioactivity)) and C is the

[†]Brownian motion is the random movement of particles suspended in a medium.

concentration of the tracer at transient equilibrium (e.g. in mg/ml or Bq/ml) [25]. Vol indicates the volume of the fluid for distributing the tracer in a transient equilibrium state between the plasma and the rest of the body.

The time T can be obtained from the equation $T = Vol/Q$ where Q is the flow of the tracer in units of ml/min [25]. When a measurement is taken over a period of time, the average T and average Q are used instead.

Clearance, CL , (or plasma clearance) is defined as the volume of a substance (tracer) cleared out from the body per unit time [25]. The definition is expressed as

$$\text{plasma clearance} = \frac{\text{amount of tracer injected}}{\text{total area under the concentration-time curve}}. \quad (1.3)$$

The unit of the injected Dose is usually in mg or Bq, and the unit for the area under the curve is usually in mg/min or Bq/min. Therefore, the calculated plasma clearance is in units of Bq/min or mg/min. In this study, plasma clearance is measured as the glomerular filtration rate. However, the tracer can be cleared from urine as well. The difference between the plasma and urine clearance is due to the redistribution of the tracer inside the body. In the following, the term plasma clearance is used merely for the renal clearance. More details of plasma clearance and the mathematical models used to estimate it are shown in Chapter 2.

1.7 Bone Densitometer

This PhD research project includes two projects. The first project is based on data collected using gamma counters while the second project is based on data collected using bone densitometers.

A bone densitometer is used to measure the bone mineral density (BMD) of pa-

tients. Bone density mainly refers to the inorganic components formed of carbonate hydroxyapatite [26]. BMD is a medical term used to define the bone mineral content (BMC) per area (being measured) given by $BMD = BMC / Area$, in units of g/cm^2 . The reason why BMD is not defined in terms of mass per volume is that the clinical measurement of BMD is mostly conducted using an X-ray technique which is based on projected 2D images. There are techniques such as quantitative computed tomography (QCT) for obtaining 3D information; however, QCT was not used in this work. Although QCT is more accurate for measuring volumetric bone density, it requires a relatively higher radiation dose than conventional CT imaging [27]. In this study, BMD was measured using absorptiometry.

Absorptiometry BMD measurements are based on the fact that cortical bone (also known as compact bone) has a higher linear attenuation coefficient than soft tissue under the same X-ray energy. The linear attenuation coefficient is the mass attenuation coefficient multiplied by the density of the material. For a given amount of incident X-ray radiation, bone absorbs more X-rays compared with soft tissue. Therefore, there will be less X-ray energy detected through the bone compared with regions of soft tissue. In single energy X-ray absorptiometry the attenuation equation of an incident X-ray beam going through an area containing both bone and soft tissue is written as:

$$I = I_0 e^{-(\mu_{\text{soft}} M_{\text{soft}} + \mu_{\text{bone}} M_{\text{bone}})}, \quad (1.4)$$

where I_0 is the intensity of the incident X-ray, I is the intensity of the X-ray going through the material, μ_{soft} and μ_{bone} are the mass attenuation coefficients (e.g. in units of g/cm^2) of soft tissue and bone and M_{soft} , and M_{bone} are the corresponding area densities (e.g. in units of cm^2/g). There are two unknown variables M_{soft} and M_{bone} in Eq 1.4. The solution to M_{bone} is obtained by surrounding the area being

measured with a soft tissue equivalent material (such as water), and measuring the beam intensity, I_{soft} , by letting the X-rays go through the area containing no bone material. In this way, the term $\mu_{\text{soft}}M_{\text{soft}}$ can be calculated as

$$\begin{aligned} I_{\text{soft}} &= I_0 e^{-\mu_{\text{soft}}M_{\text{soft}}} \\ \mu_{\text{soft}}M_{\text{soft}} &= -\ln \frac{I_{\text{soft}}}{I_0}, \end{aligned} \quad (1.5)$$

from which it is possible to calculate the bone density as

$$M_{\text{bone}} = -\frac{1}{\mu_{\text{bone}}} \left(\ln \frac{I}{I_0} + \mu_{\text{soft}}M_{\text{soft}} \right). \quad (1.6)$$

Because of the requirement of surrounding the area under study in water, single-energy X-ray absorptiometry is limited to the usage of certain body parts, such as the forearm [28].

At the present, dual energy X-ray absorptiometry (DEXA or DXA) developed from single energy X-ray absorptiometry (SXA) is more commonly used to distinguish the density of soft tissue from that of bone. The two X-ray energy levels can be generated by either using a K-absorption filter to separate a continuous energy X-ray spectrum into a high and low energy beam, or using a switch to change the voltage applied to the X-ray tube [29]. Given an incident X-ray beam with two energy levels, the measured intensities I_0 and I'_0 are given by:

$$\begin{aligned} I &= I_0 e^{-(\mu_{\text{soft}}M_{\text{soft}} + \mu_{\text{bone}}M_{\text{bone}})} \\ I' &= I'_0 e^{-(\mu'_{\text{soft}}M_{\text{soft}} + \mu'_{\text{bone}}M_{\text{bone}})} \end{aligned} \quad (1.7)$$

where the prime notation is used to indicate parameters that are affected by the

different X-ray energy. Equations 1.7 can be manipulated to solved for the density of both soft tissue and bone to give:

$$\begin{aligned} M_{\text{bone}} &= \frac{\mu'_{\text{soft}} \ln \frac{I_0}{I} - \mu_{\text{soft}} \ln \frac{I'_0}{I'}}{\mu'_{\text{soft}} \mu_{\text{bone}} - \mu_{\text{soft}} \mu'_{\text{bone}}} \\ M_{\text{soft}} &= \frac{\mu'_{\text{bone}} \ln \frac{I_0}{I} - \mu_{\text{bone}} \ln \frac{I'_0}{I'}}{\mu'_{\text{bone}} \mu_{\text{soft}} - \mu_{\text{bone}} \mu'_{\text{soft}}}. \end{aligned} \quad (1.8)$$

DXA not only improves the accuracy of the calculated bone density but also allows more sites such as spine and hip to be measured since it does not require the measurement site to be surrounded by water (as required by SXA) [29]. This is an important advantage of DXA since most BMD measurements focus on the lumbar spine or the upper part of the hip because of the high incident rate of fractures occurring around these areas. The radiation of a whole body bone density scan using a bone densitometer is approximately 1/10 that of a chest X-ray [30].

Figure 1.4 illustrates two BMD images obtained from a General Electric (GE) Lunar Prodigy bone densitometer machine in the department of Nuclear Medicine at the General Hospital in St. John's, NL. The image in Figure 1.4 (a) is a BMD scan of the lumbar spine known as the third major region of the spine. The sections L1-L4 labeled with grids in Figure 1.4 (a) were used in the second project described in this thesis (Chapters 8 to 10). The image in Figure 1.4 (b) is of the hip region. The region labeled femoral neck is the proximal part of the femur which surrounds the hip joint. A femoral fracture can be identified as a hip fracture if the fracture involves the femoral neck, head or the shaft of the femur. The BMD value of the femoral neck is important for the diagnosis or prediction of hip fracture. The edges of the spine and hip can be drawn either automatically using software or manually for defining the Regions of Interest (ROIs) for performing the BMD calculations. The brighter areas in

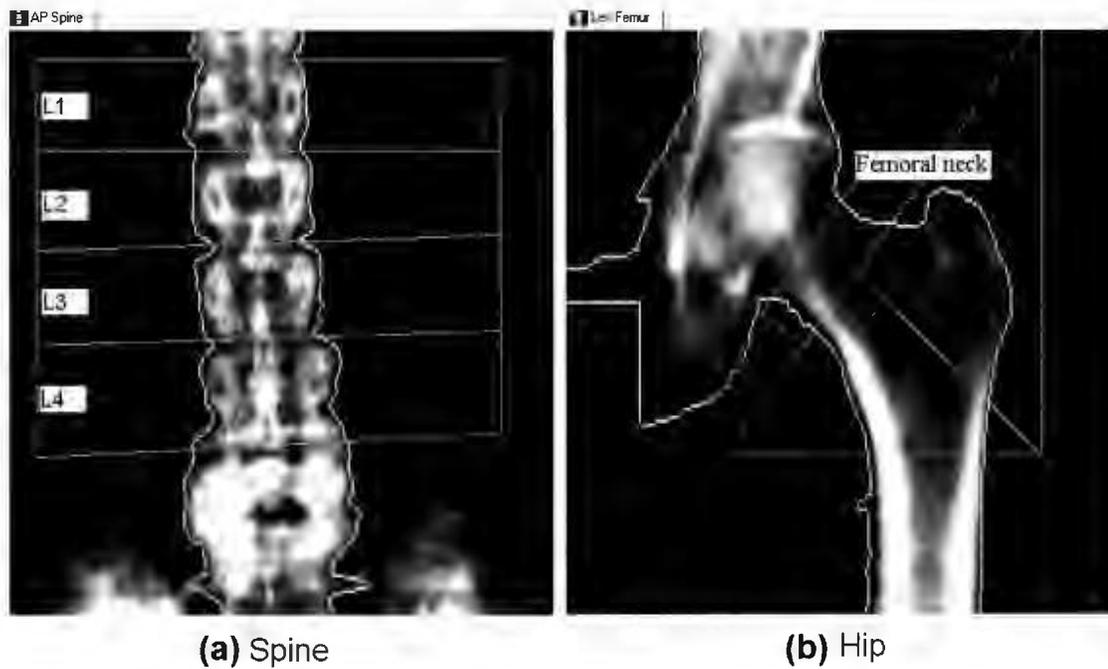


Figure 1.4: (a) Demonstration of a BMD image of the lumbar spine and (b) the hip, obtained using a GE Lunar Prodigy Dual-energy X-ray absorptiometry densitometer.

the image indicate bone because bone attenuates (absorbs) X-ray more than the tissue. The BMD value of each examination site for the hip is given, such as the femoral neck, trochanter, and wards, etc. The calculation of BMD values in the spine is the averaged bone density. For example, $BMD_{L1-L4} = (BMC_{L1} + BMC_{L2} + BMC_{L3} + BMC_{L4}) / (Area_{L1} + Area_{L2} + Area_{L3} + Area_{L4})$ where BMC is the bone mineral content. Similarly, the BMD calculation of the total hip is the mean BMD value of all BMD values of each examination site involved in the hip region.

The DXA densitometer used at the General Hospital in St. John's, NL is the Lunar Prodigy manufactured by General Electric (GE) and the DXA densitometer used at the St. Clare's Mercy hospital in St. John's, NL is the QDR 4500 manufactured by Hologic. The GE densitometer operates at photon energies of 40 and 70 keV while the Hologic densitometer has an X-ray tube of 70 and 140 keV. A bone density test normally takes 5 to 10 minutes.

DXA densitometers from different manufacturers perform differently in obtaining BMD measurements. The second project discussed in this thesis uses data from both the GE and Hologic densitometers. Both densitometers use fan-beam design instead of pencil beam. In general, the BMD values from the GE densitometer are larger by approximately 11.7% than those from the Hologic densitometer for the same patient [31]. The main reason for the difference is due to the fan beam design and the method used for calculating the BMD value [32]. The GE Lunar Prodigy densitometer has a smaller X-ray aperture and fewer detectors, therefore providing a smaller X-ray flux. As a result, the GE Lunar prodigy densitometer interrogates a narrower anatomical region, requiring multiple scans to be collected to complete a single measurement. The scan dimension for the GE Lunar Prodigy densitometer is 197.5×60 cm while for the Hologic QDR densitometer is 195.6×67 cm [33]. As

a result, the examination time of the GE Lunar Prodigy densitometer is longer than that of the Hologic QDR 4500 densitometer. Under the same scan speed, the GE Lunar Prodigy densitometer has less precision and image quality than the Hologic QDR 4500 densitometer. However, the Hologic QDR 4500 densitometer suffers from magnification problems due to the wide angle X-ray beam which means that objects scanned at a closer distance have less magnification than those scanned at a further distance.

1.8 My Contributions

In this section, I explain my role in both the *GFR* and BMD projects with respect to the theory, programming, data analysis, and physics (quality control of the instruments).

I started my PhD and became involved in the *GFR* project at the stage of developing the theory of the Tikhonov regularization. At this stage, the idea of using Tikhonov regularization for solving the gamma variate function had already been conceived, however, the implementation of the idea with regard to the mathematical equations and the programming was not done. I contributed to the derivation of the theory, such as finding the error propagation terms for *GFR* and *Vol*, which are introduced as *ERG* and *ERV* in Section 2.6.4. I also contributed to finding the best minimization method, the Nelder Mead method, after trying various minimization methods.

I contributed significantly to the computer programming. The initial code for the Tk-GV model with the Nelder Mead minimization was created by me in 2008. I did the programming for all the subsets (Leave-one-out, Leave-two-out, Leave-three-out, Leave-four-out, Bootstrapping, Jackknife, and the hump subsets) for all four

models between 2008 and 2009 on my own. I performed 90% of the calculations for the results and also verified the results from Mathematica with those from other software packages such as Matlab and Visual Basic 6, in 2009. I created the code for the patients' data from Dr. Burniston in 2010, including defining the scaling factor for the concentration values, performing all the calculations using the full samples and subsets, and analyzing the results. Between 2010 and 2011, I created the final software which enabled the clinical application of the Tk-GV method for the technologists at the General Hospital. This software provided a user-friendly interface that the technologist could use under the Microsoft Excel environment (I implanted the required linkage between Excel and Mathematica) without the need to learn how to use Mathematica. This software was also adopted and used by researchers from other institutions. I calculated approximately 90% of all the results and analyzed them in collaboration.

In addition, I worked on developing a website (www.renalfunction.org) for the application of the Tk-GV model. The goal of this website was to attract users to use the Tk-GV method for estimating the *GFR* values by inputting the initial information, including the concentration of the radioactivity and the collection time. Although I built the website at the request of my supervisor, he has since decided to discontinue it by allowing the domain of the website to expire in 2010. In 2009, I developed a program for calculating the dosage for thyroid function examination using Visual Basic 6, which has since been used by the technologists in the department of Nuclear Medicine at the General Hospital.

Moreover, I have improved the application of the Tk-GV method from a medical physics perspective by implementing quality control and quality assurance of the instruments used for the plasma clearance method, as well as modifying the laboratory

protocol and educating the technologists. The scintillation well counter was calibrated and tested on a regular basis using a linearity test, counting statistics, and Chi-square tests. To perform these tests, one needs to understand the basic physics principles of the instruments and how to analyze the results from the instruments. In 2010, I spotted the problem of the multi-well counter in counting and repaired it by taking the counter apart and tuning the mechanical parts inside the instrument under the instructions of the technical service of the manufacturer. Moreover, the appropriate use of the air displacement pipetting made a significant difference to the accuracy of the samples. The sampling time was recorded using a stop-watch, which was initiated by me. All the radioactivity counting was considered for decay correction.

Between 2010 and 2011, I participated in the application of the Tk-GV model to 10 local patients' studies, during which I performed the quality control of the gamma counter the day before the study and the day during the test, introduced the use of an electronic balance for better controlling the pipetting error, recorded the sampling times, counted all the samples, calculated the results and reported them to the physicians. All the results of the 10 patients' *GFR* values and data analysis based on these 10 patients were performed on my own. A paper written by me outlining the results of these 10 patients was presented to my supervisor in 2011.

The idea of the BMD project was proposed in 2009 by my supervisor. In order to initiate this work I wrote and applied for a research grant and the human ethical approvals on my own. I collected all the BMD data from both the General Hospital and St. Clare's Mercy Hospital in 2009 and completed all the data processing, programming, and analysis on my own. My supervisor provided advice on what distribution function to use for fitting the histogram. I performed multiple tests on different distribution functions using Mathematica in order to obtain the final distri-

bution function. I also wrote the code for performing Monte Carlo simulations for the BMD project, however, these results could not be included in my thesis. After presenting these findings at a conference, I realized there was a need to expand the investigation of the BMD study from the local dataset to a national dataset, in order to determine the general range of the calculated criterion. I wrote an application to the Canadian multi-center osteoporosis study (CaMos) for sharing their data, and this request was permitted in 2010. I performed a thorough calculation and data analysis on the CaMos dataset on my own. I have independently completed two paper drafts based on my findings for this BMD project.

Chapter 2

Glomerular Filtration Rate (*GFR*) Determination

This chapter introduces the concept of glomerular filtration rate (*GFR*) along with the methods commonly used for measuring *GFR*. The plasma clearance method is discussed in detail due to its importance to this work. A literature review of plasma clearance methods is given and the motivation of this study is also presented.

2.1 Glomerular Filtration Rate

The kidneys are important for regulating body fluid, osmolarity, electrolyte and pH balance, excreting metabolic and foreign substances and secreting hormones [34]. Urine is the major waste product from the kidneys [35]. The volume of urine is controlled by the balance between fluid intake and output.

Figure 2.1 shows a schematic representation of the anatomy of the kidney. This figure is a general demonstration of the kidney and some of the concepts discussed in the following are not shown in this figure. The kidney is composed of two major

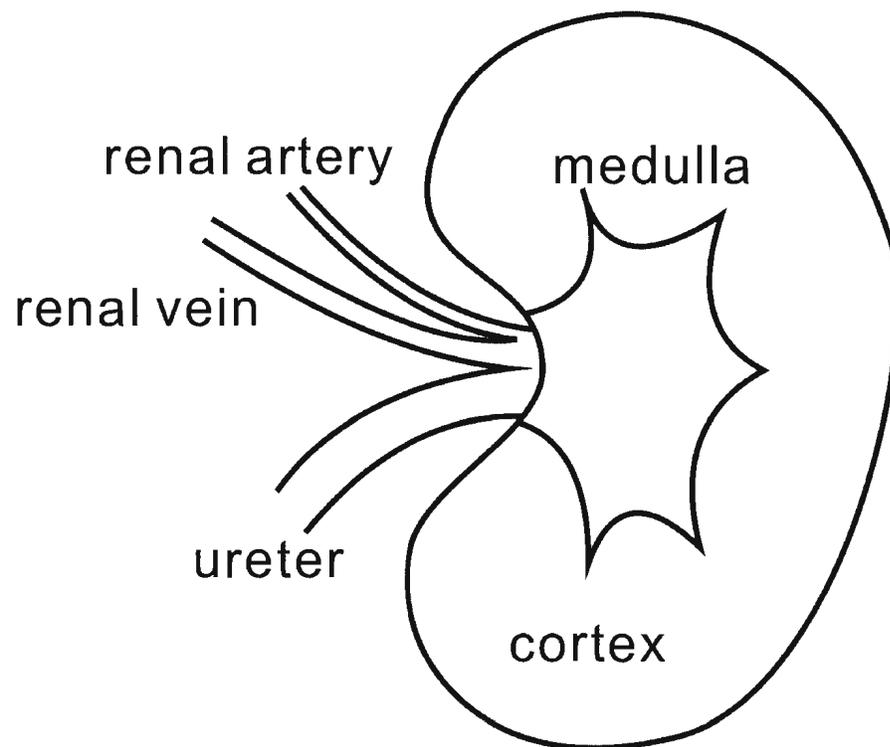


Figure 2.1: Schematic representation of the anatomy of a human kidney.

regions: the cortex (the outer region) and the medulla (the inner region). Inside the medulla there are multiple cone-shaped masses called renal pyramids [35]. Blood enters the kidney via the renal artery, going through smaller arteries and eventually coming out from the afferent arteriole. This is where the blood enters the glomerulus. The solution that comes out of the glomerulus contains the waste product being excreted from the ureter.

The smallest functional unit of the kidney is the nephron. For a healthy adult, there are approximately 1.5 million nephrons in one kidney. Each nephron is composed of a glomerulus through which fluid is filtered, and a long tubule in which fluid is converted to urine [35]. The filtered fluid first enters the proximal tubule and then the loop of henle. Coming out the loop of henle, the fluid enters the distal convoluted tubule and then the collecting duct that collects the urine. The blood flow to both kidneys is approximately 21% to 25% of the cardiac output. There are four mechanisms involved in the nephron: filtration, excretion, reabsorption, and secretion.

The total function of all the nephrons in the kidneys is called the glomerular filtration rate (*GFR*). *GFR* is defined as the volume of fluid filtered from the glomerular capillaries into the Bowman's capsule per unit time [35]. *GFR* is typically expressed in milliliter per minute (ml/min). *GFR* indicates how fast the kidneys filter blood and is proportional to the clearance of wastes from the body. *GFR* is usually used to indicate the level of kidney function. The normal range of *GFR* for a healthy adult is between 100 and 130 ml/min/1.73 m² which represents approximately 170 liters (L) of fluid processed by the kidneys per day [35]. This is equivalent to refreshing all the fluid in the body three times per day. The unit ml/min/1.73 m² is commonly used for describing estimated kidney function because it is realized that

the absolute value of the calculated GFR values (in units of ml/min) is insufficient to reflect the kidney function for different body habitus. For example, the normal range of GFR values for obese people is higher than that for the non-obese people [36]. Therefore, in order to better compare the estimated GFR results for these patients, one recommendation is to calibrate the GFR using the ratio of the patient's body surface area (BSA) to the standard BSA value 1.73 m^2 . BSA is also commonly used to indicate the patient's metabolism. A commonly used formula for calculating BSA is the Haycock formula [37], $BSA(\text{m}^2) = 0.02425 \times \text{height}(\text{cm})^{0.3964} \times \text{mass}(\text{kg})^{0.5378}$. The BSA correction for GFR is written as $GFR_{\text{corrected}} = GFR \times 1.73 / \text{BSA}$ [38]. The standard BSA for healthy adults is 1.73 m^2 and less for children. As an example, if the patient's calculated GFR is 100 ml/min and his/her BSA is 1.95 m^2 , the corrected $GFR = 100 \times 1.73 / 1.95 = 88.72 \text{ ml/min}/1.73 \text{ m}^2$. There have been studies questioning how well the BSA can calibrate the GFR difference among different people [39], however, this is not the focus of this study. We used the BSA corrected GFR mainly for the purpose of comparing our results with the GFR values used in the guidelines for determining kidney diseases.

GFR is closely related to the renal blood supply. If renal blood flow (RBF) is defined as the volume of blood delivered to the kidneys per unit time, another measurement called the filtration fraction (%) = (GFR/RBF) is used to describe the kidney function. This filtration fraction is approximately 20% for a healthy adult [35].

2.1.1 Impact of Kidney Function

Kidney function varies with different factors such as age, gender, race, body habitus, diet, etc. For example, women generally have lower GFR values than men at the same age and Asians tend to have lower GFR values than Caucasians. The peak

value of GFR is achieved after puberty. After the age of 40, the amount of functioning nephrons starts to decrease by approximately 10% per decade [35]. This gradual reduction is not life-threatening since the human body adapts to these changes correspondingly [40]. However, acute loss or abnormal loss of nephrons should be monitored for further investigations.

Accurate early diagnosis of kidney function decrease can effectively prevent the onset of more severe kidney diseases. Kidney diseases include acute renal failure (over hours or days) and chronic renal failure (over 3 months) [35]. With respect to pathology, kidney diseases can occur at the renal blood vessels, the glomeruli, the tubules, or parts of the urinary tract and bladder, etc. Kidney diseases can also result in complications such as cardiovascular disease, anemia, malnutrition, and so on [35]. As a result, a precise quantitative measurement of GFR is necessary for evaluating kidney function.

Chronic kidney disease (CKD) is defined as the progressive loss of kidney function over time. In Canada, there are nearly 3 million people with CKDs [41]. GFR is used to identify the different stages of CKDs. Table 2.1 shows the five stages of CKDs (CKD1, ..., CKD5) classified by the National Kidney Foundation in 2002 [42]. When the patient's GFR loss is greater than 90% of his/her expected GFR (hence, the remaining GFR is approximately less than 15 ml/min/1.73 m²), the patient is diagnosed as being at the end stage of renal disease or kidney failure. At this stage (CKD5), dialysis or kidney transplant is required.

The accuracy of the GFR measurement is important for monitoring kidney function changes, especially for patients at the boundaries of different CKD stages or at the end stage. Some acute kidney failure can be reversible if early diagnosis is made available, however, miscalculations of GFR might lead to inappropriate treat-

Table 2.1: The stages of chronic kidney disease (CKD). All GFR values are in units of ml/min/1.73 m² (1.73 m² represents the standard body surface area (BSA) of a healthy young adult).

Symbol	Stage	$GFR_{cr-corrected}$
CKD1	Early Stage	≥ 90 and evidence of kidney damage
CKD2	Mild	(60, 89) and evidence of kidney damage
CKD3	Moderate	(30, 59)
CKD4	Severe	(15, 29)
CKD5	Kidney Failure	< 15

ment plans and cause potential damage to the kidneys. For example, overestimation of GFR may misdiagnose the disease stage and delay treatment while underestimation of GFR may lead to unnecessary treatment or over-medication. Precise GFR measurements are important for CKD pre-screening programs such as the Kidney Early Evaluation Program (KEEP) in the United States and the Kidney Foundation of Canada [43, 44]. Preliminary results from KEEP have shown that there is an approximate 50% delay in the early diagnosis of CKD which significantly increases the potential burden for treating later stage kidney diseases [45]. Moreover, because kidney diseases are correlated with many other disorders, accurate diagnosis of CKD can benefit the detection of other relative health issues such as hypertension and diabetes [35].

2.2 GFR Measurements

2.2.1 Biomarkers

In situ biomarkers for GFR measurements include blood urea nitrogen (BUN) and serum creatinine (S_{cr}) [46]. A BUN test measures the amount of urea nitrogen in

the blood whereas S_{cr} measures the level of serum creatinine in the blood. BUN is a natural waste product from protein metabolism. Serum creatinine (S_{cr}) is a waste product of creatine phosphate in muscle and therefore the S_{cr} level depends on the muscle mass of the body. Both biomarkers are produced internally and eliminated mainly through filtration by the kidneys. Normal human blood contains 7 to 21 mg/dL of BUN and a S_{cr} level of 0.5 to 1.0 mg/dL for women and 0.7 to 1.2 mg/dL for men. An increase in either BUN or S_{cr} indicates a potential decrease in GFR . However, the level of BUN or S_{cr} can be affected by many other factors besides kidney diseases. For example, diabetes and hypertension can result in elevated BUN. High muscle content, meat digestion or certain medications can increase the S_{cr} level. The major drawback of using BUN or S_{cr} for GFR calculation is that their concentrations will not change significantly until two thirds of the kidney function is lost. As a result, S_{cr} or BUN is insufficient for detecting early stage kidney diseases.

A more accurate GFR measurement is based on the creatinine concentration which involves the collection of both blood samples and 24-hour urine samples. The collection of the 24-hour urine samples is performed by collecting all the urine from the patient during a 24-h period [47]. From this, the GFR can be calculated as:

$$GFR_{cr} = \frac{U_{cr}V}{S_{cr}}, \quad (2.1)$$

where the U_{cr} (mg/ml) is the urine creatinine concentration, V (ml/min) is the flow rate of urine (determined by the volume of urine divided by the period of time used for the urine collection) and S_{cr} (mg/ml) is the serum creatinine concentration. This calculated GFR_{cr} can be further corrected using BSA as $GFR_{cr\text{-corrected}}$ [38],

$$GFR_{cr\text{-corrected}} = \frac{GFR_{cr} \times 1.73}{BSA}. \quad (2.2)$$

Besides the *In situ* biomarkers, there are external biomarkers available which are administered intravenously or orally for estimating GFR values. One example is inulin, a natural polysaccharide found in many types of plants. Inulin is known as the ideal GFR biomarker because it is neither reabsorbed nor excreted by the glomerulus in the kidneys. Studies have shown that more than 98% of inulin is filtered by the kidneys [46]. Another study has indicated that inulin can also be reabsorbed by the tubules in the kidneys besides the glomerulus which leads to an overestimation of the GFR value by approximately 15% [48]. Inulin can also cause allergic reactions such as anaphylaxis [49]. Moreover, the inulin method sometimes requires the collection of urine samples which adds to the length in time and uncertainty of the examination procedure. In some cases, patients need to be catheterized to ensure the accuracy and completeness of the urine collection [50]. The inaccuracy in the 24 h urine collection can be as much as 20% [47]. Inulin based GFR measurements also suffer the same limitation as BUN or S_{cr} where the inulin clearance will remain normal until 30% of the kidney function is lost, due to hyperfiltration (body is adapting to the early loss of nephrons) of the remaining nephrons.

The gold standard for measuring GFR uses a constant infusion (injection) of inulin [51]. Constant infusion techniques require a system to deliver the solution (herein inulin) at a constant rate (e.g. 0.5 ml/min) to the body. The injection process normally takes approximately 15 to 30 minutes but in some cases can take hours [51, 52]. During the infusion, blood samples, with or without timed urine samples, are collected for estimating the GFR values. If urine samples are included in the method, the GFR is estimated using

$$GFR_{In} = \frac{U_{In}V}{S_{In}}, \quad (2.3)$$

where the U_{In} (mg/ml) is the urine inulin concentration, V (ml/min) is the flow rate of urine and S_{In} (mg/ml) is the serum inulin concentration. If the method does not require urine samples, the *GFR* is estimated using

$$GFR_{\text{In}}^{\text{non-urine}} = \frac{I_{\text{In}} \times R_{\text{In}}}{S_{\text{In}}}, \quad (2.4)$$

where I_{In} is the concentration of inulin in the infusion injection and R_{In} is the infusion rate [53].

The advantage of using constant infusion of inulin is that it can avoid the acute side effects of inulin in certain patients. It also allows for a transient equilibrium of the tracer to be reached in order to calculate the plasma clearance. The disadvantage of using constant infusion of inulin method (regardless of whether or not urine samples are used) is that for patients with impaired kidney functions, their *GFR* values cannot be accurately measured because the concentration of the tracer continuously increases for a long period of time, hence making it difficult to assess the elimination stage [53, 54]. Studies have shown that inulin-based *GFR* measurements can overestimate the real *GFR* by approximately 8.1 ml/min [51]. Therefore, despite being an excellent tracer, inulin is not commonly used for estimating kidney function. We did not use inulin in our study, but used ^{99m}Tc -DTPA instead as the tracer for estimating *GFR* values.

Another external biomarker iohexol is known as a radiocontrast agent for medical imaging. The clearance of iohexol is recorded using time-based blood samples collected after the injection [55]. The iohexol based *GFR* is calculated as

$$GFR_{\text{Iohexol}} = \frac{1}{0.0016 + t/Vol} \ln \frac{\text{Dose}}{C_t \times Vol}, \quad (2.5)$$

where Dose is the injected dose of iohexol (in units of mg), t is the time after the injection (in units of min), C_t is the iohexol concentration (in units of mg/ml) obtained at time t , and Vol is the volume of distribution expressed as $Vol(ml) = 187 \times \text{mass (kg)} + 732$, and the calculated GFR_{Iohexol} is in ml/min [55]. The estimated GFR_{Iohexol} can also be corrected using BSA or other body scaling methods.

2.2.2 Estimated *GFR* (*eGFR*)

Biomarker-based *GFR* estimations are usually inaccurate, especially for detecting low *GFR* values or mild *GFR* changes at early stages of CKD [34, 56]. One important limitation of the current *GFR* estimation is the inappropriate consideration for the various body habitus between individuals. For example, females and males differ in their lean body mass, therefore, S_{cr} based *GFR* needs to be corrected for gender. The BSA correction can compensate for some of the inaccuracies, however BSA itself is a biased estimation of body habitus (see Chapter 7). More variables besides BSA (BSA considers two parameters: body mass and height) should be considered to describe the expected metabolism or kidney function appropriately [51, 57]. Moreover, the criterion for determining the CKD stage has been obtained from a study group involving mainly Caucasians without the consideration of African Americans (with relatively higher metabolic rate) or Asians [42].

In order to take into account more parameters for estimating *GFR*, a method called estimated *GFR* (*eGFR*) was produced [57, 58]. The two formulae most commonly used for estimating *eGFR* are the Cockcroft-Gault (CG) formula, and the Modification of Diet in Renal Disease (MDRD) formula [57, 58]. The CG formula

developed in 1976 is given by:

$$eGFR_{CG} = \frac{(140 - \text{Age}) \times \text{Mass} \times (0.85 \text{ if Female})}{72 \times S_{cr}}, \quad (2.6)$$

where the resulting $eGFR_{CG}$ value is in mg/dL, Age is in years, Mass is in kg and S_{cr} is in mg/ml. The MDRD formula uses four variables: S_{cr} , Age, ethnicity and gender and is given by:

$$eGFR_{MDRD} = 186 \times S_{cr}^{-1.154} \times \text{Age}^{-0.203} \times (1.212 \text{ if Black}) \times (0.742 \text{ if Female}). \quad (2.7)$$

The $eGFR$ formulae have been shown to improve the accuracy of conventional S_{cr} based GFR measurements [59]. At this point in time, there is not yet a generally accepted $eGFR$ formula partly due to the relatively small study group used for obtaining each equation. As a result, new formulae for calculating $eGFR$ have been reported [60–62].

The latest development of an $eGFR$ equation is the Schwartz formula [62], which was derived for evaluating the GFR values of children. The formula contains five variables: height, S_{cr} , Cystain C, BUN and gender, and is given by:

$$eGFR_{Schwartz} = 39.1 \times (\text{Height}/S_{cr})^{0.516} \times (1.8/\text{Cys})^{0.294} \\ \times (30/\text{BUN})^{0.169} \times (1.099 \text{ (if Male)}) \times (\text{Height}/1.4)^{0.188}, \quad (2.8)$$

where Cys is the serum level of Cystatin C. The limitation of this formula is that it is only applicable to children.

2.2.3 Renography

Renography is a type of kidney imaging technique based on the renal excretion of a radiopharmaceutical. Renographs are typically acquired using gamma cameras depending on the uptake sites of the radiopharmaceutical in the kidney [1]. SPECT and PET can both be used for renography study [63]. However, SPECT is the most commonly used imaging modality for collecting renographs. Various radiopharmaceuticals such as ^{99m}Tc -DTPA, ^{99m}Tc -DMSA, ^{99m}Tc -MAG3, ^{51}Cr -EDTA, and ^{123}I -OIH can be used for renography using SPECT. A standard renography analysis includes a series of digital images taken with a 10 second interval exposure collected over a period of approximately 30 minutes. *GFR* can be estimated from the contrast difference between the kidneys and the surrounding tissues in the renographs. However, the renograph only gives a relative *GFR* value instead of the absolute *GFR* value because the estimation is not based on the actual clearance of plasma but the characteristics of the obtained images as an indirect representation. The renographs are analyzed using computer software to obtain the distribution of radioactive counts over time in the kidneys and surrounding tissues of interest (e.g. aorta and bladder). These numerical data can be plotted as a time-activity curve (radioactivity as a function of time) known as a renogram.

A renograph directly shows the location and relative sizes of both kidneys [64]. A renogram is plotted as the radioactive counts (a measure of the concentration of the injected radiopharmaceutical) versus time (min), see Figure 2.2 in the following section. A renogram has three phases, shown in Figure 2.2 [65]. The first, phase I, represents the perfusion stage, appearing as a steep positive slope in the curve which occurs within the first few minutes. The second, phase II, extends from the perfusion to the peak value of the concentration value in the renogram. Phase III is the concave

negative slope part of the renogram indicating the elimination of the radiopharmaceuticals from the kidneys. Renograms are usually composed of time-activity curves for the left and right kidney, as well as the bladder. Normal kidneys are symmetric on both sides, leading to an approximately 50/50 split in kidney functions. Malfunctioning kidneys may have a different ratio between the two sides. There are cases where both kidneys are equivalently malfunctioning but still maintaining a 50/50 ratio. An abnormal renogram may indicate obstructions or insufficiencies in the kidneys. However, renograms cannot be used alone as the criterion for diagnosing kidney diseases. Other factors should be considered such as the total and individual kidney function, pathology findings and other quantitative measurements. A limitation of the time-activity curve method is that the time range of the renogram does not usually extend beyond 30 minutes, which may omit some characteristics of the later elimination stage of the kidneys. A detail description of renographs and renogram analysis were not given since they were not used in this study for estimating GFR values.

2.3 Plasma Clearance

As mentioned in Section 2.2, S_{cr} based GFR estimation is obtained from plasma clearance data. In this study, the plasma clearance of radiopharmaceuticals was the method of choice for estimating GFR . The plasma clearance method measures the filtration rate of kidneys and other organs (e.g. liver or bladder). Plasma clearance results indicate how fast plasma is filtered from the kidneys. The use of radiopharmaceuticals for estimating GFR in nuclear medicine dates back to the 1960s. At the present, the choice of radiopharmaceuticals includes ^{51}Cr -EDTA, ^{99m}Tc -DTPA, and ^{125}I -iothalamate, etc. The mathematical tools for estimating the GFR value from observed plasma data are usually called pharmacokinetic models.

General Theory of Plasma Clearance

There are two essential parameters used in the study of plasma clearance: *GFR* (*GFR* is used to represent *CL* described in Section 1.6 since *GFR* describes the clearance through the kidneys specifically) and the volume of distribution (*Vol*).

GFR is obtained from the measured radioactivity of a tracer in the plasma per unit volume using the following equation:

$$GFR = \frac{\text{Dose}}{\int_0^{\infty} C(t)dt}, \quad (2.9)$$

where Dose is the amount of administered radiopharmaceuticals measured in units of MBq in this study, and $C(t)$ is the concentration of the plasma (radioactivity measured from the blood samples collected from a patient) in units of MBq/ml at a given time t [66].

The volume of distribution (*Vol*) is very important for the investigation of kidney function and interpreting *GFR* results because *Vol* can help determine whether or not the estimated *GFR* is physiological or physical. *Vol* is the apparent volume used to quantify the distribution of a drug between the plasma and the rest of the body after its administration. In this case, *Vol* is defined as the theoretical volume in which the total amount of the tracer needs to be uniformly distributed to achieve the expected plasma concentration [35]. The *Vol* value is related to the estimated *GFR* value such that decreasing *GFR* usually causes *Vol* to increase due to the slower plasma clearance rate and fluid retention in the body. The *Vol* value is also compared with the total body fluid volume in order to determine whether or not the patient is dehydrated. *Vol* is calculated from the *GFR* as:

$$\langle Vol \rangle = MRT \times \langle GFR \rangle, \quad (2.10)$$

where the symbol $\langle \rangle$ indicates a time integrated average and *MRT* is the mean residence time defined as the average time a given tracer stays in the body, which is expressed as[†] [67].

$$MRT = \frac{\int_0^{\infty} t \cdot C(t) dt}{\int_0^{\infty} C(t) dt}. \quad (2.11)$$

Both *GFR* and *Vol* estimations are highly dependent on the observed concentration $C(t)$ values. However, in practice, $C(t)$ is obtained by taking a finite number of measurements at different time intervals. Therefore, in order to be able to perform the integration in Eq 2.9, it is necessary to represent the set of collected clearance data $(t, C(t))$ by a mathematical equation. Because the clearance data is sometimes provided with as few as two data points, it is imperative that the model used to describe the kidney function be as accurate as possible. Due to its importance, different models have been used to provide a mathematical description of $C(t)$ [51,68–74]. The validity of these models introduces variations in the accuracy and precision in the estimated *GFR* results. Various models used to present the clearance data are the subject of the next section.

2.3.1 Pharmacokinetic Models

The theoretical basis of plasma clearance calculations relies on the pharmacokinetic model analysis that uses mathematical models to simulate how the kidneys eliminate substances. There are two commonly used pharmacokinetic models: the *compartmental* model and the *non-compartmental* model [75].

[†]There is a similar concept called mean transit time *MTT* which is the average time a given tracer spends in the kinetic system. $MRT \neq MTT$ for non-instantaneous administration of the tracer. However, in most cases, $MRT = MTT$

The compartmental model is the most commonly used pharmacokinetic model in the study of plasma clearance as well as in other research fields such as biomedicine and epidemiology [67]. The compartmental model is a key topic of this thesis because of its limitations in describing the mechanisms of the kidneys. The compartment model treats the body or organs of interest as a combination of several compartments, each having a fixed volume of fluid. The compartments can be either closed to the external environment (entities only allow exchange between the compartments) or open (entities undergo dynamic interactions with the outside). Under the compartmental models in the plasma clearance study, each compartment is assumed to be homogeneous; the entities inside each compartment are exchanged at constant rates; entities can reach transient equilibrium instantly; and no chemical reactions occur between the entities and the carriers. Among all the assumptions, the instantaneous mixing and instantaneous transient equilibrium of the contents inside the compartments are the most inaccurate (see Chapter 3 for more details).

The non-compartmental model is simply viewed as a model that is independent of using compartments [67]. The non-compartmental model reduces the assumptions required for estimating *GFR*. For example, the non-compartmental model does not require the exchange rate between the system and input (or output) to be constant, nor that the exchange rate be uni- or bi-directional. Whereas the compartmental model is based on a linear or nonlinear mathematical description of the concentration $C(t)$, the non-compartmental model relies on a numerical integration of the concentrations to estimate the total drug (radiopharmaceuticals) exposure. Although the amount of injected dose is known, how the drug is distributed and eliminated from the body is critical for evaluating the effect of the drug. Therefore, the total drug exposure is related to both the injected dose and the distributing process of the drug

inside the body. Knowing the total drug exposure from the non-compartmental model can help determine the kidney function as well. The essential parameters for the non-compartmental models are the mean residence time (MRT), volume of distribution (Vol), area under the curve (AUC) and glomerular filtration rate (GFR). The

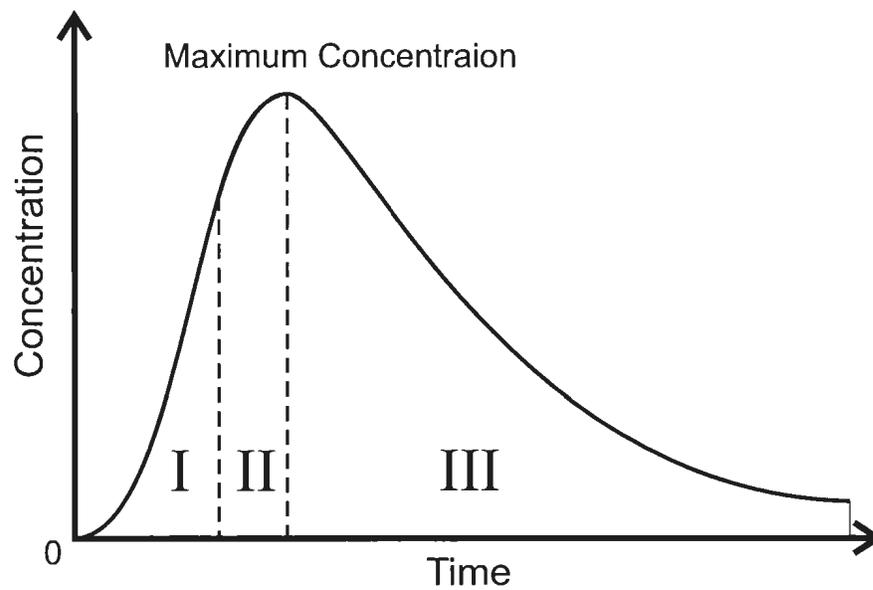


Figure 2.2: Schematic representation (lightly shaded area) of the Area Under the Curve (AUC) used for calculating the elimination rate of a tracer. The plasma concentration of the tracer is plotted as a function of the time (with time=0 set at the finishing point of the injection) after the administration. The three phases in the renograph are labeled as I, II and III.

denominator in Eq 2.9 represents the area under the curve and is a key parameter for both the compartmental and non-compartmental models. Figure 2.2 shows a schematic representation of the plasma concentration of a tracer versus time used to

calculate the $AUC = \int_0^{\infty} C(t)dt$ (the shaded area). The time $t = 0$ indicates when the tracer is administered. The time interval from the starting time to the peak concentration (maximal concentration) of the tracer is the perfusion stage. After the perfusion the concentration starts to decline gradually with time, which is called the clearance (or elimination) stage. How quickly the tracer is cleared from the body is obtained from the curve analysis of the AUC, which is used to estimate the plasma clearance. As stated above, the concentration-time curve is not produced from continuous data points but from finite individual measurements of the radioactivity of the blood samples. Therefore, the concentration-time curve is highly dependent on the sampling time schedule used to measure the concentration $C(t)$. The sampling schedule includes the sampling start time (beginning of the curve), the interval between each sampling and the range of the entire sampling time. Since the AUC calculation is based on the concentration-time curve, the longer the sampling time and the closer the time points (or the more time points), the more accurate the AUC calculation becomes.

The compartmental model has the advantage over the non-compartmental model to predict the concentration $C(t)$ at any time t provided the appropriate compartmental model is used [67]. However, in most cases, the compartmental model is inaccurate and imprecise for describing the kidney system because not all the required assumptions can be met [68]. On the other hand, the non-compartmental model is less dependent on the model used for estimating the AUC and clearance.

The non-compartmental model is theoretically more precise than the compartmental models assuming that the time-activity curve can be plotted accurately, thus allowing the better estimation of GFR values. However, the drawback of the non-compartment model is that the estimation of the AUC is highly dependent on the

sampling schedule. For example, T_{\max} and C_{\max} values are determined directly from the plotted activity-time curve. If there are missing samples at the critical point (the maximum concentration at T_{\max}), inaccuracy can be introduced into the results. Secondly, for patients with very low *GFR*, it is difficult to decide the sampling schedule in order to have sufficient data points for plotting the time-activity curve. That is why one usually needs to collect several blood samples for plotting the time-activity curve using the non-compartmental model (for example between 6 and 10 samples) [76–78]. Lastly, the non-compartmental model cannot predict the concentration at certain time points because it does not adopt the use of a model expression, therefore there are no parameters for predicting the concentration values at a certain time. For the purpose of better control of sampling schedule and accuracy in estimating *GFR*, the non-compartmental method was not used in this study.

In the following, an introduction of how the compartmental models are applied to this study and the limitations of the models are presented.

2.3.2 One-compartment Model

The one-compartment model treats the plasma as a single compartment, as illustrated by the circle in Figure 2.3. The two arrows indicate the flow of the radiopharmaceuticals in and out of the compartment. A key variable in the one-compartment model is the concentration of the tracer (radiopharmaceutical) $C(t)$ in the plasma [67]. In the one-compartment model, the rate of change of the concentration $dC(t)/dt$ is assumed to be proportional to the concentration $dC(t)/dt = -\alpha C(t)$ where α is the constant elimination rate of the tracer. The concentration is therefore expressed as $C(t) = Ae^{-\alpha t}$ where A is the concentration at the initial injection time $t = 0$.

The one-compartment model is denoted as the sum of one exponential term $C(t) =$

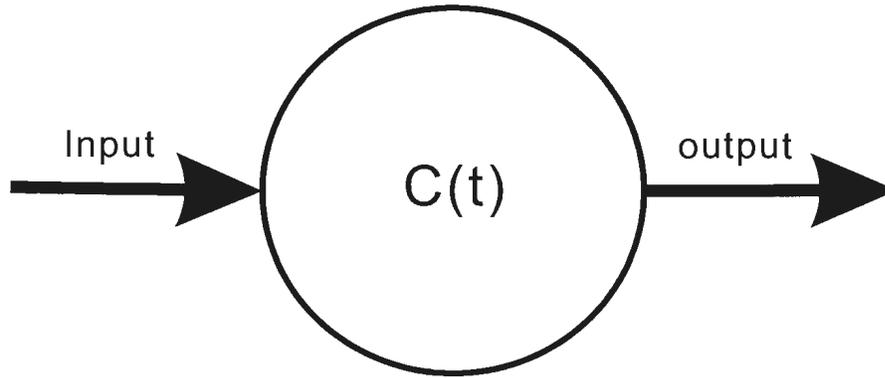


Figure 2.3: Schematic representation of the one-compartment model used to model the plasma clearance. The circle represents the plasma as the compartment. The two arrows indicate the flow of the radiopharmaceuticals in and out of the compartment.

$Ae^{-\alpha t}$ and is abbreviated as *SET1*. The *SET1* model has two unknown parameters A and α . Using the definition of *SET1* in Eq 2.9, the *GFR* can be calculated as:

$$\begin{aligned}
 GFR &= \frac{\text{Dose}}{\int_0^{\infty} C(t) dt} \\
 &= \frac{\text{Dose}}{\int_0^{\infty} Ae^{-\alpha t} dt} \\
 &= \frac{\text{Dose}}{A \left(-\frac{1}{\alpha} \right) e^{-\alpha t} \Big|_{t=0}^{t \rightarrow \infty}} \\
 &= \frac{\text{Dose}}{A \left(-\frac{1}{\alpha} \right) (e^{-\alpha \times \infty} - e^{-\alpha \times 0})} \\
 &= \frac{\text{Dose}}{A \left(-\frac{1}{\alpha} \right) (-1)} \\
 &= \frac{\text{Dose}}{A/\alpha} \\
 &= \text{Dose} \times \alpha/A.
 \end{aligned} \tag{2.12}$$

where Dose is the administered radiopharmaceutical. In order to determine the values of parameters A and α , at least two blood samples are required. If the concentrations of the two blood samples collected at time T_1 and T_2 are denoted as C_1 and C_2 (e.g.

in units of counts/min) respectively, the two equations to be solved are:

$$\begin{aligned} C_1 &= Ae^{-\alpha T_1} \\ C_2 &= Ae^{-\alpha T_2}. \end{aligned} \tag{2.13}$$

where the two parameters A and α are given by

$$\begin{aligned} A &= C_1 \exp\left(\frac{\ln(C_1/C_2)}{T_2 - T_1} T_1\right) \\ \alpha &= \frac{\ln(C_1/C_2)}{T_2 - T_1}. \end{aligned} \tag{2.14}$$

Therefore, *GFR* from Eq 2.12 can be written as

$$\begin{aligned} GFR &= \text{Dose} \times \alpha/A \\ &= \text{Dose} \times \frac{\ln(C_1/C_2)}{T_2 - T_1} / \left\{ C_1 \exp\left(\frac{\ln(C_1/C_2)}{T_2 - T_1} T_1\right) \right\} \\ &= \text{Dose} \times \left(\frac{\ln(C_1/C_2)}{T_2 - T_1}\right) / \exp\left(\ln C_1 + \frac{T_1 \ln(C_1/C_2)}{T_2 - T_1}\right) \\ &= \text{Dose} \times \left(\frac{\ln(C_1/C_2)}{T_2 - T_1}\right) / \exp\left(\frac{T_2 \ln C_1 - T_1 \ln C_1 + T_1 \ln C_1 - T_1 \ln C_2}{T_2 - T_1}\right) \\ &= \text{Dose} \times \left(\frac{\ln(C_1/C_2)}{T_2 - T_1}\right) / \exp\left(\frac{T_2 \ln C_1 - T_1 \ln C_2}{T_2 - T_1}\right). \end{aligned} \tag{2.15}$$

The injected dose of the tracer cannot be measured using a gamma counter due to its high radioactivity. It is therefore measured using a dose calibrator. In order to compare the radioactivity of the injected dose and the collected plasma samples without the detector difference, and also to standardize the injected dose used between different tests, it is necessary to introduce a conversion factor. The term Dose in Eq

2.15 is scaled using the standard dilution of the same radiopharmaceutical:

$$\text{Dose} = \frac{(D - R) \times ADS \times DS}{AS} \quad (2.16)$$

where D is the radioactivity of the injected dose (measured using a dose calibrator), R is the radioactivity of the residue of the injected dose (mostly in the syringe used for the injection) after injection, ADS is the radioactivity of 1 ml of the standard solution, DS is the volume of the standard solution (e.g. 500 ml) and AS is the radioactivity of the standard solution [69]. The reason for scaling the radioactivity of the plasma samples using a standard solution is to make the injected dose measurement from the gamma counter comparable with the plasma radioactivity measurement using the same gamma counter. The standard solution is prepared and measured in the same fashion as the injected dose, and the radioactivity of the injected dose is approximately 10 times that of the standard. In some institutions, a weighting technique (e.g. using a electrical balance) is used instead of measuring the volumes of the blood sample and the standard solution.

2.3.3 Two-compartment Model

The two-compartment model is shown schematically in Figure 2.4. Different from the one-compartment model that uses one compartment to describe the plasma clearance, the two-compartment model uses two compartments (C_1 and C_2) to represent a more complicated mechanism for the plasma clearance [67]. C_1 is the compartment that includes the plasma where the glomerular filtration occurs (similar to that used in the one-compartment model). C_2 represents the less accessible portion of the plasma distribution space. C_2 is assumed to exchange the tracer with C_1 at a rate proportional to the amount of tracer in each compartment (C_1 or C_2) [79]. Each compartment has

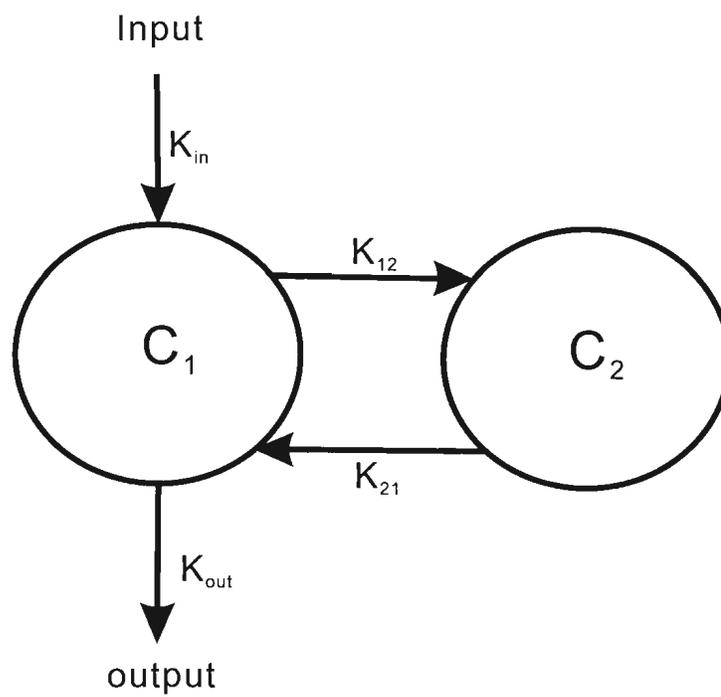


Figure 2.4: The two-compartment model used to describe plasma clearance. The two circles C_1 and C_2 represent compartments for the plasma and secondary volume of distribution of the tracer respectively. The flow in and out of the two compartments is indicated with arrows showing the rate of exchange K_{in} , K_{out} , K_{12} and K_{21} .

a volume of distribution and four exchange rates: K_{in} , K_{out} , K_{21} , K_{12} . Therefore, the two-compartment model describes the concentration $C(t)$ of the whole plasma system as the sum of two exponential terms, denoted as *SET2* given by $C(t) = K(\alpha e^{-\lambda_1 t} + e^{-\lambda_2 t})$. The parameters K , α , λ_1 , and λ_2 are related to the rates K_{in} , K_{out} , K_{12} , K_{21} in Figure 2.4. The parameters λ_1 , λ_2 and K are independent of one another. The ratio α of the two exponential terms $e^{-\lambda_1 t}$ and $e^{-\lambda_2 t}$ shows the percentage of the fast elimination rate to the slower elimination rate. This parameter α is unrelated to the parameter α used for the *SET1* model (e.g. Eq 2.13). As a result of having four parameters (K , α , λ_1 , and λ_2), at least four blood samples are required to use the two-compartment model.

Only the one-compartment and two-compartment models are introduced and used in this work. Other models with more compartments such as the three-compartment model or the four-compartment model are not used because of their requirement of more blood samples. In addition, using more compartments (more than 2) does not necessarily provide physiological meaning for each compartment. One of the goals of this work is to find the best model to calculate the plasma clearance based on four blood samples.

2.4 Limitations of the Compartment Models

The *SET1* and *SET2* models both belong to the general model called the *sum of exponential terms*, denoted as *SETs*, where s can be any positive integer $s = 1, 2, 3, \dots$. The *SETs* models have been used for over 50 years for estimating plasma clearance rate [80]. As a result of numerous clinical applications of the *SETs* models, there have been increasing reports of their limitations in estimating *GFR* values [68]. For example, the *SET1* model occasionally gives negative *GFR* values for patients with

very low kidney function while the *SET2* model has been reported to have difficulty in converging during the fitting of the time-activity curve. Also the repeatability of the *SETs* models is very poor if different sampling times are used for the same patient [81].

There have been several debates on the choice of sampling times [71,81]. Different laboratory protocols use different sampling times. For example, there are suggestions indicating that the following sampling times should be used for the *SET1* model: (1 h, 3 h), (1 h, 4 h), (2 h, 4 h) or (3 h, 4 h) [71,82,83]. In order to ensure that the samples are collected during the elimination stage instead of the perfusion stage, the earliest collection time is recommended to be 1 hour post-injection [71]. If the first sample was collected during the perfusion stage and the second sample was collected during the elimination stage, the *SET1* model would have difficulty fitting the time-activity curve. As a result, the estimated *GFR* would likely be underestimated if not non-physical. If the patient is undergoing a redistribution of the tracer, this could result in $C(t_1)$ being smaller than $C(t_2)$ (where $C(t_1)$ should be larger than $C(t_2)$). Therefore, this would indicate that the concentration $C(t)$ is increasing with time due to an accumulation process instead of an elimination process, also leading to a negative *GFR*. A negative *GFR* is physiologically impossible. Because the *SET1* model is dependent on the sampling time, estimations using different combinations of sampling times can vary largely, leading to the major source of instability of the *SET1* model [71].

Monte Carlo simulations have been used to investigate the error introduced by the different sampling schedules while using the *SETs* models. For example, Sadeler et al. showed that the errors in estimated *GFR* results using the *SET1* model due to the uncertainty of sampling schedule cannot be reduced through Monte Carlo

simulations [84]. Dr. Russell pointed out, in his study using Monte Carlo simulations, that the overall sampling time is more important than the number of samples used for estimating *GFR* values [85]. Murray et al. used 20,000 Monte Carlo simulations for finding the best sampling schedule using the *SET2* model and suggested using a minimum of 5 blood samples to obtain an accurate estimation of *GFR* values [86]. Results using Monte Carlo simulations have shown that thorough consideration should be given to the sampling schedule while using the *SETs* models for estimating *GFR* values.

The description of the elimination rate of the tracer using the *SETs* models is also problematic. Neither the elimination factor α in the *SET1* model $C(t) = Ae^{-\alpha t}$, nor the factors λ_1 and λ_2 in the *SET2* model $C(t) = K(\alpha e^{-\lambda_1 t} + e^{-\lambda_2 t})$, are time-dependent. The fact that the rate of clearance declines with time and with the tracer concentration requires a more sophisticated model to describe the elimination rate. Although adding more exponential terms might better approximate the changing elimination rate, a better solution is to replace the constant elimination rate by a time-dependent variable.

The injection method used to administer the tracer conflicts with the assumptions of the *SETs* models. During a bolus injection, the dose is injected within a few seconds, however, the *SETs* models assume that the plasma concentrations in the compartment achieve instant equilibrium. This is an over simplification of the perfusion stage because if patients have any obstructions in their plasma flow or experience radiopharmaceutical adhesion with the tubules in the kidneys, the perfusion would be delayed by minutes or hours. This clearly affects *GFR* results. A possible solution is to use a constant infusion technique for injecting the radiopharmaceuticals which allows a slow mixing of the radiopharmaceuticals into the blood stream.

After a constant infusion, a transient equilibrium status can more easily be achieved. However, performing constant infusion requires special equipment and is more time consuming and as a consequence is less convenient for the patients.

Because of the limitations of the *SETs* models, several attempts have been proposed to improve their performance. The first attempt was to increase the number of compartments in the *SETs* models. The difference between the observed concentration-time curve and the fitted curve using the *SETs* models appear to decrease with an increasing number of compartments. One explanation is that increasing the number of compartments increases the number of exponential functions used to fit the data. However, the drawback of having more compartments in the *SETs* model is the requirement for more blood samples to be taken. For example, the three-compartment model treats the plasma clearance with three compartments: plasma, interstitial fluid and site of degradation [87]. The required number of blood samples is at least 6 for the three-compartment model.

A common attempt to modify the estimated *GFR* has been to introduce a correction factor by comparing the *SETs* models (mostly the *SET1* model) with other models. There are several types of corrections. The first is to use a factor of 0.87 (known as the “Chantler” correction factor) for correcting the difference between the two-sample *SET1* model and the urinary clearance method, given by $GFR_{\text{corrected}} = 0.87 \times GFR_{\text{SET1}}$ [60]. The second is to subtract a constant of 8.1 ml/min to the estimated *GFR* obtained from the *SET1* model given by $GFR_{\text{corrected}} = GFR_{\text{SET1}} - 8.1(\text{ml}/\text{min})$ [88]. The third is to use a power function as a correction. For example, a correction for the results obtained from the *SET1* model using the tracer $^{51}\text{Cr-EDTA}$ has been proposed by using $GFR_{\text{corrected}} = GFR_{\text{SET1}}^{0.964}$ [71]. This power function index 0.964 varies with different radiopharmaceuticals, such as 0.979

for ^{99m}Tc -DTP. The fourth is to apply the BSA correction to the estimated *GFR* by using $GFR_{\text{corrected}} = GFR \times 1.73/\text{BSA}$. However, BSA has been questioned as a poor indication of body metabolism or expected kidney function despite its wide use [39]. Although various body scaling factors have been used to calibrate the estimated *GFR* values, such as age, gender, race, height, body mass, and the extracellular fluid volume [38, 89–91], none of the *GFR* correction formulae have been accepted as a standard. Moreover, all these *GFR* correction formulae were obtained from relatively non-representative (small sample size) study groups and therefore cannot be used for general applications [92, 93]. In Chapter 7, an alternative *GFR* correction is discussed.

As discussed in Section 2.3.1, the most important parameter for estimating *GFR* is the calculation of the AUC value. Attempts have also been made to apply more complicated mathematical regression methods for the AUC curve fitting. The conventional regression method is based on OLS algorithm. The principle of an OLS method is to minimize the sum of the squares of all distances between the observed values and the fitted values [94]. The OLS method has limitations in assuming that there is no collinearity within the variables. Unfortunately, this assumption cannot always be met. Moreover, an OLS based regression cannot always find a solution. There have been attempts to try to use different regression methods to replace the use of the OLS method. For example, Dr. Russell used the Bayesian nonlinear regression for both the *SET2* and the *SET3* models [73, 95]. Suitable estimated parameters of the plasma clearance models were found using statistical analysis such as the Chi-square test. The employment of the Bayesian regression indeed improved the curve fitting, however, the Bayesian method still suffers from all the disadvantages of the compartmental model such as that the instant equilibrium assumption is not satisfied.

2.5 Motivation

The motivation of this study is to develop a more precise and accurate model for estimating GFR . One of the biggest problems of the current methods (the $SETs$ models) is that they occasionally give negative GFR values (physiologically impossible) for patients with poor renal function. Even in the extreme cases where the patient has no kidneys, he/she should still have a GFR value greater or equal to 0 ml/min (never negative). Studies performed on dogs strongly indicated that the lowest value of GFR can reach 0 provided there is no elimination of the tracer, which indicates the possibility of having 0 kidney function [96].

In most cases, GFR is overestimated by the $SETs$ models, causing under-diagnosis or in some cases false-negative results [68]. In addition, the kidney function results are highly dependent on the blood sampling times, which means the estimated GFR can be altered if the blood samples are collected using different time schedules compared with the previous tests, indicating poor prediction and repeatability. Moreover, the current methods are incapable of estimating GFR for children since most GFR calculation formulae were obtained from an adult population [69]. To solve these problems a new pharmacokinetic model is proposed.

2.6 Theory

2.6.1 Gamma Variate (GV) Model

The first approach for improving the compartmental models was to adopt a dynamic mixing mechanism for describing the kidney function. The description of dynamic perfusion and dilution (or elimination) of the tracer in time is accomplished by the

proposed Gamma Variate (*GV*) model [97].

The *GV* model is derived from fractal theory. A fractal is a fragmented shape that can be split into parts, each of which is (or approximately is) a reduced-size copy of the original shape. Fractal theory has been used in different areas in medicine for decades such as for studying the structure of lungs, blood vessels, etc [98,99]. The key to fractal theory is the self-similarity between the sub-branches and the “parent” branch from which the sub-branches emanate.

In this study, there are two ways of applying fractal theory to plasma clearance study. One way is to treat the anatomy of the kidneys as a fractal structure where the finer structures in the fractal approximate the renal structures (to the level of nephrons). Another way to describe the flow in and out of the kidneys is by using the image of a “parent” branch and sub-branches. The similarity in the dynamic properties of the inflow and outflow of the tracer is described as a time-based power function. A power function (or power law function) is a typical mathematical expression for a fractal system. The most commonly used power function is $f(t) = kt^a$ where k (scaling exponent) and a (power index) are constants. The main property of a power function is the scaling invariance which means that if the variable t of function $f(t)$ is multiplied by a constant c , the resulting function is simply multiplied by the constant c^a , shown as $f(ct) = k(ct)^a = c^a f(t) \propto f(t)$. This property is consistent with the self-similarity of fractal theory in the way that scaling a function with a constant will not alter the property of the original function. The other property of power functions is the universality, where the whole system with the same characteristic exponents (e.g. power index a) can represent identical (or similar) basic behaviours. This property allows the simplification from a complicated system (with a nested “parent” branch and sub-branches) down to a critical basic system (the

“parent” branch alone). On the other hand, the properties of this critical point can be expanded back to investigate the whole system.

Different from the expression of the concentration $C(t)$ in the *SET1* model ($C(t) = Ae^{-\alpha t}$), the *GV* model describes the concentration as

$$C(t) = Kt^{\alpha-1}e^{-\beta t}, \quad (2.17)$$

where a time-dependent power function $t^{\alpha-1}$ is multiplied by the conventional exponential term $e^{-\beta t}$ to allow for the description of dynamic exchange between the plasma and the rest of the body. There are three parameters in the *GV* model: the constant K , the shape factor α used to describe the mixing rate (mixing the tracer with the plasma or the rest of the body to achieve a transient equilibrium distribution of the tracer) and the rate constant β for describing the renal elimination. This parameter α is unrelated to the parameter α used for the *SETs* models. The *GV* model estimates plasma clearance using the time integration of the concentration $C(t)$ from the area under the curve (AUC) analysis given by Eq 2.9. The physical (or physiological) range of the two parameters α and β is: $0 < \alpha \leq 1$ and $0 \leq \beta$. The reason for the boundaries of the two parameters is given in the next section.

The history of using the gamma variate function (*GV* model) for biomedical research dates back to 1963 when the *GV* model was used to describe the peripheral dilution curve of a cardiac output [100,101]. The peak value of the concentration of the tracer $C_{\max}(t)$ and the time to peak value can provide the first-order kinetics values (e.g. perfusion rate) indirectly. The *GV* model was also used to study extravascular lung water [102], hepatic function [103], and pharmacokinetics of drug disposition [100,104]. Notably, the *GV* model has been applied to model tracer kinetics in medical imaging studies such as PET and MRT perfusion [105,106]. However,

although the *GV* model has been proposed, it has not yet, to our knowledge, been applied successfully for plasma clearance analysis [68, 74].

2.6.2 Rate of Exchange, *GFR* and *Vol*

The main property that makes the *GV* model attractive for investigating the kidney function is the dynamic exchange rate of the concentration. Taking the derivative of $C(t)$ with respect to t gives

$$\begin{aligned} \frac{dC(t)}{dt} &= \frac{d(Kt^{\alpha-1}e^{-\beta t})}{dt} \\ &= -K\beta t^{\alpha-1}e^{-\beta t} + (\alpha - 1)Kt^{\alpha-2}e^{-\beta t}, \end{aligned} \quad (2.18)$$

which can be rewritten as a differential equation using the concentration $C(t)$ as:

$$\frac{dC(t)}{dt} = \left(\frac{\alpha - 1}{t} - \beta \right) C(t). \quad (2.19)$$

The calculations of *GFR* and *Vol* (volume of distribution) are based on the time integration of the concentration $C(t)$ from $t = 0$ to $t = \infty$ given by

$$\begin{aligned} GFR &= \frac{\text{Dose}}{\int_0^\infty C(t)dt} = \frac{\text{Dose}}{\int_0^\infty Kt^{\alpha-1}e^{-\beta t}dt} \\ &= \frac{\text{Dose} \times \beta^\alpha}{K\Gamma(\alpha)}. \end{aligned} \quad (2.20)$$

In Eq 2.20 Dose is the administered radiopharmaceuticals scaled by a standard solution of the same tracer (see Eq 2.16) and $\Gamma(\alpha)$ is the gamma function of α . The mean residence time *MRT* is calculated according to Eq 2.9 and is given by:

$$MRT = \frac{\int_0^\infty t \cdot C(t)dt}{\int_0^\infty C(t)dt} = \frac{\int_0^\infty t \cdot t^{\alpha-1} \cdot e^{-\beta t}dt}{\int_0^\infty t^{\alpha-1} \cdot e^{-\beta t}dt} = \frac{\alpha}{\beta}. \quad (2.21)$$

The volume of distribution Vol is then calculated from GFR and MRT as

$$Vol = \frac{\alpha}{\beta} \cdot GFR = \frac{\text{Dose} \cdot \alpha \cdot \beta^{\alpha-1}}{K\Gamma(\alpha)}. \quad (2.22)$$

Recall that the physical range for the two parameters α and β is: $0 < \alpha \leq 1$ and $0 \leq \beta$. The reason that α cannot be 0 is that $\alpha = 0$ will result in $MRT = 0$ and $Vol = 0$ (see Eq 2.11 and Eq 2.10). When $\alpha < 0$, the estimated MRT and Vol are also negative. The term $(\alpha - 1)/t$ in Eq 2.19 is assumed to represent the loss of the tracer into the interstitium. When $\alpha > 1$, $(\alpha - 1)/t$ is positive which means that the kidneys are not solely responsible for eliminating the tracer from the body. Therefore, the result from $\alpha > 1$ would represent the plasma clearance for the whole body instead of the renal clearance. When $\beta < 0$, the exponential term $e^{-\beta t}$ is associated with a negative elimination rate, indicating that the concentration of the tracer is increasing with time (which is non-physical for the plasma clearance study). However, for patients with extremely low GFR value (e.g. kidney failure), their expected elimination rate are very close to 0. It is, therefore, allowed to have $\beta = 0$, which is associated with the condition when the patients have no kidneys. When $\beta = 0$, the MRT and Vol should be close to infinity, indicating that the tracer will take an infinite time to be cleared out of the body.

In order to obtain the concentration $C(t) = Kt^{\alpha-1}e^{-\beta t}$, at least three blood samples are required. A nonlinear regression using an OLS algorithm was used to try to solve the GV model (see Section 3.4). However, the OLS algorithm sometimes has difficulty in finding a solution and occasionally gives non-physical GFR results. Therefore, a more advanced mathematical method is required for solving the GV model, as introduced in the following.

2.6.3 Ill-posed Problem

The nature of the plasma clearance study is ill-posed because the estimation of the plasma clearance is obtained by approximating the concentration-time curve using finite samples within a limited time period.

Ill-posed problems are problems that are highly sensitive to changes (or fluctuations) in the initial input data or the conditions of these data (e.g. herein, sampling schedule) [107]. For example, if the blood collection time for the kidney function test has a small deviation (e.g. one of the sampling times is 5 min away from the supposed or recorded collecting time), the estimated *GFR* value from the compartmental models may differ by a factor of 2 or more.

The plasma clearance studies are also inverse problems. An inverse problem is defined as the process of obtaining the parameters of a model from the original experimental data while a non-inverse (or classical) problem deals with the original experimental data using a known model with known parameters. Inverse problems are usually difficult to solve because: there may be no known model that fits the data well (or appropriately) because the existing model sometimes is insufficient for describing the data from the system, and there may be more than one solution. This occurs especially in models with multiple parameters. Therefore, the problem becomes finding the best solution not only for the current input data but also for the other sets of input data (e.g. from repeated studies) to be fitted in the future. Since the input data usually originate from a finite sample size with unknown noise level, inverse problems are usually ill-posed.

The plasma clearance problem is an ill-posed inverse problem because of the complexities involved during the perfusion, dilution, mixing and elimination procedures of the tracers in the body. Besides the uncertainty intrinsic to the plasma clearance, the

laboratory measurements might add to the complexity, such as the injection method, sampling times, number of samples, etc. The limited sampling schedule (either the length of the sampling in time or the number of the samples being collected) does not provide sufficient information to solve the problem. As a result, during our calculation of the patients' data (see description of the data in Chapter 3), the estimated *GFR* can vary by up to approximately 40 ml/min if there is a slight change or error in any of the above factors.

The first principle for solving ill-posed problems is to introduce more information than what the original input data alone can provide. The supplemental information is usually in the form of penalty functions which reduce the complexities of the problem and restricts the occurrence of non-physical results. The penalty function method attempts to replace a constrained regression (either linear or nonlinear) by a series of unconstrained regressions. The process of using penalty functions for solving ill-posed problems is called regularization. Depending on the characteristics of the ill-posed problem, regularization can be linear or nonlinear, mono-variable or multivariable, etc.

2.6.4 Tikhonov Regularization

Because the *GV* model is ill-posed, it is important to choose a regression method to ensure that a solution can be found and that the solution is physically realistic. The OLS linear regression was initially used for solving the *GV* model, which is named as OLS-*GV* model in this study. However, it is observed that physiologically impossible *GFR* results were occasionally obtained using the OLS-*GV* model, shown in Sections 4.3 and 5.1. This has limited the usage of the *GV* model in the study of plasma clearance. Therefore, there is a need for improving the OLS regression

method and that is why Tikhonov regularization has been chosen, denoted as the **TiK**honov regularization adapted **G**amma **V**ariate (**Tk-GV**) model.

Tikhonov regularization was invented by A. N. Tikhonov and it is one of the most commonly used regularization methods for solving ill-posed problems [108–110]. The original Tikhonov method has been modified by A. E. Hoer and renamed as the ridge regression, which is known as the generalized Tikhonov regularization [111]. The principle of Tikhonov regularization is to apply constrains or penalty functions to the regression algorithm in order to lead the ill-posed problem to a solution. In this study, a damping factor or shrinkage factor λ is used for controlling the minimization procedure even when the initial guess of the solution is far from the final solution. This shrinkage factor λ is unrelated to the symbols λ_1 and λ_2 used in the *SET2* model.

How to apply the Tikhonov regularization to the *GV* model is described below. The *GV* model describes the concentration as $C(t) = Kt^{\alpha-1}e^{-\beta t}$. Taking the logarithm of this equation gives

$$\ln C(t) = \ln K + (\alpha - 1) \ln t - \beta t, \quad (2.23)$$

which is now a linear regression equation of the form:

$$y = d + mx_1 + cx_2, \quad (2.24)$$

where the values obtained from the measurements of the radioactivity of the plasma samples ($\ln C(t)$, $\ln t$, t) are replaced with the variables (y , x_1 , x_2) and the parameters needed to be determined ($\ln K$, $(\alpha - 1)$, $-\beta$) are replaced with (d , m , c) respectively. Because several samples were taken for each patient, the variables y , x_1 , x_2 are in fact single column matrices instead of single numbers. It is therefore more appropriate

to use a matrix notation for the following description. The matrix for the dependent variable $\ln C(t)$ is written as \mathbf{Y} with dimension $n \times 1$. In order to remove the constant term d in Eq 2.24, \mathbf{X} must be written as a $n \times 3$ matrix of the form:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{21} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{pmatrix}.$$

The \mathbf{Y} and \mathbf{b} matrices are then written as follows:

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} d, m, c \end{pmatrix}.$$

Eq 2.24 is then rewritten as

$$\mathbf{Y} = \mathbf{bX}, \tag{2.25}$$

If the residuals of Eq 2.25 are given as $(\mathbf{Y} - \mathbf{bX})$, the OLS approach for solving Eq 2.25 is to minimize the sum of squares of residuals, which is equal to minimizing $(\mathbf{Y} - \mathbf{bX})^T(\mathbf{Y} - \mathbf{bX})$ with respect to the variable \mathbf{b} . In this last equation, the symbol T represents transpose of the vector $(\mathbf{Y} - \mathbf{bX})$. The minimum is obtained by taking the derivative of this equation with respect to \mathbf{b} and setting it equal to zero as follows:

$$\frac{\partial((\mathbf{Y} - \mathbf{bX})^T(\mathbf{Y} - \mathbf{bX}))}{\partial \mathbf{b}} = -2\mathbf{X}^T(\mathbf{Y} - \mathbf{bX}) = 0. \tag{2.26}$$

Rearranging Eq 2.26 and solving for the estimated coefficients gives:

$$\mathbf{b} = \mathbf{b}_{\text{OLS}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}, \tag{2.27}$$

where \mathbf{b}_{OLS} indicates the use of OLS method. The term $\mathbf{X}^T\mathbf{X}$ usually contains collinearity between the variables, which leads to the case where the solution is highly sensitive to very small changes in the values of the variables (known as ill-posed). Details of ill-posed problems are described in Section 2.6.3. The OLS algorithm is insufficient in solving ill-posed problems, which may lead to very poor results [109,111]. Therefore, Tikhonov regularization was used for solving this equation by adding a term containing a shrinkage factor (with value between 0 and 1) giving:

$$\mathbf{b}_{\text{ridge}} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{Y}, \quad (2.28)$$

where \mathbf{I} is the identity matrix. When $\lambda = 0$, $\mathbf{b}_{\text{OLS}} = \mathbf{b}_{\text{ridge}}$.

Solving the linear regression equation Eq 2.25 now becomes a question of choosing the proper shrinkage factor since the coefficient $\mathbf{b}_{\text{ridge}}$ is a function of λ . The characteristics of λ include: 1) it controls the size of the coefficients, 2) it controls the amounts of regularization, 3) when $\lambda \rightarrow 0$, the regression becomes OLS, and 4) when $\lambda \rightarrow \infty$, the estimated $\mathbf{b}_{\text{ridge}}$ is zero [111]. Since the solution is determined using λ , different choices of λ may lead to different solutions. One way to find the shrinkage factor λ is by plotting the estimated coefficient \mathbf{b} as a function of different choices of λ values until a λ is found that gives a stable value of \mathbf{b} . However, this is not the best way to find the shrinkage factor because a slight change (e.g. 10^{-4}) in λ can lead to a 10 ml/min (approximately 10% of the normal *GFR* value) fluctuation in the estimated *GFR* value. It is necessary to have a more accurate formula to determine the best value of λ . In this study, the shrinkage factor λ was found by minimizing the relative error of the estimated *GFR* using the error propagation function, as introduced below.

Since the uncertainty of the estimated *GFR* depends on the uncertainties of its

variables α , β , and K , it is possible to describe the relative error of the estimated *GFR* using the relative errors of these variables. This is the principle of error propagation, which uses a second-order differential of a function to magnify the effect of a parameter or variable on the final result [112]. For example, a function $f(u, v, \dots)$ with variables u, v, \dots has a propagation of errors given by

$$S_f^2 = S_u^2 \left(\frac{\partial f}{\partial u} \right)^2 + S_v^2 \left(\frac{\partial f}{\partial v} \right)^2 + 2S_{uv} \left(\frac{\partial f}{\partial u} \right) \left(\frac{\partial f}{\partial v} \right) + \dots, \quad (2.29)$$

where S_f^2 is the standard deviation of $f(u, v, \dots)$. More details on how to calculate the S_f^2 are given in Appendix A.2. Using the definition above, the standard deviation of *GFR* (S_{GFR}^2) is expressed as,

$$\begin{aligned} S_{GFR}^2 = & S_\alpha^2 \left(\frac{\partial GFR}{\partial \alpha} \right)^2 + S_\beta^2 \left(\frac{\partial GFR}{\partial \beta} \right)^2 + S_K^2 \left(\frac{\partial GFR}{\partial K} \right)^2 \\ & + 2S_{\alpha\beta} \left(\frac{\partial GFR}{\partial \alpha} \right) \left(\frac{\partial GFR}{\partial \beta} \right) + 2S_{\alpha K} \left(\frac{\partial GFR}{\partial \alpha} \right) \left(\frac{\partial GFR}{\partial K} \right) \\ & + 2S_{\beta K} \left(\frac{\partial GFR}{\partial \beta} \right) \left(\frac{\partial GFR}{\partial K} \right). \end{aligned} \quad (2.30)$$

When the *GFR* is calculated using the *GV* model in Eq 2.20, Eq 2.30 becomes:

$$\begin{aligned} S_{GFR}^2 = & S_\alpha^2 \left(\frac{\text{Dose} \cdot \beta^\alpha \ln \beta}{K\Gamma(\alpha)} - \frac{\text{Dose} \cdot \beta^\alpha \Gamma'(\alpha)^2}{K\Gamma(\alpha)} \right)^2 \\ & + S_\beta^2 \left(\frac{\text{Dose} \cdot \alpha \beta^\alpha \ln \beta}{K\Gamma(\alpha)} \right)^2 + S_K^2 \left(-\frac{\text{Dose} \cdot \beta^\alpha}{K^2\Gamma(\alpha)} \right)^2 \\ & + 2S_{\alpha\beta} \left(\frac{\text{Dose} \cdot \beta^\alpha \ln \beta}{K\Gamma(\alpha)} - \frac{\text{Dose} \cdot \beta^\alpha \Gamma'(\alpha)^2}{K\Gamma(\alpha)} \right) \left(\frac{\text{Dose} \cdot \alpha \beta^\alpha \ln \beta}{K\Gamma(\alpha)} \right) \\ & + 2S_{\alpha K} \left(\frac{\text{Dose} \cdot \beta^\alpha \ln \beta}{K\Gamma(\alpha)} - \frac{\text{Dose} \cdot \beta^\alpha \Gamma'(\alpha)^2}{K\Gamma(\alpha)} \right) \left(-\frac{\text{Dose} \cdot \beta^\alpha}{K^2\Gamma(\alpha)} \right) \\ & + 2S_{\beta K} \left(\frac{\text{Dose} \cdot \alpha \beta^\alpha \ln \beta}{K\Gamma(\alpha)} \right) \left(-\frac{\text{Dose} \cdot \beta^\alpha}{K^2\Gamma(\alpha)} \right). \end{aligned} \quad (2.31)$$

The symbol $\Gamma'(\alpha)$ in Eq 2.31 is the derivative of the gamma function $\Gamma(\alpha)$ with respect to α . In this study, the relative error of the estimated *GFR*, labeled as *ERG*, is expressed as the ratio of S_{GFR}^2 to GFR^2 , and is expressed as

$$\begin{aligned}
 ERG = \frac{S_{GFR}^2}{GFR^2} &= S_{\alpha}^2 \left(\ln \beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) + S_{\beta}^2 \left(\frac{\alpha}{\beta} \right)^2 + \frac{S_K^2}{K^2} \\
 &+ \frac{2S_{\alpha\beta}}{\beta} \left(\ln \beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) - \frac{2S_{\alpha K}}{K} \left(\ln \beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) - \frac{2S_{\beta K}}{K\beta}, \quad (2.32)
 \end{aligned}$$

where $\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ is a digamma function denoted as $\psi(\alpha)$. Hence, Eq 2.32 becomes

$$\begin{aligned}
 ERG = \frac{S_{GFR}^2}{GFR^2} &= S_{\alpha}^2 (\ln \beta - \psi(\alpha)) + S_{\beta}^2 \left(\frac{\alpha}{\beta} \right)^2 + \frac{S_K^2}{K^2} \\
 &+ 2 \left(\frac{S_{\alpha\beta}}{\beta} - \frac{2S_{\alpha K}}{K} \right) (\ln \beta - \psi(\alpha)) - 2 \frac{S_{\beta K}}{K\beta}. \quad (2.33)
 \end{aligned}$$

Since the parameters α , β , and K are all functions of λ , *ERG* is also a function of λ . *ERG* is then minimized (using the Nelder Mead or Simulated Annealing method, see Section 3.3) to find the convergence, hence, the best value of λ . For patients with normal or high *GFR*, the scale of λ is very small (e.g. 10^{-4}) while patients with very low *GFR* usually require more regularization for the regression, hence, a larger value of λ (e.g. close to 1) is needed.

ERG can be simplified with known variables before performing the minimization. The first step is to exclude the terms that contribute little (ignorable) to *ERG*. S_K is found to be approximately 10^{-14} times less than either S_{α} or S_{β} , making it possible to truncate the terms containing $S_{\alpha K}$, $S_{\beta K}$, and S_K^2 . However, the truncations might omit important information and introduce a bias in the estimation. Therefore, an alternative way is used by replacing the parameter K with an equation containing the known variables t and $C(t)$ along with α and β as

$K = \exp(\overline{\ln C(t)} - (\alpha - 1)\overline{\ln t} + \beta\bar{t})$, where $\overline{\ln C(t)}$, $\overline{\ln t}$ and \bar{t} indicate the mean value of $\ln C(t)$, $\ln t$ and t respectively. *GFR* is then rewritten from Eq 2.20 as

$$GFR = \frac{\text{Dose} \times \beta^\alpha}{\exp(\overline{\ln C(t)} - (\alpha - 1)\overline{\ln t} + \beta\bar{t})\Gamma(\alpha)}. \quad (2.34)$$

Taking the derivative with regard to parameter α and β according to the definition of error propagation in Eq 2.29, *ERG* is then rewritten from Eq 2.33 as

$$\begin{aligned} ERG &= \frac{S_{GFR}^2}{GFR^2} \\ &= S_\alpha^2 (\overline{\ln t} + \ln \beta - \psi(\alpha)) + S_\beta^2 \left(\frac{\alpha}{\beta} - \bar{t} \right)^2 \\ &\quad + 2 \frac{S_{\alpha\beta}}{\beta} (\overline{\ln t} + \ln \beta - \psi(\alpha)) \left(\frac{\alpha}{\beta} - \bar{t} \right). \end{aligned} \quad (2.35)$$

ERG is used to calculate the shrinkage factor λ and hence the *GFR* and *Vol* values. However, during the results and discussion sections, only the *GFR* and *Vol* values are presented. An example on how to use the expressions for α , β , $\ln K$, *ERG* and *ERV* as functions of λ are shown in Appendix A.2. The numerical methods used to minimize *ERG* are introduced in the next chapter.

Chapter 3

Data and Analysis Methods

This chapter will discuss the data and methods used in the *GFR* project while Chapter 9 will give details of the data and methods used in the BMD project. The numerical methods used to investigate the pharmacokinetic models are introduced. The statistical analysis used to compare the performances of different models is also presented.

3.1 Data

The study of plasma clearance is usually performed using the two-blood sample *SET1* method [113]. By 2008, there were 42 *GFR* patients' data analyzed using the two blood sample *SET1* method at the General Hospital in St. John's, NL. There is no database from the General Hospital with 4 or more blood samples. Over and above the limited sample size (two samples per patient), the *GFR* measurements obtained at the General Hospital from 2004 to 2008 were not performed under appropriate laboratory quality control. For example, the multi-well gamma counter was used without routine calibration or quality control. The minimum requirement for assuring

quality control of the gamma counter includes checking the background count rate, testing the energy resolution and energy response of the counter (by viewing the output spectrum), testing the sensitivity using a calibration material such as Cesium-137 or Cobalt-57, and performing a linearity test for assessing the counting range of the radioactivity. Calibration of the counter should be performed at least yearly. If there are concerns during the usage of the counter, a thorough calibration should be performed to avoid possible loss of precision in the measurements.

In the following Chapters 4 and 5, an analysis of a new *GFR* model will be presented based on a group of 46 patients. Because these 46 patients' data were provided by others, there was no need for ethics approval for using these data according to the human investigation committee at Memorial University of Newfoundland. Of the 46 *GFR* patients, 41 patients' data were collected and provided by Dr. Charles Russell from the University of Alabama in the United States. Each of these 41 patients was injected with ^{99m}Tc -DTPA [71]. Eight blood samples were drawn with standard anticoagulated tubes at 10, 20, 30, 45, 60, 120, 180, and 240 min after the injection from the opposite arm where the initial injection was given. A standard solution using the same radiopharmaceutical was prepared and diluted to scale the radioactivity of the plasma measured using the well counter. The remaining 5 of the 46 patients' data were collected and provided by Dr. Barbara Y. Croft formally from the National Cancer Institute of the United States (present location Duke University). Each of these 5 patients had 9 blood samples drawn at 5, 10, 15, 20, 60, 70, 80, 90 and 180 min after the injection of ^{99m}Tc -DTPA. These 46 patients were the primary dataset used for the *GFR* part of this thesis for investigating the advantages of the new Tk-GV model compared with the *SETs* models. Despite the fact that early samples, such as 5 min or 10 min, were included in the patients' results, no concentration-time curve

from the obtained data indicated an observed perfusion stage.

3.2 Measurement Uncertainty

There are many factors that contribute to the uncertainty of estimated *GFR* values which are calculated from quantities obtained from clinical procedures. The purity of the radiopharmaceutical ^{99m}Tc -DTPA can affect the uncertainty of the value of the injected dose administered to patients [114]. The radioactivity of the syringe and gloves used during the injection of the radiopharmaceutical are measured and subtracted from the injected dose which introduces an uncertainty. The uncertainty due to the dose calibration using a dose calibrator can be as high as 5% [115]. Although the sampling time can be recorded to the fraction of a second using a stop watch, the actually blood collection can last from seconds to minutes depending on the patients' condition. From our own experience in the department of Nuclear Medicine at the General Hospital, we know that for patients with ideal vein conditions, the collection can be done within 5 to 10 seconds. However if there is difficulty in drawing blood from patients such as children or cancer patients, it can take much longer. Given the half-life of ^{99m}Tc -DTPA of approximately 6 hours, a 1 min difference in the blood extraction time introduces an error of approximately 0.3%. The protein binding rate of ^{99m}Tc -DTPA can also induce an error in the estimated *GFR* values. Protein binding effect raises more concern for patients with diabetes or for patients consuming an abnormally high-protein diet such as athletes [116]. Pipetting is used to extract a fix volume of plasma for performing counting measurements. An air displacement pipette can induce an uncertainty approximately less than or equal to 6% of the volume depending on the temperature, humidity and handling of the instrument [117]. The single-well counter can also introduce a measurement uncertainty due to the po-

sition of the sample in the well, the background noise and the detector efficiency. In general, the uncertainty during the radioactivity counting using the gamma counter is approximately the square root of the counts [1, 20]. The uncertainty due to the mathematical modelling and calculations also contribute to the uncertainty of the estimated *GFR* value. For calculations that require patient information such as body mass and body height, uncertainties can be introduced by factors such as what the patient is wearing, food intake before the test and so on. For pediatric patients, the uncertainty of the body mass and body height can also be caused by the positioning of the patient.

For this study, where data (the 46 patients from Dr. Russell and Dr. Croft) were provided by others, it is difficult to determine the specific error for each of the factors listed above, a general pipetting error of 3% is used for the measurements [118].

3.3 Numerical Methods

The software used in this study was Mathematica 6.0.2. Mathematica is a powerful software package used for performing complex mathematical calculations. The software is written in C language and its built-in functions are comparable with those of other numerical software packages such as Matlab and SPSS. The gamma variate function, the digamma function and all the minimization and regression algorithms used in this work are available in Mathematica. The calculation time for all the models using Mathematica was on the order of seconds.

The parameters α , β , K , *ERG* and *ERV* are all expressed as functions of the shrinkage factor λ using the Tikhonov regularization (see Section 2.6.4). After the minimization of *ERG* with respect to λ , the obtained value of λ was used to calculate the *GFR* and *Vol* values.

Several methods were tested for finding the global minimum of ERG such as Simulated Annealing algorithm [119, 120], the Nelder Mead method [121], the Differential Evolution method, the random search method, and the maximum likelihood estimation method [122, 123]. Among all these methods, the Nelder Mead algorithm gave the smallest residual sum of squares value and therefore was chosen as the best approach for finding the global minimum.

3.4 Regression Methods

Four models for calculating GFR were tested in this study: $SET1$, $SET2$, OLS-GV and Tk-GV models. For the $SET1$, $SET2$ and OLS-GV models, an ordinary least squares (OLS) method was used to find the solutions. For the Tk-GV model, the Tikhonov regularization was used.

The OLS method is the most commonly used regression algorithm. The OLS method is applied to both linear and nonlinear regressions. Depending on the format of the independent variables, the OLS linear regression was used for solving the $SET1$ model with two data points of the form $(t, \ln C(t))$.

The OLS nonlinear regression was also used for solving the $SET2$ and OLS-GV models using 3 and 4 data points of the form $(t, C(t))$. As will be discussed in greater detail in Chapter 5, the OLS regression used for the OLS-GV model occasionally gave non-physical results. The reason is the limitation of the OLS algorithm in converging during the regression. The parameters of the curve to be fitted were obtained from the minimization of the sum of squares of the residuals, instead of minimizing the absolute residuals. The assumptions required for using the OLS algorithm could not always be satisfied in this study. For example, one assumption for using the OLS method required that there be no collinearity among the independent variables.

The variables used in the OLS-GV model are of the form $(t, \ln t, \ln C(t))$ where the variables t and $\ln t$ are correlated. Because the plasma clearance study in this work does not satisfy all the assumptions of the OLS algorithm, the ridge regression method was used to solve the *GV* model. The Tk-GV method was performed in the way that the shrinkage factor λ was found by minimizing the relative error of the *GFR* values, which in return minimizes the difference between the observed and predicted values of concentration of the radiopharmaceutical as well.

The coefficient of determination (R^2) was used to describe the performance of the linear regression [124]. R^2 is defined as

$$R^2 = 1 - \frac{\sum (y_{\text{est}} - \bar{y})^2}{\sum (y - \bar{y})^2}, \quad (3.1)$$

where y_{est} is the estimated value and \bar{y} is the mean value of the observed y values.

Regression was used not only for finding the shrinkage factor, but also for comparing different methods. If two methods provided the same or similar test results, their R^2 values should be high (e.g. 0.99). However, R^2 alone cannot detect whether or not there is a constant or proportional difference between the methods. During the *GFR* project, R^2 values were used for comparing the performance of difference models while in the BMD study, R^2 was used for showing the performance of specific distribution functions in fitting the histogram.

Chapter 4

Results

In this chapter, the results of the *SET1*, *SET2*, OLS-GV and Tk-GV models from the 46 patients are presented. The calculated results include the *GFR* and *Vol* values as well as the estimated parameter values of each model. When we refer to the full samples, it implies that the data from all 46 patients were used in the calculations. Subsets consist only of data from the 41 patients provided from Dr. Charles Russell because of their uniform sampling times. During this Chapter, some minor discussions are included along with the results of each model; however, the major discussion regarding all of the models is given in Chapter 5.

4.1 *SET1* Results

The *SET1* model is expressed as $C(t) = Ae^{-\alpha t}$ with two estimated parameters A (scaling coefficient) and α (elimination factor). The *GFR* calculations were performed using a linear regression on the logarithm of the concentration $\ln C(t)$ as a function of time t . Table 4.1 lists the mean, minimum and maximum values of the estimated *GFR* and *Vol* values for the 46 patients from the full samples and two subsets. Two-sample

subsets were chosen because the minimum requirement of the number of samples for using the *SET1* model is 2. In many institutions, it has become common practice to calculate *GFR* from the plasma clearance method using the *SET1* model with 2 blood samples [71, 125]. There are 28 different combinations of two-sample subsets from patients with 8 samples each ($\binom{8}{2} = 28$) and 36 different combinations from patients with 9 samples ($\binom{9}{2} = 36$). Table 4.1 only lists the results using the (60, 180)min and (10, 180)min subsets because these two subsets were contained within all 46 patients, and (60, 180)min has been recommended as good sampling times for using the *SET1* model [71]. The calculations using the 2-sample subsets were based on the absolute solution (see Eq 2.15).

Table 4.1: *GFR* (ml/min) and *Vol* (L) values calculated from the *SET1* model for the 46 patients using the full samples (denoted as “full”) and subsets: (60, 180)min and (10, 180)min.

<i>GFR</i>	full	(60,180)	(10,180)	<i>Vol</i>	full	(60,180)	(10,180)
min	4.5	53.3	66.5	min	6.0	7.8	5.9
mean	103.0	326.0	189.9	mean	25.8	15.9	11.4
max	208.4	3041.9	1209.0	max	242.8	27.4	23.2

In Table 4.1, the estimated *GFR* values from the full samples ranged from 4.5 ml/min to 208.4 ml/min. There were 6 out of 46 patients with estimated *GFR* values less than 30 ml/min. There were, however, no negative *GFR* values observed. Although the time interval between the two sample collection times (120 min for the (60, 180)min subset and 170 min for the (10, 180)min subset) were relatively long, the 2-sample subsets significantly increased the minimum *GFR* value from 4.5 ml/min to larger than 50 ml/min. The maximum *GFR* values calculated from both subsets were beyond the expected range of normal *GFR* reported as between 100 and 130 ml/min/1.73 m² [66]. The percentage of *GFR* values larger than 130 ml/min was

approximately 40% indicating there were at least 16 patients from 46 with potentially overestimated GFR values.

There are two reasons why GFR values may be larger than 130 ml/min. For estimated GFR values between 130 and 200 ml/min, one common explanation is that the patients have hyperfiltration in their kidneys which is a phenomena common among patients with diabetes [116,126]. Other causes for hyperfiltration include the consumption of a high protein diet and other underlying kidney diseases that alter the glomerular pressure [116,127]. Unfortunately, there was no additional information provided such as family history for the 46 patients from Dr. Russell making it difficult to determine the causes of abnormally high GFR values. It is observed that hyperfiltration can rarely lead to GFR values exceeding 200 ml/min [116,126–128]. Therefore, GFR values larger than 200 ml/min could be considered as physiologically impossible. However, the GFR values discussed in the literature were calculated using either the creatinine method or the $SETs$ models with limited sample size, and are not reliable. Therefore, it is difficult to set an upper limit on the estimated GFR and hence for the purpose of this analysis only negative GFR values were defined as non-physical.

The mean GFR values from the two subsets in Table 4.1 were approximately 1.8 to 3.2 times larger than those from the full samples indicating the degree of overestimation obtained from using $SET1$ with a limited sample size. In addition, there were variations observed in the results from using the different 2-sample subsets. The (60, 180)min sampling schedule has been chosen as a reference sampling schedule by most health institutions since 1985 [71,125]. However, when compared with the full samples, the (10, 180)min subset was found to perform better than the (60, 180)min subset in estimating the GFR values because of the longer time interval (170 min

versus 120 min) between the samples. When the *SET1* model was used to fit the concentration-time curve from two samples, the closer the time difference between the two collected samples, the more likely the *SET1* model risked being affected by unexpected fluid redistribution or slow mixing occurring within a short time interval. As a result, the *SET1* model applied during a non-elimination process might misrepresent the concentration time curve. In an extreme case such as redistribution, the mono-exponential fitting might lead to negative elimination.

When patients have kidney diseases, their *Vol* values are expected to be higher than normal indicating a longer time is required for eliminating the tracer from the body. For example, the *Vol* value for Patient 19 (Pt19) was 242.8 L and the corresponding *GFR* value of Pt19 was 4.5 ml/min. In Table 4.1, the relative fluctuation of the *Vol* values from the full samples and the two subsets was less than that of the *GFR* values. The mean *Vol* value from the full samples was 25.8 L with a range from 6.0 L to 242.8 L, and the mean *Vol* value from the (60, 180)min subset was approximately 10 L less than that from the full samples. For the same reason that estimated *GFR* values larger than 200 ml/min/1.73 m² were not considered as non-physical, the upper limit of the estimated *Vol* values was not discussed either, and hence only negative *Vol* values were considered non-physical in this study.

Table 4.1 listed the results from only two 2-sample subsets. It is useful to investigate how the overall choice of the 2-sample subsets affects the performance of the *SET1* model. Table 4.2 gives a summary of the average value of the *GFR*, indicated as \overline{GFR} , and the average value of the *Vol*, indicated as \overline{Vol} , from all possible $41 \times \binom{8}{2} = 1,148$ subsets chosen from the 41 patients in order to keep the choice of sampling times consistent. The data in Table 4.2 is presented in quartiles. Quartiles are used to divide data (displayed in an increasing order) into 4 equal groups. The

1st quartile is the cut-off value for the lowest 25% of the data. The 2nd quartile is the cut-off value for the lowest 50% (equal to the median value) while the 3rd quartile is the cut-off value for the lowest 75% of the data [97]. In the following, quartiles will be presented for all models with different sampling schedules to demonstrate the range of estimated GFR values and to determine whether or not there are outliers during the curve fitting.

Table 4.2: The effects of the 1,148 2-sample subsets on the values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min) and \overline{Vol} (L) using the $SET1$ model. \overline{GFR} (ml/min) is the mean GFR , and SD_{GFR} (ml/min) is the standard deviation of the GFR values. \overline{Vol} (L) is the mean Vol , and SD_{Vol} (L) is the standard deviation of the Vol values.

Quartile	\bar{A}	SD_A	$\bar{\alpha}$	SD_α	\overline{GFR}	SD_{GFR}	\overline{Vol}	SD_{Vol}
min	0.004	0.0002	0.003	0.002	52.1	5.8	7.0	0.9
1st quartile	0.007	0.001	0.007	0.003	70.1	13.0	11.4	1.9
2nd quartile	0.008	0.001	0.010	0.004	94.2	18.6	13.9	2.8
3rd quartile	0.009	0.002	0.011	0.006	136.3	31.5	15.5	3.5
max	0.014	0.005	0.014	0.012	883.2	1101.4	25.3	6.6

The first 4 columns in Table 4.2 list the minimum, maximum and quartile values of the mean and standard deviations of the parameters A and α from $C(t) = Ae^{-\alpha t}$. The SD_A values were less than the SD_α values indicating that the sensitivity of the constant elimination rate α was greater than that of the scaling coefficient A . The dispersion of the estimated \overline{GFR} and \overline{Vol} values using the overall 1,148 2-sample subsets was better than that from the subsets (60, 180)min or (10, 180)min alone. For example, the maximum \overline{GFR} values were 883.2 ml/min from the 1,148 2-sample subset (Table 4.2), 3041.9 ml/min from the (60, 180)min and 1209.0 ml/min from the (10, 180)min (Table 4.1). However, the maximum \overline{GFR} , 883.2 ml/min, from the 1,148 subsets is still beyond the expected range of GFR values. The range of

estimated SD_{GFR} is larger than that of the estimated SD_{Vol} values.

Table 4.3: The effects of the 3-sample subsets (2,296 for all 41 patients) on the values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min), and \overline{Vol} (L) using *SET1*. \overline{GFR} (ml/min) is the mean *GFR* and SD_{GFR} (ml/min) is the standard deviation of the *GFR* values. \overline{Vol} (L) is the mean *Vol* and SD_{Vol} (L) is the standard deviation of the *Vol* values.

Quartile	\bar{A}	$\bar{\alpha}$	\overline{GFR}	SD_{GFR}	\overline{Vol}	SD_{Vol}
min	0.71	0.10	6.3	3.0	4.5	2.8
1st quartile	19.24	0.11	62.2	28.2	6.2	3.9
2nd quartile	23.86	0.14	86.3	51.6	8.2	5.3
3rd quartile	34.82	0.19	119.2	67.6	12.2	8.8
max	64.10	0.21	203.4	129.0	153.0	295.1

Table 4.3 and Table 4.4 list the quartile values of the mean estimated parameters and mean *GFR* and *Vol* values using 3-sample subsets and 4-sample subsets from the 41 patients, respectively. There were 2,296 subsets generated by taking 3 samples out of 8 samples from the 41 patients and 2,870 different 4-sample subsets. The estimated \overline{GFR} values ranged from 6.3 ml/min to 203.4 ml/min using the 3-sample subsets (Table 4.3) and 5.8 ml/min to 169.2 ml/min using the 4-sample subsets (Table 4.4). The estimated \overline{Vol} values ranged from 4.5 L to 153.0 L using the 3-sample subsets (Table 4.3) and 3.9 L to 220.6 L using the 4-sample subsets (Table 4.4). Notice here that the maximum \overline{GFR} values estimated using 3-sample or 4-sample subsets were less than that obtained using the 2-sample subsets, due to the increased number of samples which improved the performance of the *SET1* model.

The 4-sample subset is the sample size of interest in this study because it meets the minimum requirement for using the *SET2* and Tk-GV models and it can also be used by the *SET1* and OLS-GV models. In order to compare the results to other models, Table 4.5 lists the *GFR* values from the 41 patients using four different

Table 4.4: The effects of the 4-sample subsets (2,870 for all 41 patients) on values of \bar{A} , $\bar{\alpha}$, \overline{GFR} (ml/min) and \overline{Vol} (L) using *SET1*. \overline{GFR} (ml/min) is the mean *GFR* and SD_{GFR} (ml/min) is the standard deviation of the *GFR* values. \overline{Vol} (L) is the mean *Vol* and SD_{Vol} (L) is the standard deviation of the *Vol* values.

Quartile	\bar{A}	$\bar{\alpha}$	\overline{GFR}	SD_{GFR}	\overline{Vol}	SD_{Vol}
min	20.05	0.17	5.8	5.47	3.9	3.4
1st quartile	55.19	0.18	51.8	38.7	5.2	4.8
2nd quartile	70.52	0.21	76.7	62.2	6.9	5.8
3rd quartile	84.78	0.25	99.2	82.9	10.2	9.6
max	140.79	0.27	169.2	146.5	220.6	927.8

Table 4.5: The *GFR* results of the *SET1* model for the 41 patients using the full 8 samples and four subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.

<i>GFR</i>	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	4.5	7.2	7.8	34.4	0.011
1st Q	68.3	69.6	50.8	109.6	0.017
2nd Q	100.9	104.7	95.6	146.3	0.021
3rd Q	146.6	148.0	182.7	201.1	0.026
max	208.3	213.3	483.6	271.4	0.044

subsets and the full samples. These four 4-sample subsets were chosen because they were representative for the use of earliest four samples, latest four samples and the two subsets that used a mixture of early and late samples. The GFR values obtained using the (60, 120, 180, 240)min subset were significantly lower than those obtained from the others. For example, a minimum GFR of 0.01 ml/min was observed using the (60, 120, 180, 240)min subset compared with 4.5 ml/min using the full samples. On the other hand, the minimum GFR value obtained using the (10, 20, 30, 45)min subset was much higher than that obtained from others because the $SET1$ model has the tendency of overestimating the GFR (compared with GFR_{full}) using early samples, while underestimating the GFR (compared with GFR_{full}) using late samples [68]. However, the maximum estimated GFR , 483.6 ml/min was obtained using the (10, 30, 120, 240) subset instead of from the (10, 20, 30, 45) subset because of outliers in the data as will be shown in Figure 5.1 c) of Chapter 5.

Another type of subset used in this study was hump subsets, which were not generated from a data resampling method but instead generated by continuously excluding data points from the original dataset. For example, for a dataset consisting of 8 data points, the hump subsets can be *1 to 4* (the first 4 data points), *1 to 5*, and *1 to 8* or *2 to 8*, *3 to 8*, *4 to 8*, and *5 to 8*, etc. Using hump subsets can help investigate the stability of the models with respect to the sampling schedule [118].

For 41 patients with 8 samples each, the choices of hump subsets for performing a linear regression were *1 to 3*, ..., *1 to 8*, *2 to 8*, ... and *6 to 8*. Figure 4.1 shows the mean estimated error in GFR given by $(\overline{GFR_{hump}} - GFR_{full})$ for each hump subset. When the earliest three samples were used (*1 to 3*), the mean GFR difference was a maximum (74.9 ml/min). This overestimation is reduced when more samples were used, such as *1 to 4*, *1 to 5*, up to *1 to 7*. On the other hand, when only the last three

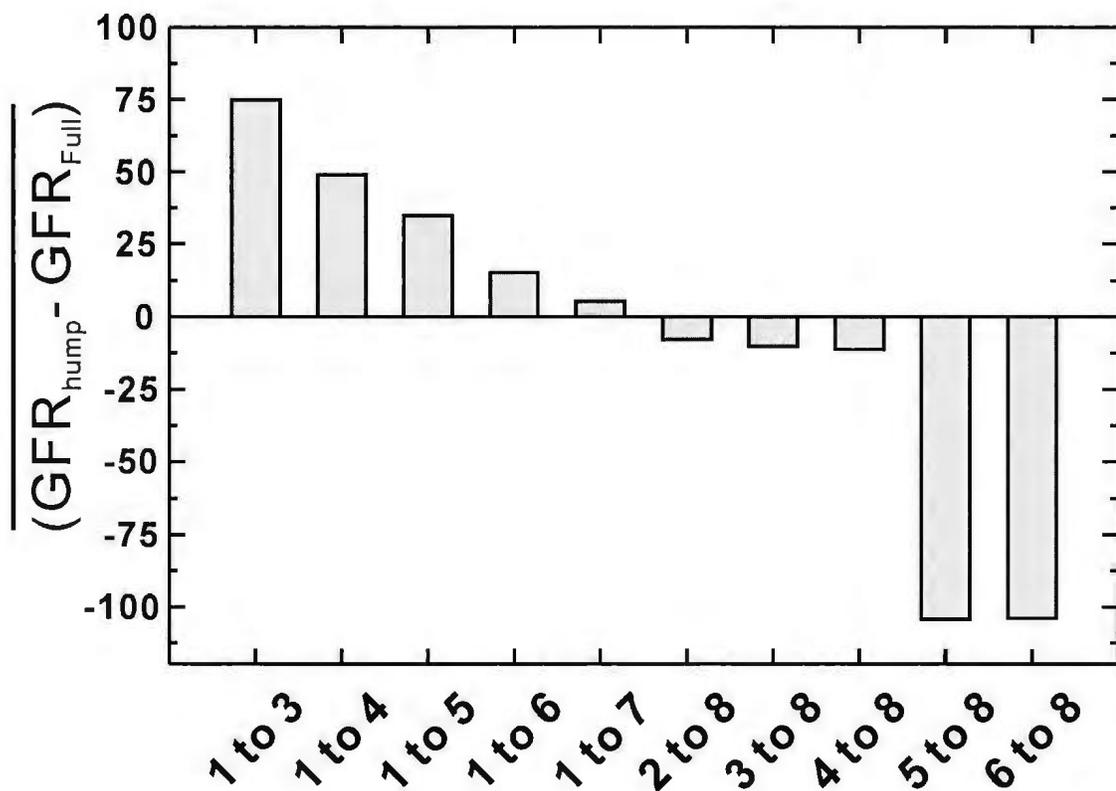


Figure 4.1: Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the *SET1* model from the 41 patients. The hump subsets are *1 to 3*, *1 to 4*, ..., *5 to 8* and *6 to 8*.

samples were used, the mean difference ($\overline{GFR_{hump} - GFR_{full}}$) became negative, -104.0 ml/min. As stated previously, this indicates that the *SET1* model tends to underestimate the *GFR* value when late samples are used. Even when the *2 to 8* subset (with sampling time from 20 min to 240 min) was used, the *SET1* model still gave a mean difference ($\overline{GFR_{hump} - GFR_{full}} = -7.8$ ml/min. In addition, the starting sampling time for the *SET1* model also affects the performance of the model. It is of interest to note that the *5 to 8* subset gave a larger difference than the *4 to 8* subset indicating the necessity of including samples collected before 1 h post-injection. In summary, the *SET1* model is found to overestimate the *GFR* value when early samples are used and underestimate the *GFR* value when late samples are used.

As a specific example, the *GFR* values for Pt1 were 260.1, 204.6, 180.7, 146.4, 111.5 ml/min using the subsets *1 to 3*, *1 to 4*, *1 to 5*, *1 to 6*, *1 to 7* and *1 to 8* respectively. The *GFR* values for Pt1 were 99.4, 95.2, 94.1, 0.02 and 0.3 ml/min for the later subsets *2 to 8*, *3 to 8*, *4 to 8*, *5 to 8* and *6 to 8* respectively. The largest *GFR* value was given using the *1 to 3* hump subset while the smallest *GFR* value was obtained using *6 to 8*. The estimated volume of distribution *Vol* from these hump subsets ranged from 0.3 L to 20.3 L. The fluctuations in the estimated *GFR* and *Vol* values were much larger when the *GFR* was expected to be very low (e.g. less than 5 ml/min). Table 4.1 and Table 4.4 clearly show the sensitivity of the *SET1* model to the number of blood samples and sampling times.

4.2 *SET2* Results

The *SET2* model was expected to improve the concentration-time curve fitting from the *SET1* model [129]. The *SET2* model is expressed as $C(t) = K(ae^{-\lambda_1 t} + e^{-\lambda_2 t})$

with parameters K , a , λ_1 , and λ_2 . Two constraints were included in the nonlinear regressions: $0 \leq a \leq 5$ and $0 \leq \lambda_2 \leq \lambda_1 < 2$. These constraints ensured that the two exponential terms had different elimination rates (therefore $SET2 \neq SET1$) with one contributing to a slower elimination process and the other one contributing to a faster elimination process.

Table 4.6: Estimated GFR (ml/min) and Vol (L) values of the $SET2$ model for the 46 patients using the full samples and the (10, 20, 60, 180)min subset.

	GFR full	(10,20,60,180)	Vol full	(10,20,60,180)
min	2.6	0.9	min 0.5	4.5
mean	80.4	82.8	mean 14.1	14.3
max	166.5	165.4	max 26.5	28.8

Table 4.6 shows the results of the $SET2$ model using the full samples (8 samples for the 41 patients and 9 samples for the 5 patients) and the (10, 20, 60, 180)min subset. The (10, 20, 60, 180)min subset is the only 4-sample subset contained within all 46 patients. All estimated GFR and Vol values were within physically acceptable ranges. The minimum GFR value obtained from the 46 patients with the full samples was 2.6 ml/min and the maximum Vol value was 26.5 L. The results from the (10, 20, 60, 180)min subset were very close to those from the full samples. The minimum GFR value using the (10, 20, 60, 180)min subset was 0.9 ml/min and the maximum estimated Vol value was 28.8 L. The mean value of GFR using the (10, 20, 60, 180)min subset was slightly larger (2.4 ml/min) than that using the full samples. The R^2 values from the regression analysis using the $SET2$ model (between 0.98 and 0.99) were significantly improved from those using the $SET1$ model.

Table 4.7 further investigates the effects of the subsets using the $SET2$ model, calculations were performed using the 4-sample subsets (10, 20, 60, 180)min, (10, 30,

Table 4.7: Estimated GFR (ml/min) calculated with the $SET2$ model using the 4-sample subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min from the 41 patients. Q stands for quartile.

GFR	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	2.6	0.9	-6.6	-474.2	0.008
1st Q	60.2	59.3	27.5	44.2	0.011
2nd Q	81.2	87.6	54.1	95.4	0.014
3rd Q	110.6	114.5	86.0	151.1	0.017
max	166.5	165.4	343.0	3838.2	0.032

120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min from the 41 patients. The (10, 20, 60, 180)min subset gave the closest results to those from the full samples. The (10, 30, 120, 240)min subset is the second best subset that gave close results, although the minimum GFR value using this subset was negative, indicating a non-physical result. The (10, 20, 30, 45)min subset gave the worst estimations of GFR results with non-physical minimum GFR and the maximum GFR value, 3,838.2 ml/min, way beyond the normal range of GFR values. The (60, 120, 180, 240)min subset still underestimated all the GFR values using the $SET2$ model, similar to that using the $SET1$ model in Table 4.5. This is probably due to the inaccuracy of the $SET2$ model in fitting late samples with a relatively flat concentration-time curve. In summary, the (10, 20, 60, 180)min subset gave the best results in estimating GFR values compared to those using the full samples. Also there were more non-physical results observed using the $SET2$ model than with the $SET1$ model.

In order to investigate how the $SET2$ model was affected by the choice of sampling time, we expand its use to a larger number of 4-sample subsets. There are $\binom{8}{4} = \frac{8!}{4! \times (8-4)!} = 70$ different combinations for taking 4 samples out of 8. The results

Table 4.8: The effects of 4-sample subsets (2,870 subsets) using the *SET2* model on values of $\bar{\lambda}_1$ and $\bar{\lambda}_2$, \overline{GFR} (ml/min) and \overline{Vol} (L) from the 41 patients. SD_{GFR} (ml/min) is the standard deviation of the *GFR* values. SD_{Vol} (L) is the standard deviation of the *Vol* values.

Quartile	$\bar{\lambda}_1$	SD_{λ_1}	$\bar{\lambda}_2$	SD_{λ_2}	\overline{GFR}	SD_{GFR}	\overline{Vol}	SD_{Vol}
min	0.19	0.23	0.18	0.24	1.9	18.8	3.3	3.6
1st quartile	0.24	0.26	0.23	0.26	41.3	39.4	7.8	8.6
2nd quartile	0.28	0.27	0.27	0.27	57.1	60.1	9.8	11.0
3rd quartile	0.30	0.29	0.29	0.29	77.3	81.2	15.6	21.4
max	0.33	0.31	0.31	0.31	139.1	455.0	458.5	517.3

shown in Table 4.8 were obtained using a total number of $41 \times \binom{8}{4} = 41 \times 70 = 2,870$ subsets. The mean and standard deviation values of the two elimination rates λ_1 and λ_2 were very close to each other. However, this does not mean that the two elimination rates are approximately the same because the mean of the λ_1 or λ_2 does not represent the ratio of λ_1 to λ_2 for each patient. The average ratio of λ_1/λ_2 was approximately 7, indicating that one compartment in the two-compartment model, on average, had a faster elimination rate than the other. The estimated \overline{GFR} and \overline{Vol} values contained no non-physical results using the *SET2* model due to the averaging of all the results of 70 subsets for each of the 41 patients, however, the estimated SD_{GFR} value, 455.0 ml/min, indicates the unstableness of the *SET2* model with respect to the choice of subsets.

Figure 4.2 shows the mean difference of *GFR* estimated from the hump subsets and the full samples. The smallest mean difference ($\overline{GFR_{hump} - GFR_{full}}$) was obtained using the 1 to 7 hump subset while the largest ($\overline{GFR_{hump} - GFR_{full}}$) was obtained using the 3 to 8 hump subset. However, unlike the *SET1* model, there was no significant trend indicating that the *SET2* model will over- or under-estimate the

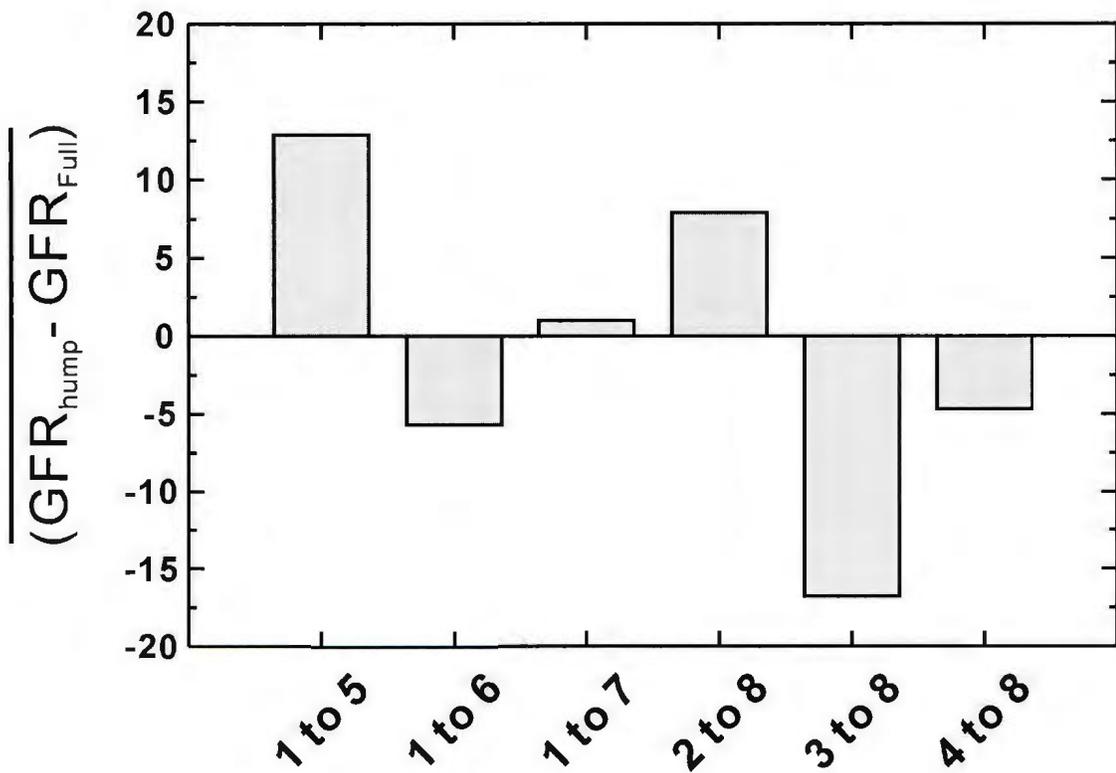


Figure 4.2: Results of the mean of the difference $(GFR_{hump} - GFR_{full})$ between the GFR_{hump} and GFR_{full} using the *SET2* model from the 41 patients. The hump subsets are *1 to 5*, *1 to 6*, *1 to 7*, *2 to 8*, *3 to 8* and *4 to 8*.

GFR results using the early or late samples. Despite that, the overall curve fitting using the $SET2$ model has been improved from the $SET1$ model because the scale of the $\overline{GFR_{hump} - GFR_{full}}$ was much smaller using the $SET2$ model than that the $SET1$ model. However, the $SET2$ model still gave non-physical results when different 4-sample subsets were used.

4.3 OLS-GV Results

Although the minimum sample size required to solve the OLS-GV model is 3, results of the OLS-GV model are presented using 4-sample subsets. The principle reason for this is that the results of the OLS-GV model using 3-sample subsets contained a large number of non-physical results because the OLS regression cannot be used to solve ill-posed problems. Also, the results of the OLS-GV model using 4-sample subsets were used here predominantly to compare with those of the Tk-GV and the $SET2$ models using 4-sample subsets. Although the OLS-GV model was used with 4 samples, the model still provided non-physical results, however, much less than when using 3-sample subsets.

Table 4.9 shows the quartiles and average results of the three estimated parameters \overline{K} , $\overline{\alpha}$ and $\overline{\beta}$, and the values of \overline{GFR} and \overline{Vol} using 2,870 4-sample subsets from the 41 patients. The estimated \overline{GFR} values using the OLS-GV model ranged from 8.3 to 89.3 ml/min and the estimated \overline{Vol} values ranged from 6.5 to 12.4 L, which was a much smaller range than those using the $SET1$ or $SET2$ models.

In order to compare the results from the 4-sample subsets using the OLS-GV model with those from the other models, Table 4.10 shows the quartile results of estimated GFR and Vol values from subsets ((10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min, and (60, 120, 180, 240)min) from the 41 patients. The (10, 20, 60,

Table 4.9: Quartiles of the mean estimated parameters \bar{K} , $\bar{\alpha}$, and $\bar{\beta}$, and the values of \overline{GFR} (ml/min), and \overline{Vol} (L) for the OLS-GV model from the 41 patients using the 2,870 4-sample subsets. \overline{GFR} is the mean GFR and SD_{GFR} is the standard deviation of the GFR values. \overline{Vol} (L) is the mean Vol and SD_{Vol} (L) is the standard deviation of the Vol values.

Quartile	\bar{K}	$\bar{\alpha}$	$\bar{\beta}$	\overline{GFR}	SD_{GFR}	\overline{Vol}	SD_{Vol}
min	35348	0.28	0.26	8.3	8.3	6.5	3.3
1st quartile	64592	0.79	0.32	30.8	32.3	7.6	4.0
2nd quartile	84513	0.85	0.34	43.8	45.8	8.2	4.6
3rd quartile	112758	0.89	0.40	61.8	63.1	9.3	5.7
max	415197	1.09	0.45	89.3	164.0	12.4	42.5

Table 4.10: The GFR and Vol results of the OLS-GV model for the 41 patients using the full 8 samples and four chosen subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.

GFR	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	1.4	0.2	0.4	-410.1	4.6E-09
1st Q	49.8	53.5	51.2	35.9	2.0E-08
2nd Q	78.4	82.8	75.3	81.9	2.4E-08
3rd Q	112.5	113.8	124.5	128.1	3.4E-08
max	153.0	148.2	382.0	291.9	6.3E-08
Vol	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	10.4	10.4	6.0	-126.5	1.67
1st Q	12.0	12.1	7.5	11.7	1.74
2nd Q	13.5	13.2	9.9	13.8	1.75
3rd Q	15.5	15.6	12.7	15.7	1.77
max	23.8	33.9	25.1	44.1	1.79

180)min subset consistently gave the best agreement with the results using the full samples. The estimated GFR values using the (10, 20, 60, 180)min subset ranged from 0.2 ml/min to 148.2 ml/min while the range using the full samples was from 1.4 ml/min to 153.0 ml/min. The results using the (10, 30, 120, 240)min subset showed large differences when compared to those from the full samples with a maximum GFR value of 382.0 ml/min which is well beyond the expected GFR range. The (10, 20, 30, 45)min subset gave non-physical results with a minimum GFR value of -410.1 ml/min and a minimum Vol value of -126.5 ml/min. The corresponding estimated parameter α using the (10, 20, 30, 45)min subset ranged from -0.484 to 1.091 which was beyond its expected range ($0 < \alpha \leq 1$). The OLS-GV model did not perform well using the earliest samples (such as the (10, 20, 30, 45)min subset). Although early samples can contain more noise due to factors such as redistribution or delayed perfusion, the early subsets do not provide sufficient information for the OLS-GV model to properly represent the clearance curve which was consistent with both the $SET1$ and $SET2$ models. The OLS-GV model did not perform well using the latest sample times either. For example, the (60, 120, 180, 240)min subset significantly underestimated the GFR and Vol values (also giving non-physical results) compared with those from the full samples. This is probably due to the limitation of the OLS algorithm used for finding the fit for the concentration-time curve.

Figure 4.3 shows the mean difference of GFR estimated from the hump subsets and the full samples. The smallest mean difference ($\overline{GFR_{hump} - GFR_{full}}$) was obtained using the 2 to 8 hump subset while the largest ($\overline{GFR_{hump} - GFR_{full}}$) was obtained using the 5 to 8 hump subset (approximately -78 ml/min). Despite the fact that the OLS-GV model has small ($\overline{GFR_{hump} - GFR_{full}}$) (between -3 ml/min to 5 ml/min) using the hump subsets 1 to 4 to 4 to 8, the large difference obtained using the 5 to

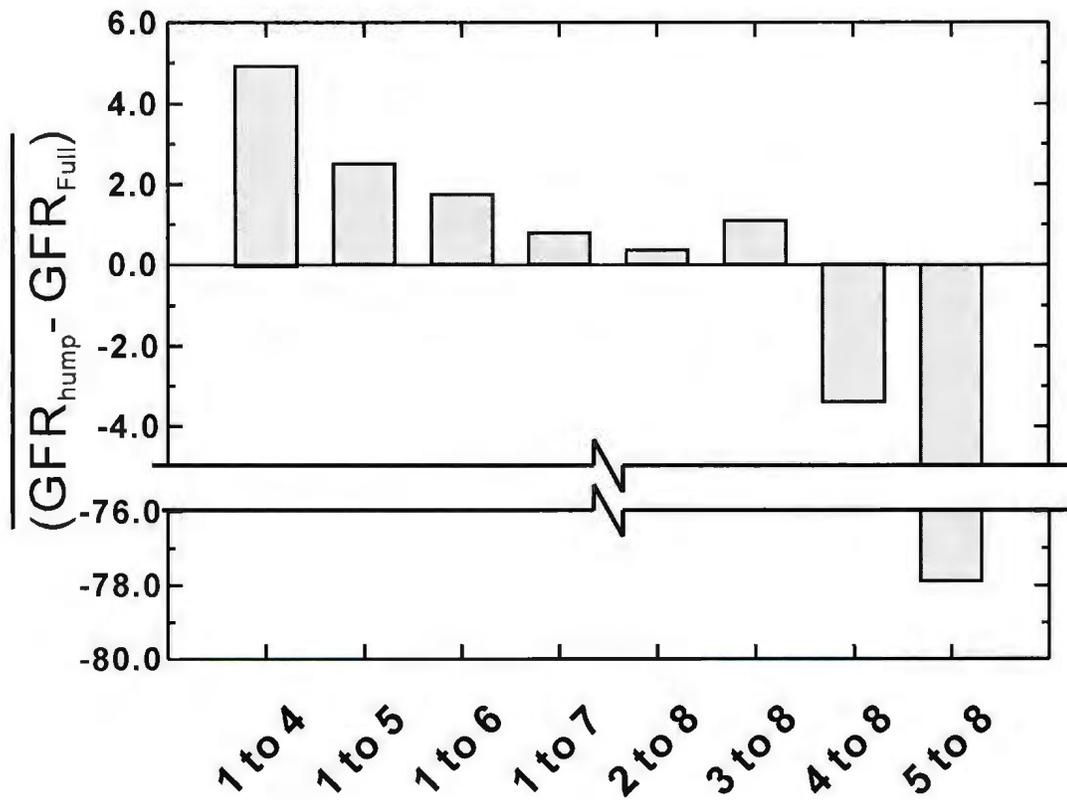


Figure 4.3: Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the OLS-GV model from the 41 patients. The hump subsets are 1 to 5, 1 to 6, 1 to 7, 2 to 8, 3 to 8 and 4 to 8.

8 hump subset indicated that the OLS-GV model can give non-physical results when later samples are used. Table 4.10 also indicates that the the OLS-GV model can give non-physical results using the earliest 4 samples. As observed for the *SET1* and *SET2* models, increasing the number of samples improved the performance of the model. This was also observed for the OLS-GV model using the hump subsets. In order to overcome the occurrence of non-physical results, it is beneficial to include samples from both early and late samples or to improve the performance of the OLS-GV model by using Tikhonov regularization, see the results below.

4.4 Tk-GV Results

Table 4.11: The *GFR* and *Vol* results of the Tk-GV model for the 41 patients using the full 8 samples and four different subsets: (10, 20, 60, 180)min, (10, 30, 120, 240)min, (10, 20, 30, 45)min and (60, 120, 180, 240)min. Q stands for quartile.

<i>GFR</i>	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	1.2	2.9	2.41	13.7	2.1
1st Q	44.4	49.2	48.5	71.8	44.7
2nd Q	74.3	76.2	76.3	107.8	67.2
3rd Q	105.5	108.0	111.9	130.0	98.4
max	157.6	154.9	161.1	205.8	159.5
<i>Vol</i>	full	(10,20,60,180)	(10,30,120,240)	(10,20,30,45)	(60,120,180,240)
min	7.4	7.3	6.9	6.3	6.5
1st Q	13.1	12.8	13.7	11.2	13.9
2nd Q	16.3	15.8	16.1	13.5	17.8
3rd Q	18.5	18.8	19.2	16.0	19.7
max	31.1	30.5	33.3	27.9	27.3

Table 4.11 shows the estimated *GFR* and *Vol* values from the 41 patients using the full samples and the same four 4-sample subsets used for the analysis of the

previous models. Excluding the (10, 20, 30, 45)min subset, the difference between the estimated GFR values using the full samples and the 4-sample subsets in Table 4.11 were much smaller than those from the $SETs$ and OLS-GV models. When considering all 4 subsets and the full samples, the minimum estimated GFR ranged from 1.2 ml/min to 13.7 ml/min and the maximum GFR ranged from 154.9 ml/min to 205.8 ml/min, which shows great improvement in the agreement among the results. Similarly, the estimated Vol values did not differ significantly when comparing the results from each subset to those obtained using the full samples. The (10, 20, 30, 45)min subset gave the worst agreement with the full samples and overestimated the minimum and maximum GFR values by 12.5 ml/min and 48.2 ml/min, respectively. The (10, 20, 60, 180)min subset gave results that were the closest to those obtained using the full samples. The results from the (10, 30, 120, 240)min subset and (60, 120, 180, 240)min subset were not significantly different from each other. This indicates that the Tk-GV model is still affected by the choice of sampling schedule, however, the deviation between the full samples and subsets was much smaller than those from the $SETs$ and OLS-GV models. Lastly, there were no non-physical results observed using the subsets in Table 4.11, even when the earliest 4 samples were used.

Table 4.12: Estimated values of $\bar{\lambda}$, $\overline{\ln K}$, $\bar{\alpha}$, $\bar{\beta}$, \overline{GFR} , \overline{Vol} and the standard deviations SD_{GFR} (ml/min) and SD_{Vol} (L) values using the 2,870 4-sample subsets from the 41 patients using the Tk-GV model.

Quartile	$\bar{\lambda}$	$\overline{\ln K}$	$\bar{\alpha}$	$\bar{\beta}$	\overline{GFR}	\overline{Vol}	SD_{GFR}	SD_{Vol}
min	0	-5.364	0.595	0.0001	3.7	7.4	0.2	0.1
1st quartile	0.012	-4.537	0.714	0.0022	50.3	13.3	1.4	0.3
2nd quartile	0.093	-4.280	0.775	0.0035	75.6	17.1	2.5	0.5
3rd quartile	0.261	-3.991	0.865	0.0046	113.2	18.5	3.7	0.7
max	2.197	-3.386	0.990	0.0091	161.1	30.2	6.2	1.3

Table 4.12 lists the estimated minimum, maximum and quartile values of the mean parameters and the standard deviation of the GFR and Vol values from the 2,870 4-sample subsets. The mean shrinkage factor $\bar{\lambda}$ was mostly between 0 and 1. However the maximum $\bar{\lambda}$ was 2.2 (for Pt19). This is due to the requirement of more regularization for patients whose GFR values are expected to be very low (less than 5 ml/min or almost 0 ml/min) and those patients would have very small plasma clearance value, and hence a very flat concentration-time curve. In those cases, the smallest degree of noise (such as redistribution) would greatly influence the shape of the clearance curve. The estimated parameters $\overline{\ln K}$, $\bar{\alpha}$ and $\bar{\beta}$ values were all within physical ranges. There was no incidence of $\alpha = 0$, $\alpha > 1$ or $\beta < 0$ in the 2,870 subsets. The \overline{GFR} and \overline{Vol} values were quite close to those using the full samples. In addition, the SD_{GFR} and SD_{Vol} values were very small, indicating that the Tk-GV model is stable in performing the study of plasma clearance using different 4-sample subsets. The maximum SD_{GFR} was 6.2 ml/min and the maximum SD_{Vol} was 1.3 L.

Figure 4.4 shows the mean difference of GFR estimated from the hump subsets and the full samples. The difference $(\overline{GFR_{hump}} - \overline{GFR_{full}})$ ranged from approximately -5 ml/min to 22 ml/min. The smallest difference was obtained using the 1 to 7 hump subset. The 4 to 8 and 5 to 8 subsets gave very similar results for the $\overline{GFR_{hump}} - \overline{GFR_{full}}$ value indicating that the Tk-GV model is more robust than the other three models with respect to using the later samples. There is also a trend indicating that the Tk-GV model is likely to overestimate the GFR results using early samples, while underestimating the GFR results using late samples, as observed with the *SET1* and OLS-GV models. A optimal sampling schedule is recommended to include both early and late samples (see Section 5.5). In the next chapter, a comparison of all the models discussed in this chapter will be presented.

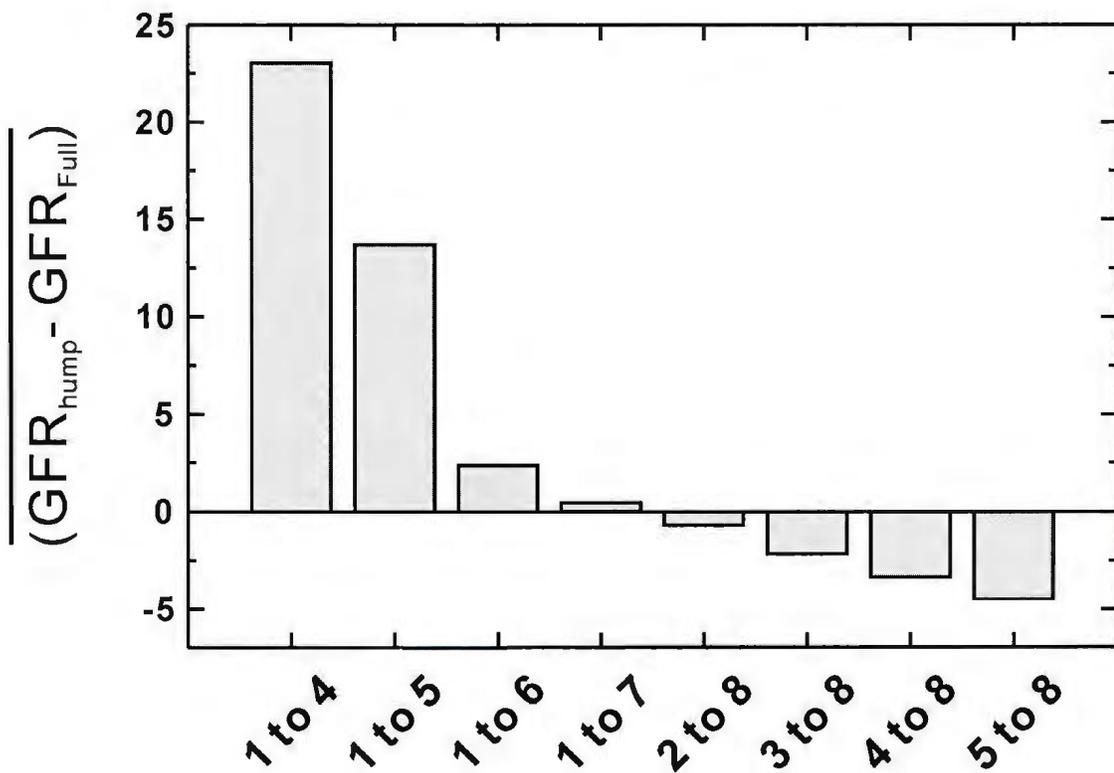


Figure 4.4: Results of the mean of the difference $(\overline{GFR_{hump} - GFR_{full}})$ between the GFR_{hump} and GFR_{full} using the Tk-GV model from the 41 patients. The hump subsets are 1 to 5, 1 to 6, 1 to 7, 2 to 8, 3 to 8 and 4 to 8.

Chapter 5

Discussion

In this chapter, the performance of the *SET1*, *SET2*, OLS-GV and Tk-GV models are compared using 1) the *GFR* estimation, 2) the effects of subsets, 3) the case study, and 4) a cross comparison. The best sampling schedule for applying the Tk-GV model and the clinical impact of using the Tk-GV model are also presented.

5.1 *GFR* Estimation

5.1.1 Non-physical Results

The most important outcome from the plasma clearance study is the estimated *GFR* value. If a model has a high frequency of giving non-physical (or physiologically impossible) *GFR* results ($GFR < 0$), it should not be used for clinical applications. In this work, when the full samples were used, none of the models gave non-physical results due to the large sample size. The comparison among the four models (*SET1*, *SET2*, OLS-GV and Tk-GV) with respect to the frequency of giving non-physical results was performed using subsets (1,148 2-sample subsets for the *SET1* model and

2,870 4-sample subsets for all the models).

When the 1,148 2-sample subsets were used for the *SET1* model for the 41 patients, there were 8 out of 1,148 subsets that gave negative GFR_{SET1} results (0.7%). When the 2,870 4-sample subsets were used for all the models, the number of non-physical results for each model was: the *SET1* model gave 2 out of 2,870 (0.07%), the *SET2* model gave 49 (1.7%), the OLS-GV gave 5 (0.17%) and the Tk-GV gave 0. It is reasonable that the *SET1* model had a higher frequency of giving non-physical results using 2-sample subsets compared with using 4-sample subsets because each model is expected to perform better when fitted to more samples (see Tables 4.1 to 4.5). However, using the 4-sample subsets, the *SET1* model gave fewer non-physical results (0.07%) compared with the *SET2* model (1.7%) because the *SET1* model has a higher probability in finding a convergence during the regression compared with the *SET2* model. Evidence of this is shown in Tables 4.5 and 4.7 where the (10, 20 30, 45)min subset from the *SET2* model gave non-physical GFR values while there were no non-physical results obtained with the *SET1* model. In Table 4.10 there was a non-physical GFR result obtained using the (10, 20 30, 45)min subset with the OLS-GV model. In Table 4.11, there was no non-physical results observed using the Tk-GV model, with all the GFR values being close to each other, irrespective of the use of different subsets. In summary of the results from the 2,870 4-sample subsets, the Tk-GV model was the only model which did not give any non-physical results. On the other hand, as discussed in Section 4.1, if we considered estimated GFR values larger than 200 ml/min as physiologically impossible, the *SET1*, *SET2* and OLS-GV models gave a higher incidence of estimated GFR values larger than 200 ml/min compared to the Tk-GV model. Because of the uncertainty in defining an upper limit for the estimated GFR results, such values are not discussed in this

chapter.

5.1.2 Variations within the Model

Another interesting aspect of the estimated *GFR* results is the degree of overestimation or underestimation caused by different subsets. Since the real *GFR* value is unknown, all the estimated *GFR* results using different models cannot be compared with an absolute value. Therefore, the alternative comparison was performed within each model. If the *GFR* results from different subsets using the same model gave a large range of values, it may indicate the potential of this model to overestimate (or underestimate) the *GFR* values compared with the results using the full samples.

The comparisons in this section focus only on the overestimation (or underestimation) of the *GFR* values within each model using different subsets. When overestimation or underestimation is mentioned in the following, it refers to the comparison of the results from using the subset to those obtained from the full samples from the same model.

SET1

For the *SET1* model, the data in Table 4.1 shows an average overestimation of *GFR* value of 223 ml/min using the (60, 180)min subset and 87 ml/min using the (10, 180)min subset. The (60, 180)min subset gave an average underestimation of the *Vol* value of approximately 10 L while the (10, 180)min subset gave an underestimation of 14 L. Table 4.5 shows that the maximum *GFR* value obtained with the *SET1* model was the largest (overestimated the *GFR* value by approximately 275 ml/min) using the (10, 30, 120, 240)min subset and smallest (underestimated the *GFR* value by approximately 200 ml/min) using the (60, 120, 180, 240)min subset. Tables 4.2,

4.3 and 4.4 summarize the effects of the 2-sample, 3-sample and 4-sample subsets in affecting the \overline{GFR} values using the *SET1* model. The range of the SD_{GFR} values decreased from 1,095 ml/min using the 2-sample subset to between 130 and 140 ml/min using the 3-sample and 4-sample subsets, indicating the improvement of model performance by using more samples. In summary, the *SET1* model has been shown to produce a large difference in the estimated *GFR* values using 2-sample subsets (e.g. (60, 180)min as recommended by the guidelines [41,42]) when compared with the full samples, ranging from approximately 50 to 2,800 ml/min.

SET2

The data in Table 4.7 shows an overestimation of the maximum *GFR* value and underestimation of the minimum *GFR* value (in fact, non-physical) using both the (10, 20, 30, 45)min and (10, 30, 120, 240)min subsets. The (10, 20, 60, 180)min subset gave the best agreement with the results using the full samples, however, this is the only subset that provided all *GFR* values within the expected range. The agreement between the range of estimated *GFR* values using the full samples and the (10, 20, 60, 180)min subset does not necessarily mean that there is similar agreement for individual patient. In fact, the *SET2* model gave large differences in the estimated *GFR* values using different subsets. For example, Table 4.8 shows that the estimated \overline{GFR} values ranged from 1.9 ml/min to 139.1 ml/min while the SD_{GFR} values ranging from 18.8 ml/min to 455.0 ml/min for the *SET2* model applied to the 4-sample subsets. The SD_{GFR} values were generally larger than those using the *SET1* model (ranging from 5.5 ml/min to 146.5 ml/min shown in Table 4.4), indicating that the *SET2* model was more unstable in giving reliable *GFR* values than the *SET1* model, not to mention providing a larger number of non-physical *GFR* results.

OLS-GV

The OLS-GV model delivered a similar performance as that of the *SET2* model in regards to the range of the estimated *GFR* values from using different subsets. Table 4.10 shows that the (10, 20, 30, 45)min still gave non-physical result for the minimum *GFR* value and the (60, 120, 180, 240)min subset significantly underestimated the *GFR* values compared with those using the full samples. The (10, 20, 60, 180)min subset consistently gave the best agreement in estimated *GFR* values compared to the full samples. The (10, 30, 120, 240)min subset, however, did not give any non-physical results as in the *SET2* model. Although, as shown in Table 4.9, the OLS-GV model had a relatively smaller range of SD_{GFR} , between 8.3 to 164.0 ml/min than that of the *SET2* model (SD_{GFR} ranging between 18.8 to 455.0 ml/min in Table 4.8). The fact that the OLS-GV model still gave large variations of *GFR* values using the earliest 4 samples or the latest 4 samples does not make it a good candidate for clinical applications.

Tk-GV

Table 4.11 shows that the Tk-GV model had the smallest variation between the estimated *GFR* values obtained from all four subsets and full samples among all four models. The Tk-GV model gave the largest variation in the minimum (12.5 ml/min) and maximum (48.2 ml/min) *GFR* values using the (10, 20, 30, 45)min subset compared with those using the other subsets. The most important improvement from the other three models is that there was no non-physical result using the Tk-GV model, and the (60, 120, 180, 240)min subset gave very close estimation of *GFR* results to those from the full samples. Because of the agreement in the estimated *GFR* values, all the *Vol* values using the Tk-GV model from the subsets were very close to those

using the full samples. For example, the minimum Vol values ranged from 6.3 to 7.4 L for the full samples and four subsets, and the maximum Vol values ranges from 27.3 to 33.3 L. A further proof of the improvement in the performance of this model is that the SD_{GFR} value ranged from only 0.2 to 6.2 ml/min in Table 4.12, while the other three models all gave ranges of estimated SD_{GFR} values on the order of hundreds of ml/min. Similarly, the range of the estimated SD_{Vol} values was much smaller using the Tk-GV model than those obtained using the other three models.

With respect to the differences in the estimated GFR values, the Tk-GV model performed the best among all the models in that it had the most stable performance from using different sampling schedules by introducing the least variations in the estimated GFR values.

5.2 Effect of Subsets

This section discusses the stability of each model with respect to the different subsets. Two types of subsets were used in this work. One type of subset was the minimum-size sample subset such as the 2-sample subset for using the $SET1$ model or the 4-sample subset for using the $SET2$ and Tk-GV models. The other type of subset was the hump subsets which contain more samples than the minimum requirement for each model, such as 1 to 6 hump subset, 1 to 7 hump subset, etc. All the comparisons in the following sections were based on the results obtained from the 41 patients.

5.2.1 4-sample Subsets

For the purpose of comparison, the 4-sample subsets introduced in Chapter 4 are the primary subsets used in this section. Figures 5.1 to 5.4 show the effects of the four

different subsets (10, 20, 30, 45)min, (10, 20, 60, 180)min, (10, 30, 120, 240)min and (60, 120, 180, 240)min in estimating the GFR values from the $SET1$, $SET2$, OLS-GV and Tk-GV models, respectively. Each graph shows the estimated GFR_{subset} values versus GFR_{full} values. The solid line indicates the best fit line to the GFR_{subset} (using a linear regression) while the dashed line indicates the identity line $Y = X$.

Figure 5.1 b) shows that the (10, 20, 60, 180)min subset gave good agreements with the full samples using the $SET1$ model. The (10, 30, 120, 240)min subset in Figure 5.1 c) had several outliers in the estimated GFR results which increased the scale of the difference between the GFR_{full} and GFR_{subset} from approximately 200 ml/min to 500 ml/min. The (10, 20, 30, 45)min subset gave a poor agreement in that the GFR_{subset} values consistently overestimated the GFR_{full} values by an average of 70 ml/min. In addition, the intercept in Figure 5.1 a) was not zero, indicating that when the estimated GFR using the full samples was very close to 0 ml/min, the estimated GFR value using the (10, 20, 30, 45)min subset was approximately 50 ml/min, which was highly inaccurate for estimating the GFR value when the expected GFR was very low. The last subset (60, 120, 180, 240)min in Figure 5.1 d) was unable to provide reliable GFR results due to the difficulty in fitting the later part of the concentration-time curve (a much flatter portion of the clearance curve).

Figure 5.2 shows that the $SET2$ model is able to improve the agreement of the GFR_{subset} values with the GFR_{full} values when compared with the $SET1$ model using the same subset. Although the fits with the $SET2$ model are generally good, there are outliers using all four subsets, particularly in Figure 5.2 a) where the GFR_{subset} results from the (10, 20, 30, 45)min subset forced the scale of the ordinate from approximately 200 ml/min to 4,000 ml/min while in Figure 5.2 d) many of the estimated GFR values were close to 0 ml/min. The (10, 20, 60, 180)min subset consistently gave the best

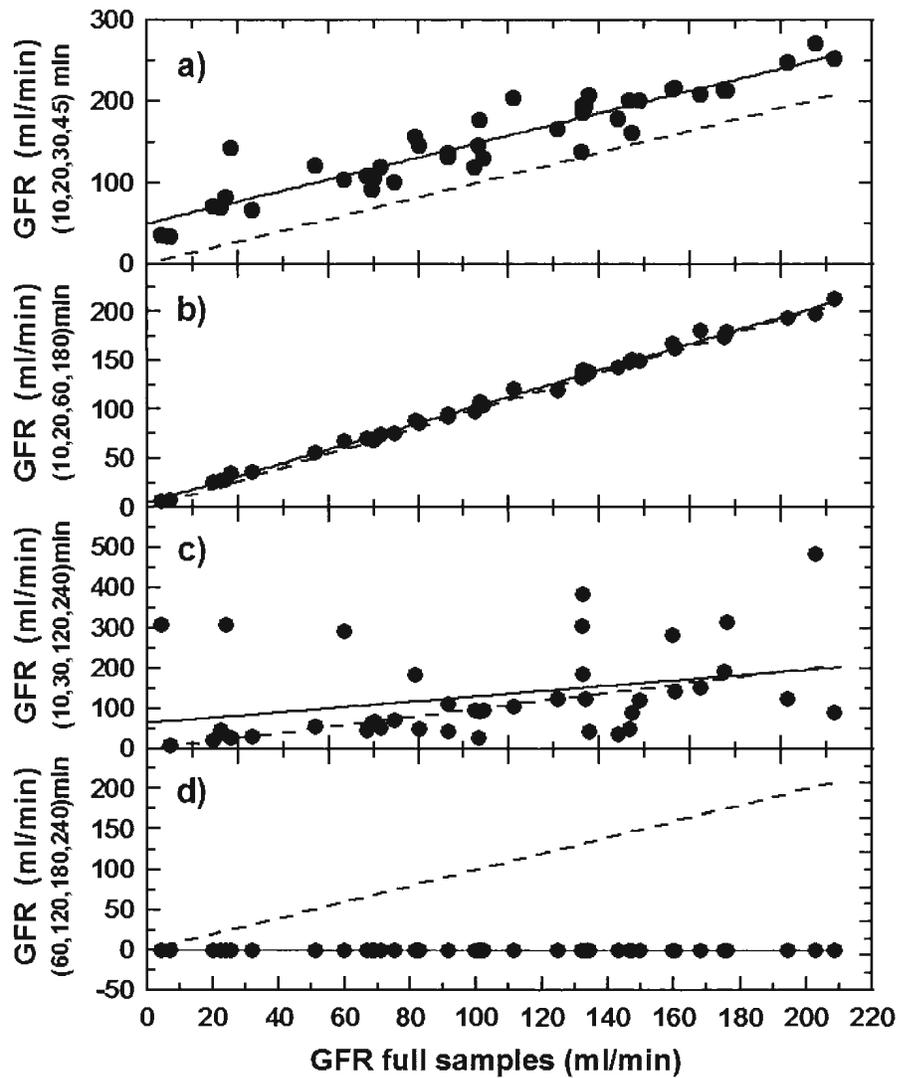


Figure 5.1: Comparison of the GFR_{SET1} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$.

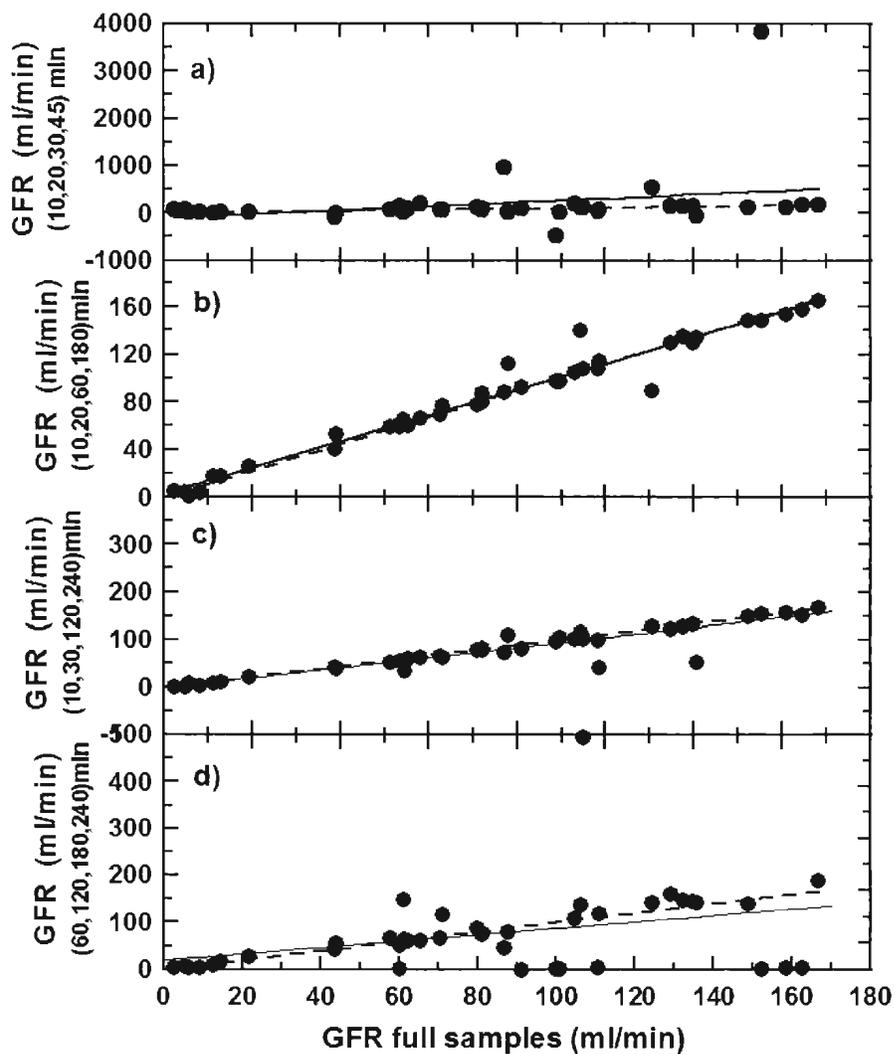


Figure 5.2: Comparison of the GFR_{SET2} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$.

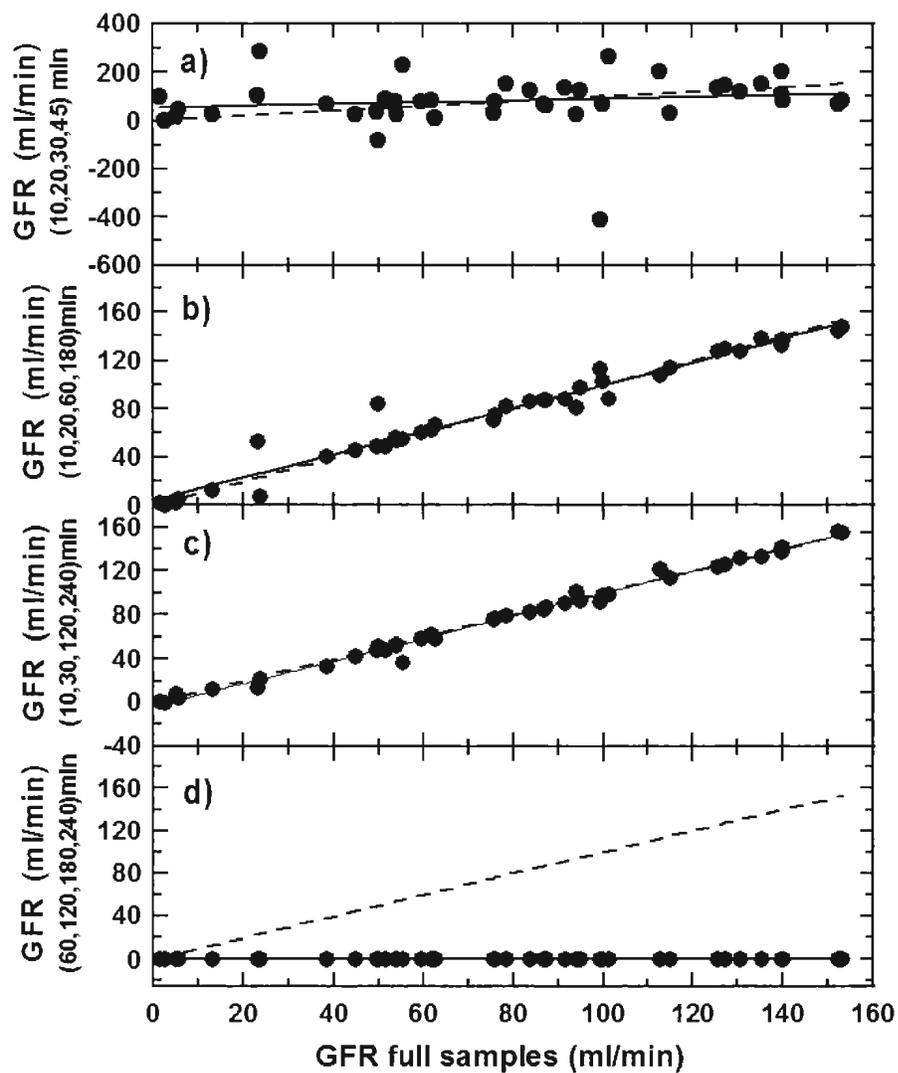


Figure 5.3: Comparison of the GFR_{OLS-GV} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$.

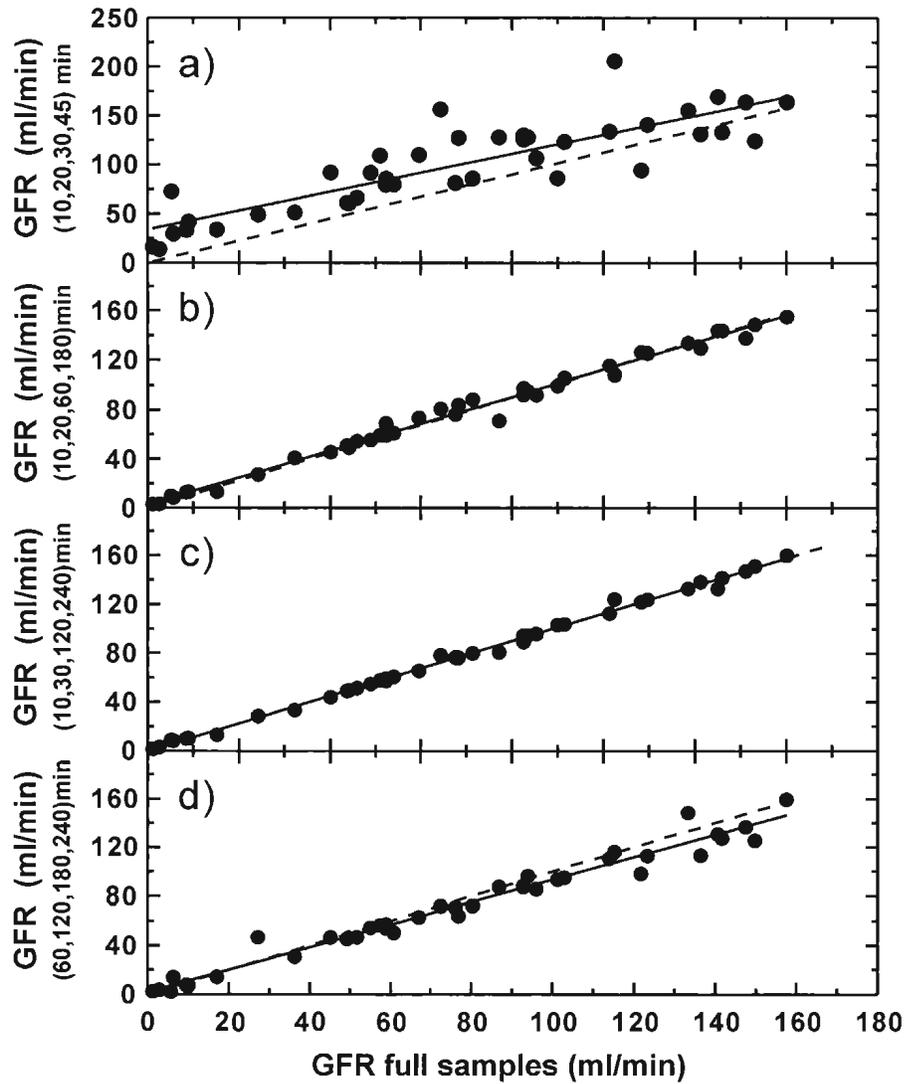


Figure 5.4: Comparison of the GFR_{TK-GV} values obtained using the 4-sample subsets to the GFR_{full} obtained using the full samples. a) (10, 20, 30, 45)min, b) (10, 20, 60, 180)min, c) (10, 30, 120, 240)min and d) (60, 120, 180, 240)min. The solid line indicates a fit of the GFR values and the dashed line is the identity line $Y = X$.

agreement with the results using the full samples among all four models.

Figure 5.3 shows that the estimated GFR_{subset} values from the two subsets, (10, 20, 60, 180)min and (10, 30, 120, 240)min gave good agreements with the GFR_{full} using the OLS-GV model. The (10, 20, 30, 45)min subset shown in Figure 5.3 a) had a few outliers which shifted the scale of the ordinate (see detail of values in Table 4.10). The OLS-GV model was still not able to provide reliable GFR results using the (60, 120, 180, 240)min subset partially due to the limitation of the OLS algorithm in solving an ill-posed GV problem.

Table 5.1: The slope and R^2 values obtained from the correlation between the results from the $SET1$, $SET2$, OLS-GV and Tk-GV models using the 4-sample subsets (shown from Figures 5.1 to Figure 5.4) and the full samples from OLS linear regression.

models	$SET1$		$SET2$		OLS-GV		Tk-GV	
4-sample subsets	slope	R^2	slope	R^2	slope	R^2	slope	R^2
(10,20,30,45)	1.365	0.716	2.731	0.074	0.894	-0.043	1.186	0.590
(10,20,60,180)	1.023	0.994	1.003	0.960	0.997	0.957	0.999	0.976
(10,30,120,240)	1.143	0.034	0.678	-0.283	1.039	-0.195	0.999	0.997
(60,120,180,240)	0.0002	-1.379	0.0001	-1.115	3E-10	0.071	0.935	0.968

All four graphs in Figure 5.4 gave good agreement (in some case excellent agreement) between the GFR_{subset} and GFR_{full} values using the Tk-GV model. The fitted lines were all very close to the identity lines, and the scales of the estimated GFR values were all within expected ranges.

In order to better compare Figures 5.1 to 5.4, Table 5.1 lists the slope and R^2 results from the linear regressions of the fit between the results using the 4-sample subsets and the full samples from all four models. The results of the slopes for the (10, 20, 30, 45)min subset range from 0.894 to 2.731 while the Tk-GV model gave a slope closer to 1. Using the last four samples, which is the (60, 120, 180, 240)min

subset, gave the worst results for all four models, however, the Tk-GV model showed the best agreement (with a slope = 0.935 and $R^2 = 0.968$) between the results using subsets and full samples (see Figure 5.4 d)), compared to the other three models which gave slopes much less than 1 (*SET1*, *SET2* and OLS-GV). The results from the (10, 30, 120, 240)min subset using all four models improved the correlation with those estimated from the full samples, with the *SET2* model giving the worst correlation, with slope = 0.678. In addition, the results from the (10, 20, 60, 180)min subset showed the best agreements for all four models giving slopes value very close to 1 and R^2 very close to 1. However, the other three subsets gave negative R^2 values using the *SETs* and OLS-GV models due to outliers (e.g. non-physical results or very large *GFR* values) of the estimated *GFR* results. In summary, the Tk-GV model gave the best values of slope and R^2 for all four 4-sample subsets among all four models.

A weakness in this comparison is that although 4-sample subsets were used for all models, not all four models require 4 samples as their minimum sample size. Considering that the *SET1* model requires 2, the *SET2* model requires 4, the OLS-GV model requires 3 and the Tk-GV model requires 4 minimum samples, it can be argued that this places the *SET1*, and OLS-GV models at an unfair advantage compared to the Tk-GV model. However, despite the extra advantage given to these two models by using samples more than the minimum requirement, the Tk-GV model was still found to give the best results in terms of frequency of giving non-physical results and curve fitting.

In summary from the results shown in Figures 5.1 to 5.4 and Table 5.1, the Tk-GV model gave the best agreement in the estimated *GFR* values between the full samples and subsets, indicating that the Tk-GV model is the model that is least affected in estimating the *GFR* values by the choice of sampling schedule.

5.2.2 Hump Subsets

The comparison of the results using the hump subsets in this study is mainly based on the four figures shown in Chapter 4 (Figure 4.1 for the *SET1* model, Figure 4.2 for the *SET2* model, Figure 4.3 for the OLS-GV model, and Figure 4.4 for the Tk-GV model). Additional discussion based on the results from the hump subsets shown in our paper is also presented below [118].

Figure 4.1 shows how the mean difference of the estimated *GFR* values using the *SET1* model between the hump subsets and the full samples ($\overline{GFR_{hump}} - GFR_{full}$) ranges from approximately -100 ml/min to 75 ml/min. The trend from the results clearly indicates that the early samples (e.g. 1 to 3 hump subset) overestimated the GFR_{full} values while the later samples (e.g. 5 to 8 hump subset) underestimated the GFR_{full} values using the *SET1* model. This finding agrees with the *GFR* results using the four 4-sample subsets with the *SET1* model shown in Tables 4.4 and 4.5. As shown in Figure 4.2, the *SET2* model is much less sensitive to the choice of the hump subsets. The mean difference ($\overline{GFR_{hump}} - GFR_{full}$) values ranged from approximately -18 ml/min to 12 ml/min which is much lower than those obtained from the *SET1* model. The ($\overline{GFR_{hump}} - GFR_{full}$) value with the *SET2* model decreases when more samples were used. The OLS-GV model greatly reduced the value of ($\overline{GFR_{hump}} - GFR_{full}$) (to less than 10 ml/min) for all the subsets with the exception of the 5 to 8 hump subset which gave an underestimated value of approximately -78 ml/min (Figure 5.3 d)). This agrees with the finding that the OLS-GV model gave a 0.17% frequency of non-physical results from the 2,870 4-sample subsets. Therefore, the OLS-GV model has improved over the *SET1* and *SET2* model, however, it still has limitations during the curve fitting. The Tk-GV model gave estimated ($\overline{GFR_{hump}} - GFR_{full}$) values ranging from approximately -5 ml/min to 22 ml/min

(see Figure 4.4). The largest difference came from the *1 to 4* hump subset (equivalent to the (10, 20, 30, 45)min subset shown in Figure 5.4 a)) which provided a few data points with significant overestimation of the *GFR* value.

In summary, with respect to the effects of subsets, the Tk-GV model is the least sensitive to the choice of subsets, indicating that the Tk-GV model can perform well in the study of the plasma clearance regardless of the choice of the sampling schedule.

5.3 Case Study

This section presents the case studies of 3 patients: Pt1 chosen randomly, Pt15 had the largest estimated *GFR* value, and Pt19 had the smallest estimated *GFR* value among all the 46 patients. For Pt1, the concentration time curve fittings using the four models is presented in Figure 5.5 and for Pt15 and Pt19, Table 5.2 lists the estimated *GFR* results using 4-sample subsets and full samples using the four models.

Figure 5.5 shows the concentration curve fitting using the full 8 samples for Pt1 using the *SET1*, *SET2*, OLS-GV and Tk-GV models. This figure provides an excellent example of how each of the four models performs in estimating the area under the curve (AUC) of the concentration-time curve. Error bars are not included in this plot because the original counts obtained from the gamma counter were unknown from the data provided by Dr. Russell and Dr. Croft. However, the pipetting error is reported to be within 3% [95]. The estimated *GFR* values using the full samples from the *SET1*, *SET2*, OLS-GV and Tk-GV models were: 111.5, 86.7, 55.4 and 67.1 ml/min respectively. The original data from the 8 samples are shown as solid circles. The *SET1* model (shown with a fine dashed line) failed to represent the general trend of data. In addition, the fit given by the *SET1* model gave the largest slope, resulting in the largest estimated $GFR = 111.5$ ml/min. The other three models (*SET2*, OLS-

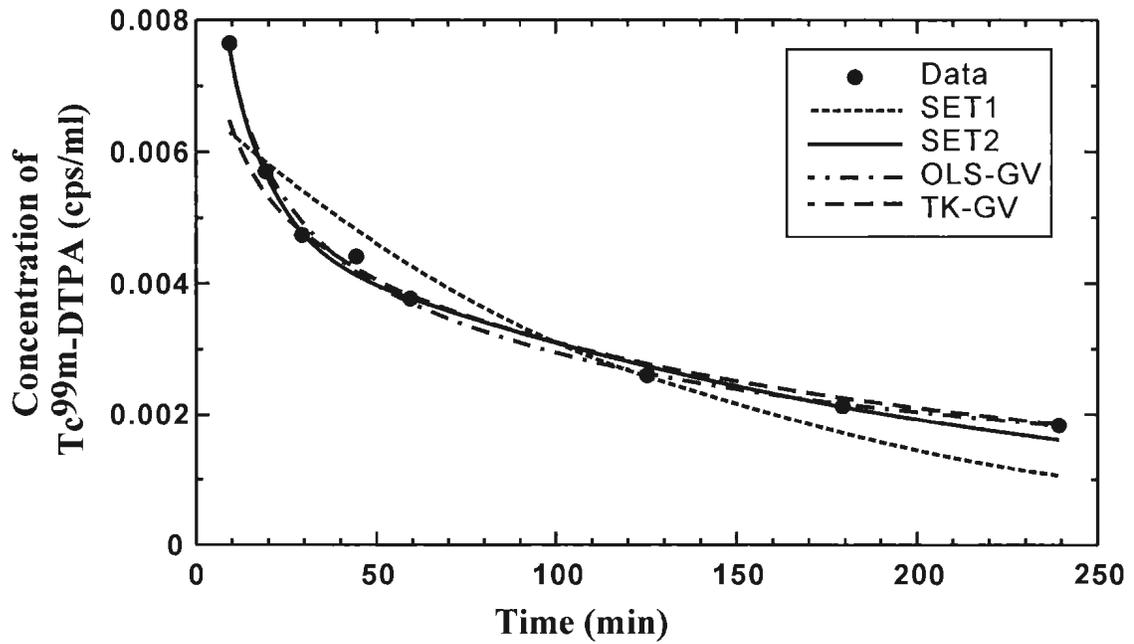


Figure 5.5: Comparison of performance of the concentration time (with time=0 set at the finishing point of the injection) curve fitting using the *SET1* (fine dashed line), *SET2* (solid line), OLS-GV (dash-dot-dash line) and Tk-GV (dashed line) models for Pt1 using the full samples. The original 8 samples are represented with dark circles.

GV and Tk-GV) all gave curves that represent the data, indicating that these three models are not significantly different in fitting the concentration-time curve when the patient's expected GFR value was not critically low. However, when less than 8 samples were used, the difference in the curve fitting among the $SET2$, OLS-GV and Tk-GV models became more concerning because the $SET2$ and OLS-GV models are more prone to giving non-physical results compared to the Tk-GV model.

Table 5.2: Estimated GFR values for Pt15 and Pt19 using the 4-sample subsets (70 subsets) using the $SET1$, $SET2$, OLS-GV and Tk-GV models. The GFR_8 (in ml/min) and Vol_8 (in L) are the estimated results using 8 samples. The \overline{GFR}_4 and \overline{Vol}_4 are the mean estimated results of GFR and Vol using 4 samples.

Pt15	GFR_8	\overline{GFR}_4	Vol_8	\overline{Vol}_4
$SET1$	208.3	123.7	6.0	3.9
$SET2$	166.5	87.0	14.3	7.5
OLS-GV	153.0	71.9	15.5	8.6
Tk-GV	157.6	161.1	16.0	17.0
Pt19	GFR_8	\overline{GFR}_4	Vol_8	\overline{Vol}_4
$SET1$	4.5	48.8	242.8	220.5
$SET2$	2.6	-60.6	11.2	11.6
OLS-GV	2.5	15.2	10.5	7.5
Tk-GV	1.2	2.9	11.6	11.3

Table 5.2 shows the GFR and Vol results for Pt15 and Pt19 using the $SET1$, $SET2$, OLS-GV and Tk-GV models. The GFR_8 and Vol_8 values were obtained using the full 8 samples. The \overline{GFR}_4 and \overline{Vol}_4 values are the mean of the estimated results of GFR and Vol . When the expected GFR value was high, as is the case for Pt15, the $SET1$, $SET2$ and OLS-GV models consistently gave larger GFR_8 values than those from the 4-sample subset. In all but one case, the ratio of GFR_8 to \overline{GFR}_4 was approximately 2. The same observation was obtained for the ratio of Vol_8 to \overline{Vol}_4 , indicating the poor agreement between the estimations using 8 samples and 4

samples. However, the Tk-GV model gave an estimated GFR_8 value of 157.6 ml/min, while 161.1 ml/min was obtained for \overline{GFR}_4 , indicating a very small difference in the GFR results using different sampling schedules. Similarly, the Tk-GV model gave an estimated Vol_8 value of 16.0 L while a value of 17.0 L was obtained for \overline{Vol}_4 . When the expected GFR value was very low, as for the case of Pt19, the difference between the performance of all four models in estimating GFR and Vol values using 8 and 4 samples became larger. The *SET1* model gave the \overline{GFR}_4 10 times larger than the GFR_8 value. The *SET2* model gave a negative \overline{GFR}_4 value of -60.6 ml/min. The OLS-GV model gave a slightly better agreement between the estimated GFR_8 and \overline{GFR}_4 values, however, still with a difference of 12.7 ml/min. The Tk-GV model showed that it can perform well in estimating GFR values when they were expected to be less than 15 ml/min. The GFR_8 and \overline{GFR}_4 values were 1.2 ml/min and 2.9 ml/min using the Tk-GV model, while 11.6 L and 11.3 L for Vol_8 and \overline{Vol}_4 , respectively. It should be made clear that the fits obtained using *SET2* and OLS-GV for the 8 samples were very close to the fits obtained using Tk-GV for both the 8 and 4 sample subsets, indicating the inherent difficulty in estimating low GFR values.

The case studies based on the results from patients Pt1, Pt15 and Pt19 have shown that the Tk-GV model consistently gives the best performance in fitting the concentration-time curve, as well as estimating the GFR and Vol values compared with other models.

5.4 Cross Comparison

Figure 5.6 shows the correlation of the estimated GFR values using the *SET1* and *SET2* models compared with the Tk-GV model using the full samples from the 46 patients. The correlation between the *SET1* and Tk-GV models gave a slope of

1.18. The reciprocal of the slope is $1/1.18 = 0.85$ which is very close to Chantler's correction factor of 0.87 used to modify the overestimation of the *SET1* model using $GFR_{\text{corrected}} = 0.87 \times GFR_{\text{SET1}}$ (see Section 2.4). This shows that the Tk-GV model has corrected the overestimation of the *GFR* values from the *SET1* model compared with the results using the standard clearance technique (which was calculated as UV/P using the urinary clearance, urine flow rate and plasma activity, see Section 2.2.1) [60].

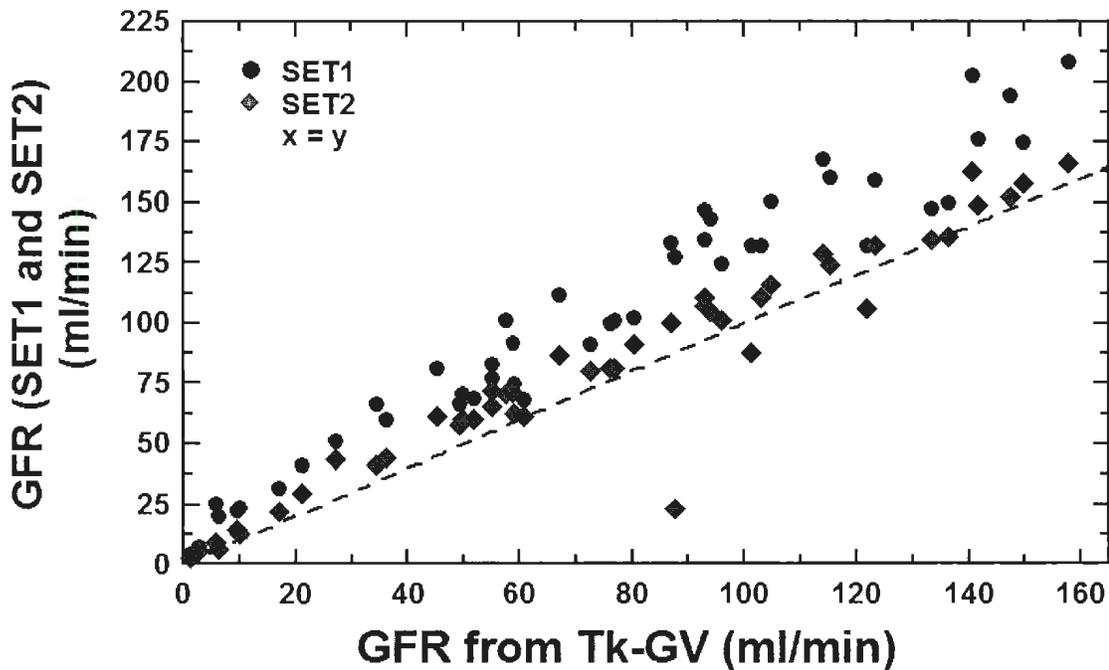


Figure 5.6: Comparison of the 46 estimated *GFR* values using *SET1* (dark circles) and *SET2* (gray rhombus) compared with Tk-GV, using the full samples. The dashed line is the identity line $Y = X$.

Florijn's study showed that there is a constant difference of 5.1 ml/min in the *GFR* values obtained using the *SET2* model compared with those from the constant infusion of inulin (known as gold standard for estimating *GFR* values) [80]. The correlation of the *GFR* results in Figure 5.6 gave a slope of $0.99 \approx 1$ and an intercept of

6.4 ml/min (between the *SET2* and Tk-GV models) which is very close to 5.1 ml/min. This indicates that the Tk-GV model has corrected the inaccuracy in estimating the *GFR* value from the *SET2* model compared with the gold standard. This agreement with the correction factor is further validated by scaling the results of the *SET2* and Tk-GV models by the mean values of *GFR* and the parameter β , which gives a correlation coefficient of approximately 1.106. This value agrees with the finding from Moore et al. where an overestimation of approximately 10% obtained by the *SET2* model was observed based on 24 h plasma clearance study [61, 118]. This shows that the Tk-GV model can consistently perform well without the requirement for further corrections.

The cross-comparison of the Tk-GV model with the *SET1* and *SET2* models has shown that the Tk-GV model reduces the overestimation of the *SET1* and *SET2* models by the same amount predicted by the Chantler's correction factor and by Moore's study. In addition, the cross-comparison with Florijn's observation shows that the Tk-GV model gave results that are close to those obtained using the gold standard.

5.5 Best Sampling Schedule

It is a useful exercise to attempt to determine the best sampling schedule (using 4 samples) in consideration of 1) estimating the *GFR* value accurately, 2) being the least invasive for the patients to follow during the kidney function test, and 3) providing the most convenience for the technologists to perform. The minimum number for samples required to use the Tk-GV model is 4.

There are $\binom{8}{4} = 70$ different combinations of choosing 4 samples out of 8 samples. The 70 estimated GFR_4 values were compared with the GFR_8 values using the full 8

samples. The results from the 41 patients were used to find the best sampling schedule using the residuals calculated from the GFR_8 and GFR_4 values. Five subsets (10, 30, 180, 240)min, (10, 30, 120, 240)min, (10, 45, 120, 240)min, (10, 20, 45, 240)min, and (10, 45, 180, 240)min were identified from the 70 subsets with residual values < 2 ml/min. Given that the normal range of GFR values is between 100 to 130 ml/min/1.73 m², a residual < 2 ml/min indicates a relative error $< 2\%$. Residuals were calculated as the difference in estimated GFR values between the chosen 4-sample subset and full samples. Among all these 5 subsets, the residuals ranged from 1.35 ml/min to 1.86 ml/min, indicating no significant difference in the relative error. In consideration of the clinical practise for collecting blood samples, the (10, 30, 120, 240)min subset is of most interest. Since patients who undergo kidney function tests using the plasma clearance method, are also required to undergo renal scans (a renal scan usually lasts 30 min), it is convenient to collect the two blood samples during the renal scan (10 min and 30 min). The collection time of 45 min was not chosen because it would require the patient to return to the department of Nuclear Medicine 15 min after the renal scan. It is interesting that the best sampling schedule includes early samples before 1 h after the injection. Remember that for the *SET1* model, the recommended starting sampling time was after 1 h [72]. Although there are concerns about how including early samples can affect the curve fitting of the concentration-time curve, studies have shown the need of including early samples using the *SETs* models. For example, a Monte Carlo study performed in 2012 recommended that the sampling schedule include 3 early samples before 1h post-injection for using the *SET2* model [86]. Another study by Dr. Russell recommended that the sampling started by 5 min while additional samples are collected within the following 90 min using the *SET1* model [85]. Fleming et al. also pointed out the importance of including early samples while using the *SETs* models [130]. The debate on whether or not to

include early samples for the *SETs* models, again, shows the robustness of the Tk-GV model in that it can perform well using both early and late samples. The 120 min collection time is more attractive because the concentration-time curve between 120 min and 240 min has a larger elimination rate than that between 180 min and 240 min. Because the Tk-GV model is very robust with respect to the choice of subsets, different health institutions can choose a sampling schedule from these 5 subsets that best meet their needs without compromising their expected *GFR* results.

5.6 Clinical Impact

It is interesting to investigate how the difference in estimating the *GFR* values using the *SET1* model and *SET2* model compared to the Tk-GV model can affect the determination of the stage of chronic kidney disease (CKD) (see Table 2.1). Figure 5.7 shows the difference in the estimated *GFR* values using the *SET1* and *SET2* models compared with those using the Tk-GV model with respect to the 5 stages of CKD. Figure 5.7 a) gives the relative difference $(GFR_{SET1} - GFR_{TK-GV})/GFR_{TK-GV}$ as a function of the GFR_{TK-GV} values for the 46 patients. The division of the 5 CKD stages was discussed in Section 2.1.1. Because the CKD stages are described based on body surface area (BSA) corrected *GFR* values, the estimated *GFR* values from the *SET1*, *SET2* and Tk-GV models were also corrected with their corresponding BSA values. From the previous discussion, the *SET1* model is known to overestimate the *GFR* values. Figure 5.7 a) shows that the *SET1* model overestimates the *GFR* results more when the expected *GFR* value is < 30 ml/min/1.73 m² (CKD4 and CKD5). There were 6 patients falling into the category of $GFR < 15$ ml/min/1.73 m². The estimated *GFR* values from these 6 patients have been significantly overestimated (from 137% to 344%) with the *SET1* model compared with that from the Tk-GV

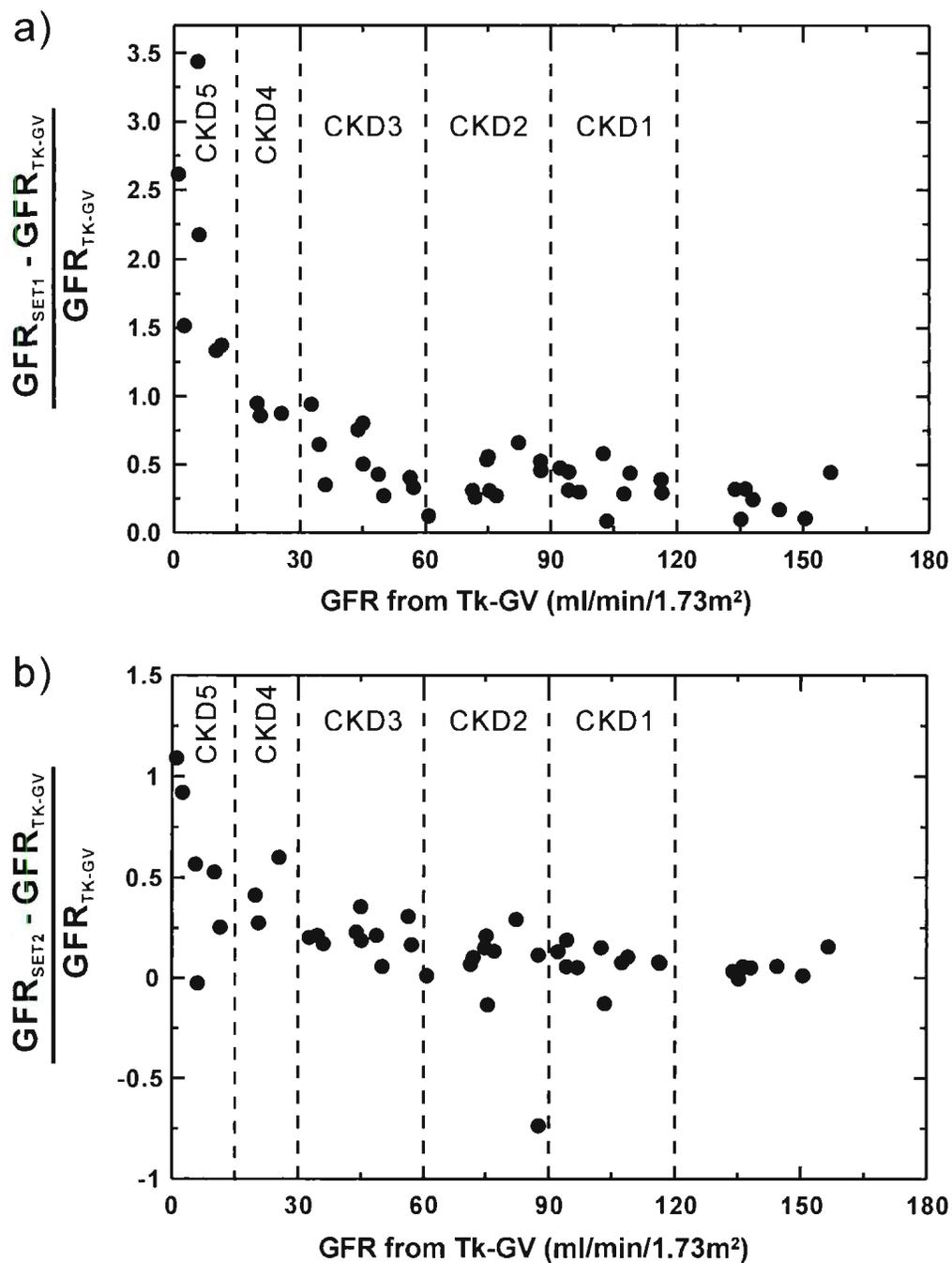


Figure 5.7: Comparison of (a) $(GFR_{SET1} - GFR_{Tk-GV})/GFR_{Tk-GV}$ from *SET1* with Tk-GV and (b) $(GFR_{SET2} - GFR_{Tk-GV})/GFR_{Tk-GV}$ from *SET2* with the Tk-GV model for the 46 patients. The vertical dashed lines indicate the CKD stages from CKD1 to CKD5.

model. This means that if the patient had a GFR value of 15 ml/min/1.73 m², the $SET1$ model would overestimate the GFR value between 21 to 51 ml/min/1.73 m². As a result, the patient's estimated GFR values from the $SET1$ model would be between 36 and 67 ml/min/1.73 m² which could alter the staging from CKD5 to CKD3 or CKD2. There were 3 patients with GFR values between 15 and 30 ml/min/1.73 m² which were overestimated using the $SET1$ model by a factor of approximately 2 compared with the Tk-GV model. The relative ratios of overestimation were between 86% and 95%, which would increase the expected GFR values from 30 ml/min/1.73 m² to 60 ml/min/1.73 m² using the $SET1$ model.

Figure 5.7 b) shows the relative error between the $SET2$ and Tk-GV model, $(GFR_{SET2} - GFR_{Tk-GV})/GFR_{Tk-GV}$ after the BSA correction indicating that the $SET2$ and Tk-GV models were closer in the analysis of the CKD stages than between the $SET1$ and Tk-GV models. For $GFR < 15$ ml/min/1.73 m², 5 out of the 6 patients in category CKD5 were overestimated by 25% to 109%. However, Pt36 had a slightly lower GFR value using the $SET2$ model (98%) compared with that from the Tk-GV model. For the rest of the CKD stages, $SET2$ had 5 patients being underestimated and 41 patients being overestimated compared with that from the Tk-GV model.

Using the $SET1$ or $SET2$ models can introduce significant differences in the estimated kidney function compared with the Tk-GV model. The importance of estimating GFR accurately at low ranges (e.g. < 15 ml/min/1.73 m²) can affect the clinical management of patients. Therefore, it is crucial to improve the accuracy and precision of the GFR measurement for patients with different ranges of expected kidney function.

5.7 Summary

From the above discussion, with respect to the number of non-physical results, the variation of the *GFR* values using different subsets within the model, the effects of different types of subsets, and a cross comparison of the results to the gold standard (constant infusion of inulin method), the Tk-GV model was consistently better than the *SET1*, *SET2* and OLS-GV models. The results from this study have been published in the Journal of Pharmacokinetics and Pharmacodynamics in 2010 [118]. A clinical protocol of how to collect blood samples for using the Tk-GV model developed at the General Hospital in St. John's, NL in 2010 is given in Appendix A.1.

Chapter 6

Applications to 24 h Data

In this chapter, the results of the Tk-GV model applied to a study sample of 10 patients with a 24 h sampling schedule are shown and discussed. These results were not discussed together with the 46 patients in Chapters 4 and 5 because: 1) all 10 patients had liver dysfunction and their kidney function was, according to guidelines, difficult to estimate accurately using the *SET1* model [130], and 2) these patients had a 24 h sampling schedule with more than 4 samples, which allows the comparison of the Tk-GV model and the *SET1* model over a longer time scale. As a result, this chapter is shown as an extension of Chapters 4 and 5, and hence is presented separately.

6.1 Data

The preliminary study based on the 4 h sampling schedule showed that the Tk-GV model performed the best among all four models. How the Tk-GV model performs on a plasma clearance study with a sampling schedule longer than 4 h is unknown. In 2010, 13 patients' data processed with a 24 h sampling schedule were collected

and provided by Dr. Maria Burniston in the department of medical physics, Royal Free Hampstead NHS Trust in London, England. The median body mass of these 13 patients was 78.4 kg (range from 51.1 kg to 105.8 kg) and median body height was 167.8 cm (range from 150.0 cm to 188.0 cm). These 13 patients who were under assessment for liver transplantation were studied using a ^{51}Cr -EDTA bolus injection. The uncertainty of the timing and activity of the samples were not provided by Dr. Burniston, however, one can follow the general estimation of the uncertainty introduced in Section 3.2. A comparison of the results based on this dataset using the Tk-GV model and *SET1* model has been presented and published in the European Journal of Nuclear Medicine and Molecular Imaging in 2011 [131].

These 13 patients had multiple blood samples (varying from 7 to 16 samples) drawn between 5 min and 24 h after the injection of the tracer. Among all 13 patients, Pt7 had the fewest samples (7 samples) and the shortest sampling time (12 h). Patients Pt1 and Pt2 also had a limited sample size (8 and 10 samples, respectively). Therefore, in order to perform a comparative investigation of the behavior of the Tk-GV model and the *SET1* model with a uniform sampling schedule, patients Pt1, Pt2 and Pt7 were not included in the results and discussion of this analysis. The 10 patients used in this work had 12 or more samples taken within 24 h. As a result of the conclusion of Chapters 4 and 5, which showed that the Tk-GV model is a better model than the *SETs* models for estimating *GFR* values, it is not necessary to again compare the *GFR* results using the *SET1*, *SET2*, OLS-GV and Tk-GV models. The application of the Tk-GV model to these 10 patients should allow us to investigate: 1) if the Tk-GV model performs consistently better than the *SET1* model in estimating the *GFR* values, and 2) the effects of using late sample times (later than 4 h) on the performance of the Tk-GV model and the *SET1* model. For completeness, a brief

discussion of the GFR values obtained using the $SET2$ model is given in Section 6.3.

6.2 $SET1$ versus Tk-GV

6.2.1 GFR Results

Table 6.1: Estimated GFR (ml/min) values from the 10 patients with 4 h, 12 h and 24 h sampling schedules, using the $SET1$ and Tk-GV models.

Pt	GFR_{SET1}			GFR_{Tk-GV}		
	24 h	12 h	4 h	24 h	12 h	4 h
3	75.1	81.0	117.8	79.7	79.6	74.4
4	43.2	48.6	85.6	45.8	43.8	43.7
5	71.7	86.9	173.5	76.5	82.7	89.6
6	52.3	58.7	83.5	56.5	57.6	60.0
8	41.7	57.9	154.8	29.9	36.7	64.4
9	42.7	48.9	74.6	46.5	47.6	49.1
10	15.6	22.1	36.4	11.4	15.9	19.5
11	21.9	25.4	45.5	21.2	19.3	24.5
12	42.1	52.5	78.4	42.5	49.1	53.3
13	43.9	58.9	92.7	36.4	45.1	54.2

This section will focus primarily on the comparison of how the $SET1$ model and the Tk-GV model perform in fitting clearance data obtained within a 24 h sampling schedule.

Table 6.1 lists the estimated GFR results using the 4 h, 12 h and 24 h sampling schedules for the 10 patients using the $SET1$ model and the Tk-GV model. The difference in GFR values between the $SET1$ model and the Tk-GV model is the largest when considering the 4 h sampling schedule (see Table 6.1). The minimum difference in the 10 GFR results using a 4 h sampling schedule is 39% while the

maximum difference is 140%. The range of differences for the 12 h sampling schedule indicates that the *SET1* model is less deviated from the Tk-GV model in estimating the *GFR* values. However, the results from the 4 h and 12 h sampling schedule show that the *SET1* model consistently overestimated the *GFR* values compared with those from the Tk-GV model. When the 24 h sampling schedule was used, both models gave similar *GFR* values. However, the 24 h sampling schedule is much more time-consuming than the 4 h sampling schedule and is more inconvenient to the patients and the technologists.

6.2.2 Curve Fit

A straightforward demonstration of the advantage of the Tk-GV model over the *SET1* model is to compare the concentration time curve fitting. Figure 6.1 shows the logarithm of the concentration as a function of time using the *SET1* model and the Tk-GV model for Pt3. The hollow circles indicate the original data as collected. The solid line and dashed line indicate the best fit curve using the Tk-GV model and the *SET1* model respectively. Figure 6.1 shows that the Tk-GV model gave a much better fit to the original data points than the *SET1* model. The difference between the two fits became more significant after 6 h. However, despite the poor fit given by the *SET1* model, the estimated *GFR* results from these two models using the data from the 24 h sampling schedule were very close (75.1 ml/min using the *SET1* model and 79.7 ml/min using the Tk-GV model).

The fact that the *SET1* model is able to provide a *GFR* value similar to that of the Tk-GV model is surprising but not inexplicable. Figure 6.2 shows the fits obtained using the *SET1* model from the 4 h, 6 h, 8 h and 24 h sampling schedules for Pt3. Error bars are not included in this plot because neither the pipetting error nor the

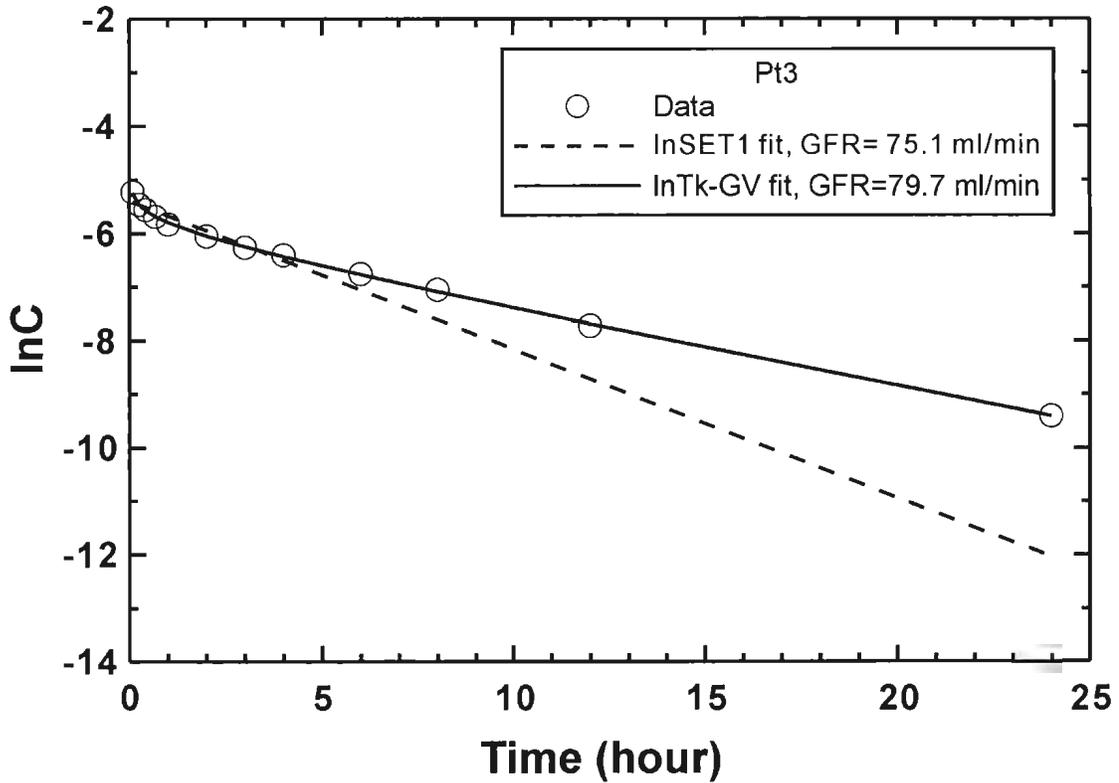


Figure 6.1: Logarithm of the concentration versus time (with time=0 set at the finishing point of the injection) using the *SET1* and Tk-GV models for Pt3. The original data is shown with hollow circles, the Tk-GV fit is indicated by a solid line while the *SET1* fit is indicated by a dashed line.

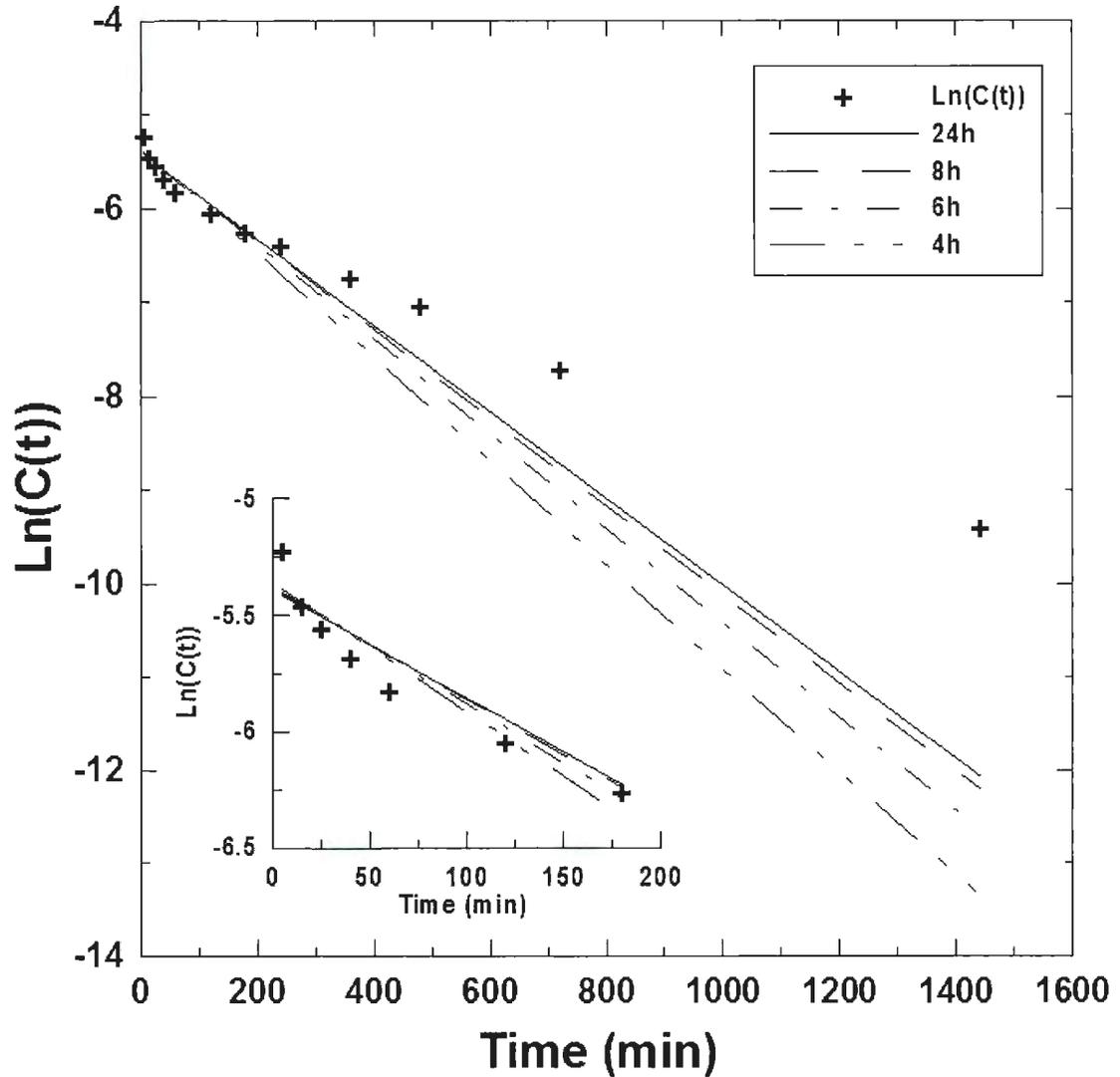


Figure 6.2: Logarithm of the concentration versus time (with time=0 set at the finishing point of the injection) using the *SET1* model for Pt3. The original data are shown as crosses (+). The straight lines indicate the fits obtained using the *SET1* model with different sampling times as shown in the legend. The insert on the bottom left shows how the *SET1* model overestimates the concentration data for earlier time.

counting error were provided by Dr. Burniston. From Chapter 2, GFR was defined as the Dose divided by the area under the clearance curve (AUC). Although the fit from the $SET1$ model using the 24 h data may seem to completely misrepresent the clearance data, the AUC is only slightly smaller than that obtained using the Tk-GV model. This is partially because the $SET1$ fit overcompensates the clearance data for the first 2 to 3 h (see insert in Figure 6.2). Hence, since the injected dose is constant, the GFR_{SET1} value is approximately equal to the GFR_{Tk-GV} value. As the $SET1$ model is fitted to fewer samples (smaller sampling schedule), the slope of the fit decreases (become more negative) and, as a result, the AUC decreases, causing the GFR_{SET1} to increase. For example, when only a 4 h sampling schedule was used, the estimated GFR value from the $SET1$ model is 117.8 ml/min while the Tk-GV model gave 74.4 ml/min, see Table 6.1. However, since the Tk-GV model is able to more consistently represent the clearance data, the GFR_{Tk-GV} remains approximately constant. As a result, the GFR_{SET1} overestimates the expected value of the GFR where a small sampling schedule is used.

6.2.3 4 h versus 24 h

Figure 6.3 shows that the results from the 10 patients using the $SET1$ model with a 4 h sampling schedule are significantly overestimated compared to those using the Tk-GV model with a 24 h sampling schedule. The reason the results of the $SET1$ model with 4 h samples are compared with those of the Tk-GV model using the 24 h samples is because the 4 h $SET1$ results can be used as reference values according to the recommendation from the national guidelines [41, 42, 69, 130]. This figure can effectively show the overestimation and relative errors in the estimated GFR values using the $SET1$ model compared with those using the Tk-GV model. The range of the

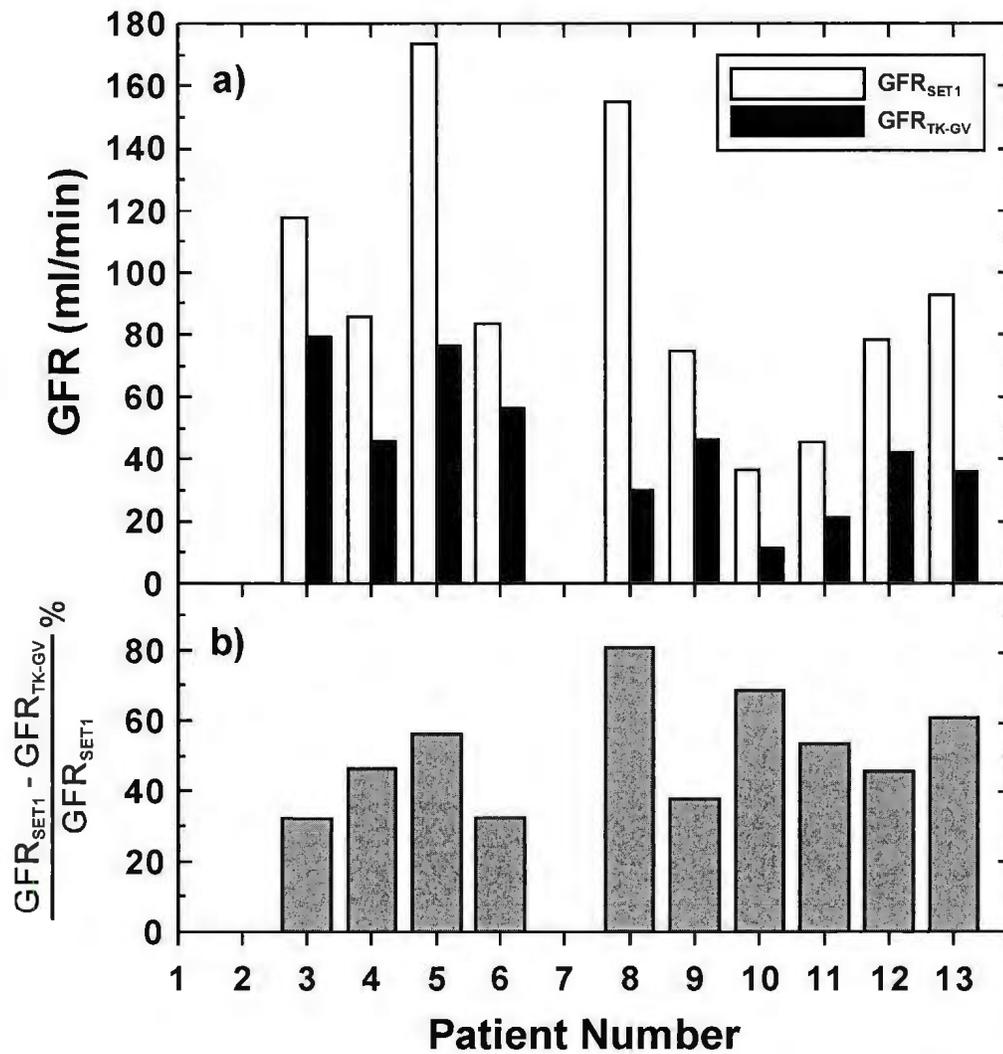


Figure 6.3: (a) Estimated GFR results (ml/min) for the 10 patients with a 24 h sampling schedule using the Tk-GV model (black bars) and the $SET1$ model (white bars) using the 4 h sampling schedule. (b) The relative difference in GFR between the two models $\frac{GFR_{SET1} - GFR_{TK-GV}}{GFR_{SET1}}$.

overestimation spanned from approximately 32% to 81%. The level of overestimation using the *SET1* model might alter the diagnosis of a patient's kidney function, and hence influence the required treatment.

Figure 6.4 a) shows that for the 4 h sampling schedule, the estimated *GFR* values using the *SET1* model were on average 1.78 times those using the Tk-GV model. On the other hand, Figure 6.4 b) shows that the estimated *GFR* values from the 24 h sampling schedule using the *SET1* model and the Tk-GV model were very close together with a slope of 0.98. The two opposite correlations show the instability of the performance of the *SET1* model in fitting the concentration time curve. As shown in Figure 6.2 and discussed above, the *SET1* model overestimates the *GFR* values using early samples, while underestimating the *GFR* values using the later samples.

This 24 h data has validated the findings in Chapters 4 and 5 that the Tk-GV model performs better than the *SET1* model. In addition, the fact that these 10 patients were studied using ^{51}Cr -EDTA (a difference tracer than ^{99m}Tc -DTPA) indicates that the Tk-GV model can be used for the study of plasma clearance with a different choice of tracer.

6.3 Effect of Subsets

This section investigates the effect of subsets on the *GFR* values estimated with each model. The *GFR* results from the 4 h, 6 h, 8 h and 12 h subsets were compared with those obtained from the 24 h sampling schedule.

Figures 6.5 and 6.6 show that the estimated *GFR* values from the different sampling schedules have different agreement with those obtained using the full samples (24 h) using the Tk-GV model and the *SET1* model respectively. The Tk-GV model

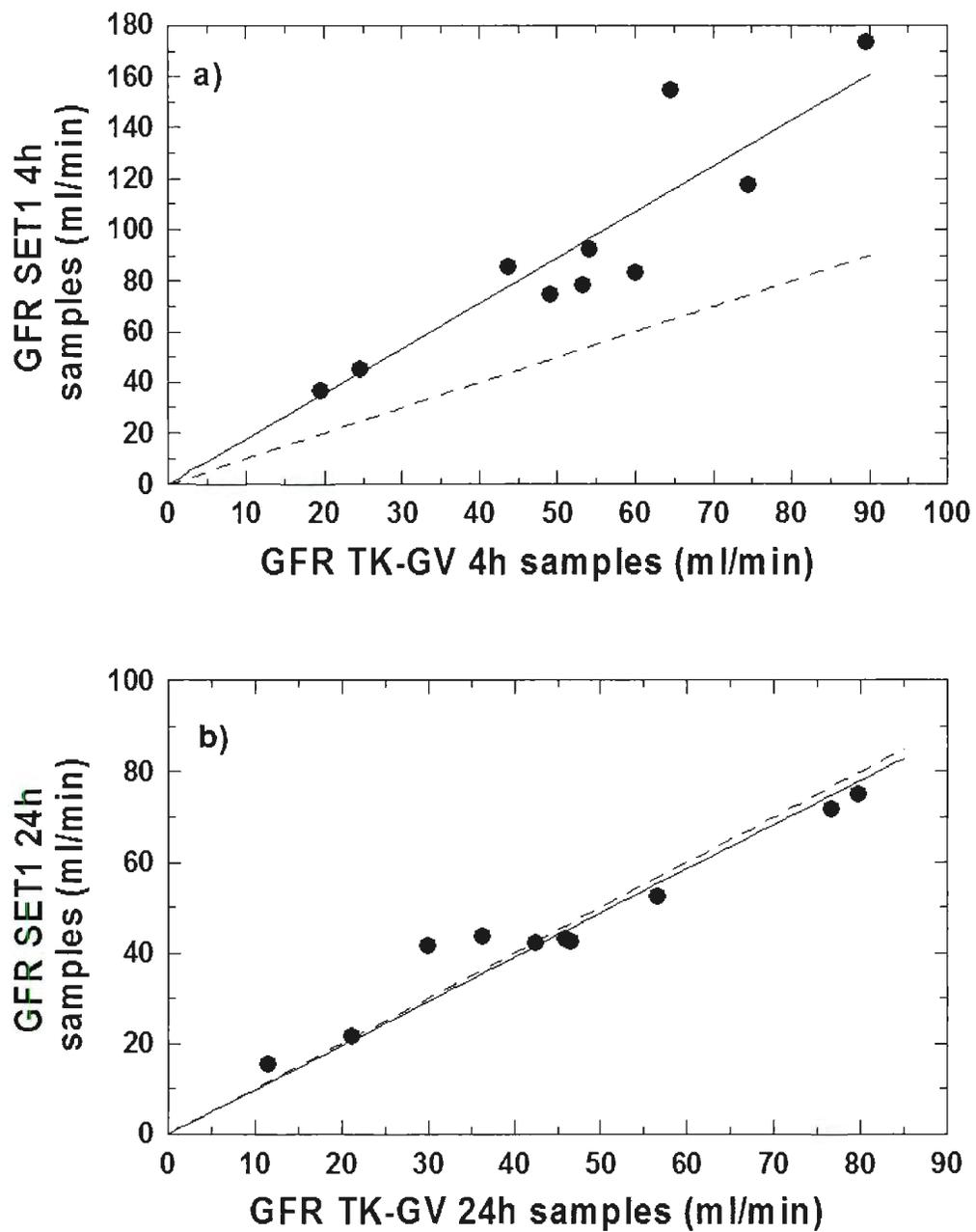


Figure 6.4: Correlation between the estimated *GFR* values using the *SET1* model and the *Tk-GV* model from a) the 4 h and b) the 24 h samples. The dashed line indicates the identity line $Y = X$ while the solid line is the fit.

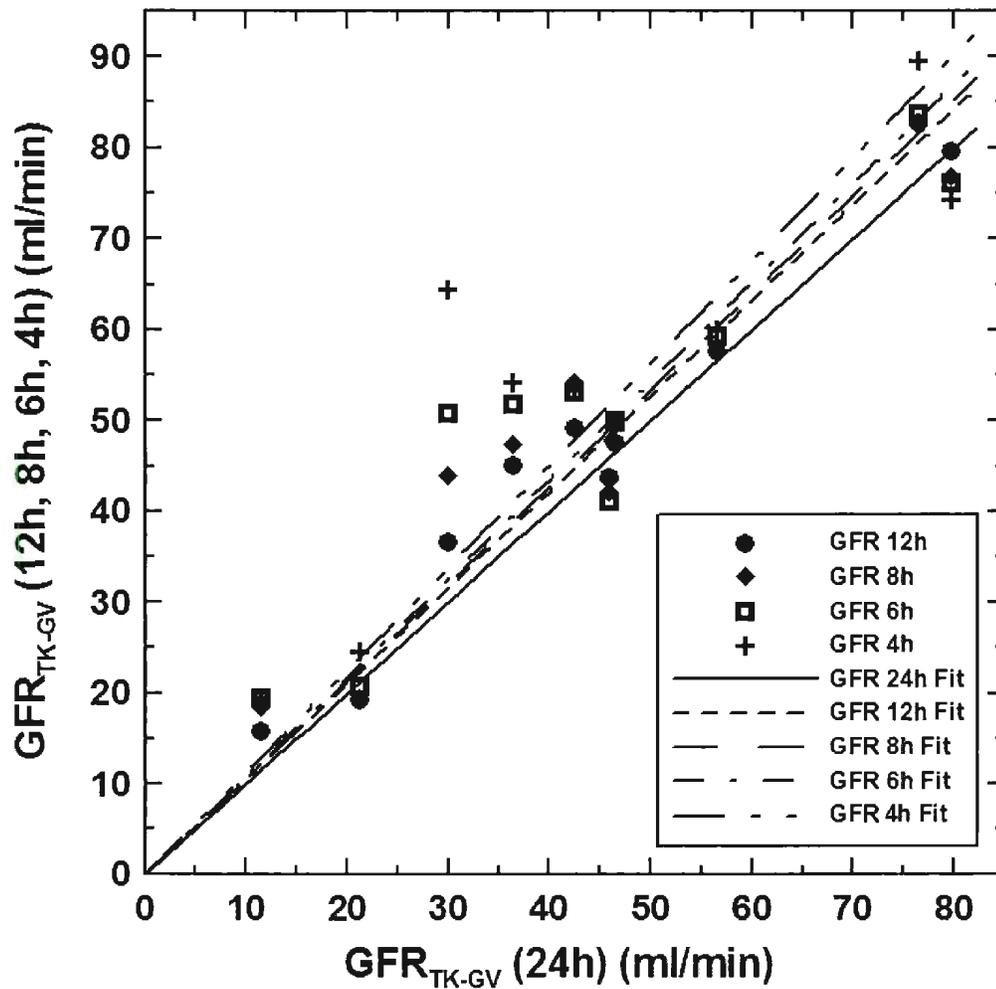


Figure 6.5: Comparison of the 10 estimated GFR values from the Tk-GV model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend.

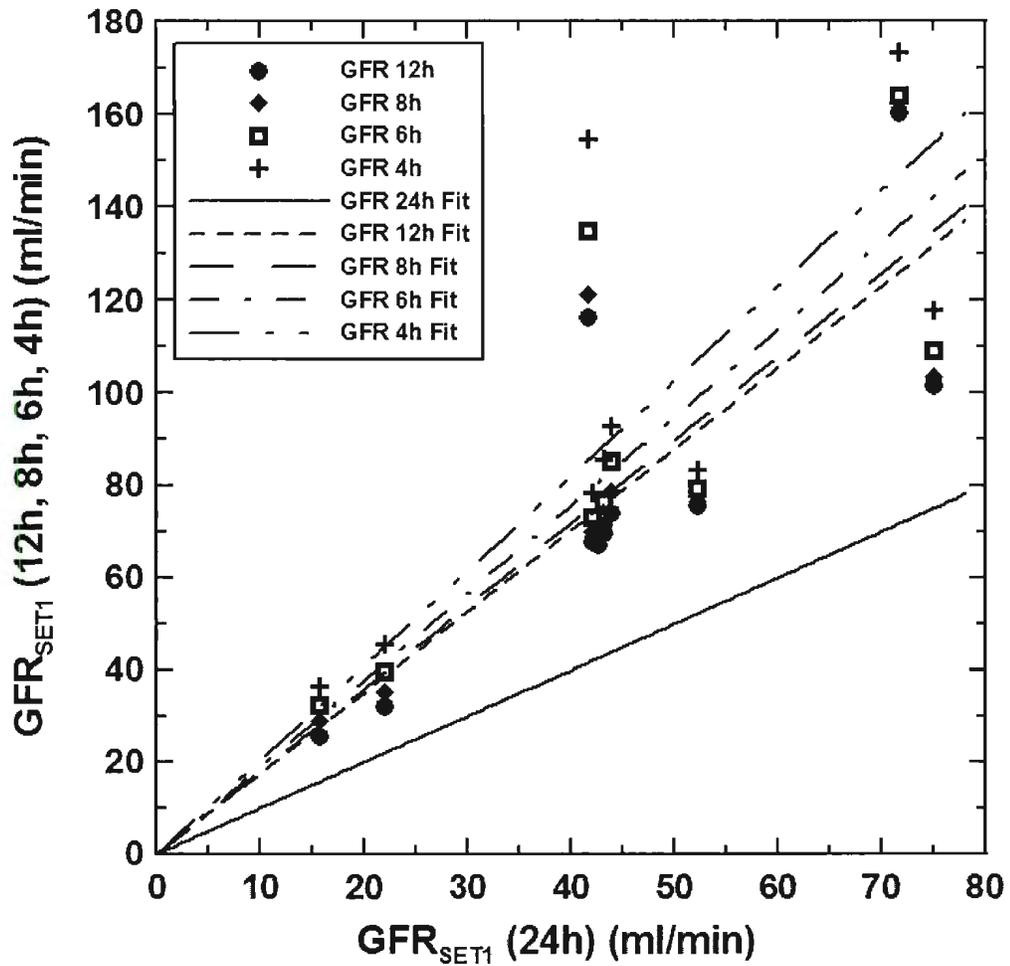


Figure 6.6: Comparison of the 10 estimated GFR values from the $SET1$ model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend.

had the least difference between the results from the 4 subsets (4 h, 6 h, 8 h and 12 h) and the 24 h samples while the *SET1* model showed large variation in *GFR* results, indicating its dependence on the choice of sampling schedule.

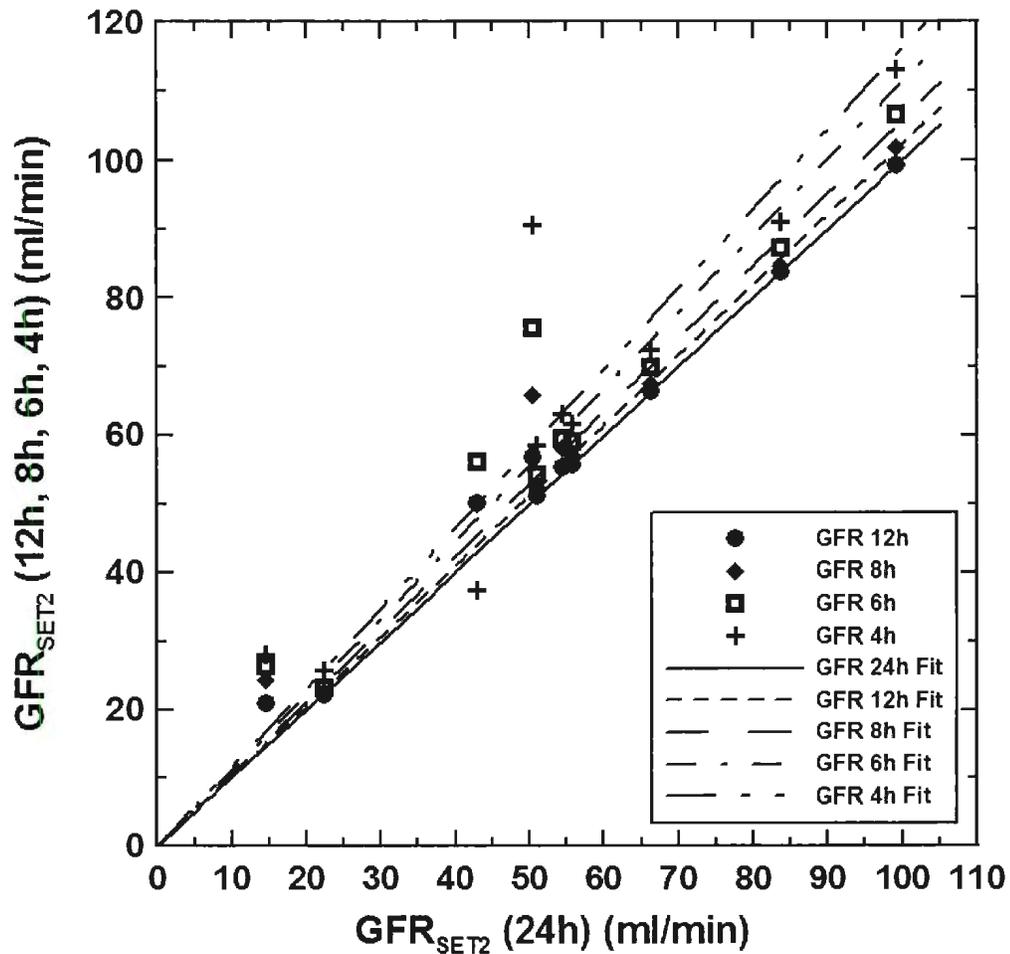


Figure 6.7: Comparison of the 10 estimated *GFR* values from the *SET2* model using 4 h, 6 h, 8 h, and 12 h samples with those using the full 24 h samples. The symbols for each subset are shown in the legend.

In addition to the comparison between the *SET1* model and the Tk-GV model, the *GFR* results obtained from using the *SET2* model with the subsets are shown in Figure 6.7. The *SET2* model performed better than the *SET1* model but poorer

than the Tk-GV model.

Table 6.2: The values of the slope and R^2 from the linear regression of the estimated GFR values obtained from the $SET1$, $SET2$ and Tk-GV models using 4 h, 6 h, 8 h and 12 h sampling schedules to those under the same models using the 24 h sampling schedule for the 10 patients.

models	$SET1$		$SET2$		Tk-GV	
subset	slope	R^2	slope	R^2	slope	R^2
4 h	2.052	0.552	1.163	0.813	1.127	0.619
6 h	1.893	0.617	1.114	0.890	1.087	0.791
8 h	1.796	0.663	1.059	0.944	1.067	0.878
12 h	1.756	0.685	1.025	0.977	1.054	0.961

Table 6.2 lists the slopes calculated from different subsets (4 h, 6 h, 8 h or 12 h) compared with those calculated using the 24 h sampling schedule. Figures 6.5 to 6.7 show that all three models ($SET1$, $SET2$ and Tk-GV) gave better GFR results when more later samples were used. For example, the Tk-GV model gave a slope of 1.127 for the 4 h samples, 1.087 for the 6 h samples, 1.067 for the 8 h samples and 1.054 for the 12 h samples. However, when the 4 h samples were used, the $SET1$ model gave the largest slope of 2.052 compared with that using the $SET2$ model (with a slope of 1.163) and the Tk-GV model (with a slope of 1.127). This indicates that the $SET1$ model introduced the largest difference in estimating the GFR values between the 4 h and 24 h sampling schedules, while the Tk-GV model had the smallest difference. Therefore, using the 4 h sampling schedule would overestimate the GFR values from the 24 h samples by 12.7% using the Tk-GV model, 16.3% using the $SET2$ model and 105.2% using the $SET1$ model. Considering that the 4 h sampling schedule is the most commonly used in clinic, the Tk-GV model remains the best choice for estimating GFR values.

6.4 Exploring the GV Function

It is interesting to observe the contribution of the power and exponential functions in the GV model. The GV model is expressed as $C(t) = Kt^{\alpha-1}e^{-\beta t}$ (Section 2.6.1), where the power function $t^{\alpha-1}$ is used to describe the mixing process of the tracer into the plasma, and the exponential function $e^{-\beta t}$ is used to describe the elimination process.

Figure 6.8 shows the percentages of the two functions (power function and exponential function) as a function of sampling time in minutes for (a) Pt3 whose estimated GFR was 79.7 ml/min and (b) Pt13 whose estimated GFR was 6.4 ml/min (from the Tk-GV model). In both cases, the exponential function (black circles) dominated the early stages of the concentration-time curve. As time increased, the power function slowly contributed more to the overall clearance rate. When the patient's expected GFR was not very low, the two functions met before $t = 24$ h. For patient Pt3, the exponential function reduced its contribution from the initial level of approximately 60% to 50% in 8 h to eventually less than 20% within 24 h period. However, it is not necessary that the two functions exchange their roles before 24 h. For Pt13, the clearance rate was very slow in the initial 24 h period and hence the power function remained the dominant contributor for that period. Figure 6.8 indicates that the clearance curve involves the contributions from both the power function and the exponential function regardless the level of the patients' GFR values. This figure shows that the Tk-GV model can effectively simulate the two components (power function and exponential function) included in the plasma clearance process, and it requires longer time for the exponential function to become dominate during the plasma clearance process for patients with low expected GFR values.

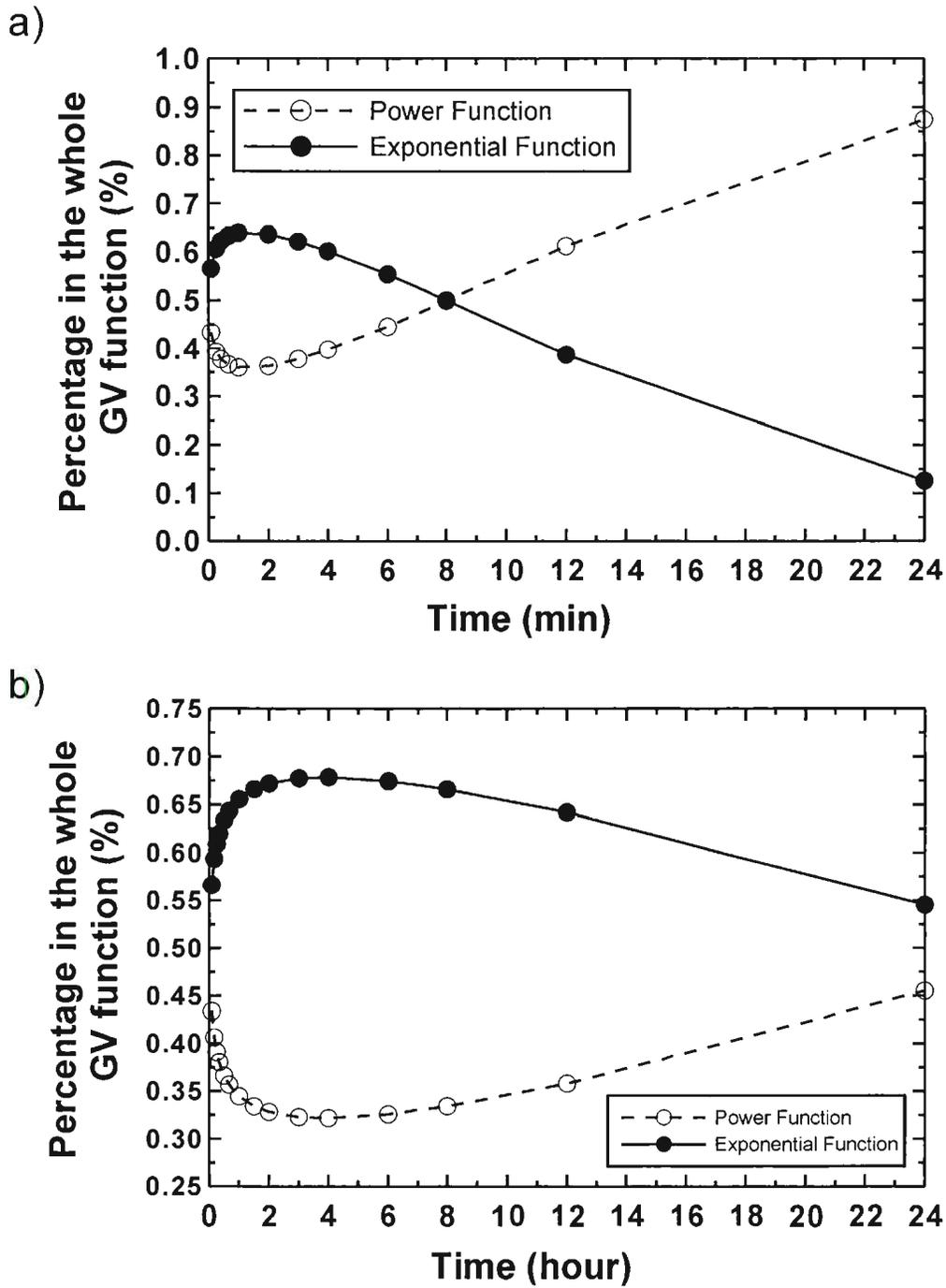


Figure 6.8: Percentage of the power function (hollow circles) and exponential function (solid circles) in the whole *GV* model as a function of sampling time t (with time=0 set at the finishing point of the injection) in min for a) Pt3 and b) Pt13.

6.5 Summary

In summary, the Tk-GV model consistently gave the best estimated *GFR* results compared to those from the *SET1* and *SET2* models, irrespective of the sampling schedules.

Chapter 7

GFR Correction

Although the Tk-GV model has been shown to perform better than the *SET1* model, it is not always possible to collect 4 blood samples, primarily due to the difficulty in drawing multiple blood samples from some patients within 4 hours. In this case, the two-sample *SET1* method is used with a body surface area (BSA) correction for adjusting the estimated *GFR* with respect to different body habitus. In this chapter, a *GFR* correction method initially developed for children is investigated to show whether it can be equally applied to adults.

7.1 Introduction

The body surface area (BSA) is used to adjust the estimated *GFR* values for patients with different body habitus. For example, the correction for the *GFR* value obtained from the *SET1* model is expressed as $GFR_{\text{corrected}} = GFR_{\text{SET1}} \times 1.73 / \text{BSA}$ where 1.73 is in units of m^2 (see Section 2.1.1) [38]. However, current BSA corrections are insufficient for adjusting the overestimation of the *GFR* value caused by the *SET1* model and the $GFR_{\text{corrected}}$ can not accurately indicate the level of patients' kidney

function [38, 92, 93].

Wesolowski et al. developed a method to assess the estimated *GFR* values from children using the 2-sample *SET1* model [132, 133]. The authors introduced a ratio called the renal sufficient index (*RSI*) for analyzing the level of the patient's kidney function instead of using the observed GFR_{SET1} value or the BSA corrected GFR_{SET1} value. *RSI* is defined as:

$$RSI = \frac{\text{estimated } GFR}{f(Vol, W)}, \quad (7.1)$$

where $f(Vol, W)$ estimates the expected *GFR* value for the same patient if he/she is experiencing normal kidney function. The estimated *GFR* refers to the result obtained from the 2-sample *SET1* method. The formula $f(Vol, W)$ obtained by Wesolowski et al, is given by

$$f(Vol, W) = 10.998Vol^{0.64717}W^{0.20185}, \quad (7.2)$$

where *Vol* (in units of L) is the volume of distribution obtained from the 2-sample *SET1* model and *W* is the patient's mass (in units of kg). The coefficients 10.998, 0.64717 and 0.20185 were obtained from a study sample of 133 children [132]. This equation is used in this study as a reference equation. The two factors *Vol* and *W* were chosen over other factors such as height, age, gender or race, etc because the power function $f(Vol, W)$ gave the best agreement with the expected *GFR* results [132]. The initial study performed in 2006 defined the cut-off value for normal *GFR* as $RSI \geq 85.89\%$ based on the Receiver Operating Characteristic (ROC) analysis [132]. If a patient's calculated *RSI* value was larger than or equal to 85.89%, he/she is likely to have normal kidney function. The goal of this study is to assess whether the *GFR* correction method using *RSI* and $f(Vol, W)$ can be applied to adults.

7.2 Data and Analysis Methods

The 46 patients discussed in Chapters 3 to 5 could not be used here because their health information were incomplete in that only the plasma clearance study data and the patients' height and weight were known; unfortunately other information such as medical history, age, gender, race or family history were unknown. Therefore, data from 26 patients, including all relevant health information, were collected from January 2009 to December 2010 in the department of nuclear medicine at the General Hospital in St. John's, NL and used in this study.

The patient data collection was approved by the Human Investigation Committee as HIC 09-64 in 2008. Each patient was injected with ^{99m}Tc -DTPA via a bolus injection. The sampling schedule was 109 ± 10 min and 152 ± 10 min post-injection as recommended by the study in 2006 [132]. Here the error ± 10 min is not an indication of the uncertainty in the timing but instead is meant to account for the difference in the starting time when collecting the samples. For example, due to the vein condition of some patients or the time patients returned to the department of Nuclear Medicine caused a delay in the blood collection time. The uncertainty used in the recorded body mass was approximately 2.0 kg to account for variations in clothing, items in pockets, and the proximity of the mass measurement to the last meal consumed. Given the uncertainty of the body mass and a typical uncertainty for Vol to be ± 0.1 L, the uncertainty in $f(Vol, W)$ is approximately 1%, which is quite small. The uncertainty in the height (used for the BSA calculations below) was taken to be 1 cm to correct for posture and the variation in the spine length with the time of the day. The average age of the 26 patients was 33.34 years with a range from 0.22 to 87.58 years old.

The *GFR* measurements were carried out under the modified laboratory protocol

for preparing and performing kidney function tests using 2 or 4 blood samples in the department of nuclear medicine at the General Hospital in St. John's, NL. All the injections and blood collections were completed by the nuclear medicine technologists Tara Page and Kelly Godin. My contributions to this study were in processing the blood samples, measuring the radioactivity and performing the calculations and data analysis. The injected dose was approximately 30 mCi for adults and 10 mCi for children. This study was not a randomized trial and most of the patients registered for the *GFR* tests were suspected of having kidney diseases or insufficient kidney function. The *GFR* study was performed using the two-sample *SET1* method for each patient. The estimated GFR_{SET1} values were corrected using the BSA Haycock equation ($BSA(m^2) = 0.02425 \times \text{height}(cm)^{0.3964} \times \text{mass}(kg)^{0.5378}$, see Section 2.1.1). As outlined in the initial paper by Wesolowski et al, the *RSI* values based on the BSA Haycock corrected *GFR* value and the $f(Vol, W)$ value were calculated and compared.

7.3 Results and Discussion

Figure 7.1 shows the estimated results from the 26 patients. Figure 7.1 a) shows the *GFR* results using the 2-sample *SET1* model. Figure 7.1 b) shows the estimated $GFR_{Haycock}$ values (dark bars, left ordinate) and estimated $RSI_{Haycock}$ values (white bars, right ordinate). Figure 7.1 c) shows the values of $f(Vol, W)$ (dark bars, left ordinate) and the values of $RSI_{f(Vol, W)}$ (white bars, right ordinate). The dashed line in Figures 7.1 b) and c) represent the 85.89% *RSI* threshold identified in [132]. The solid line in Figure 7.1 b) indicates the Haycock criterion for identifying normal kidney function given by $GFR_{Haycock} \geq 100 \text{ ml/min}/1.73 \text{ m}^2$.

Based on all relevant medical information, patients Pt6, Pt7, P12 and P21 were diagnosed as having normal kidney function. All other patients were diagnosed as

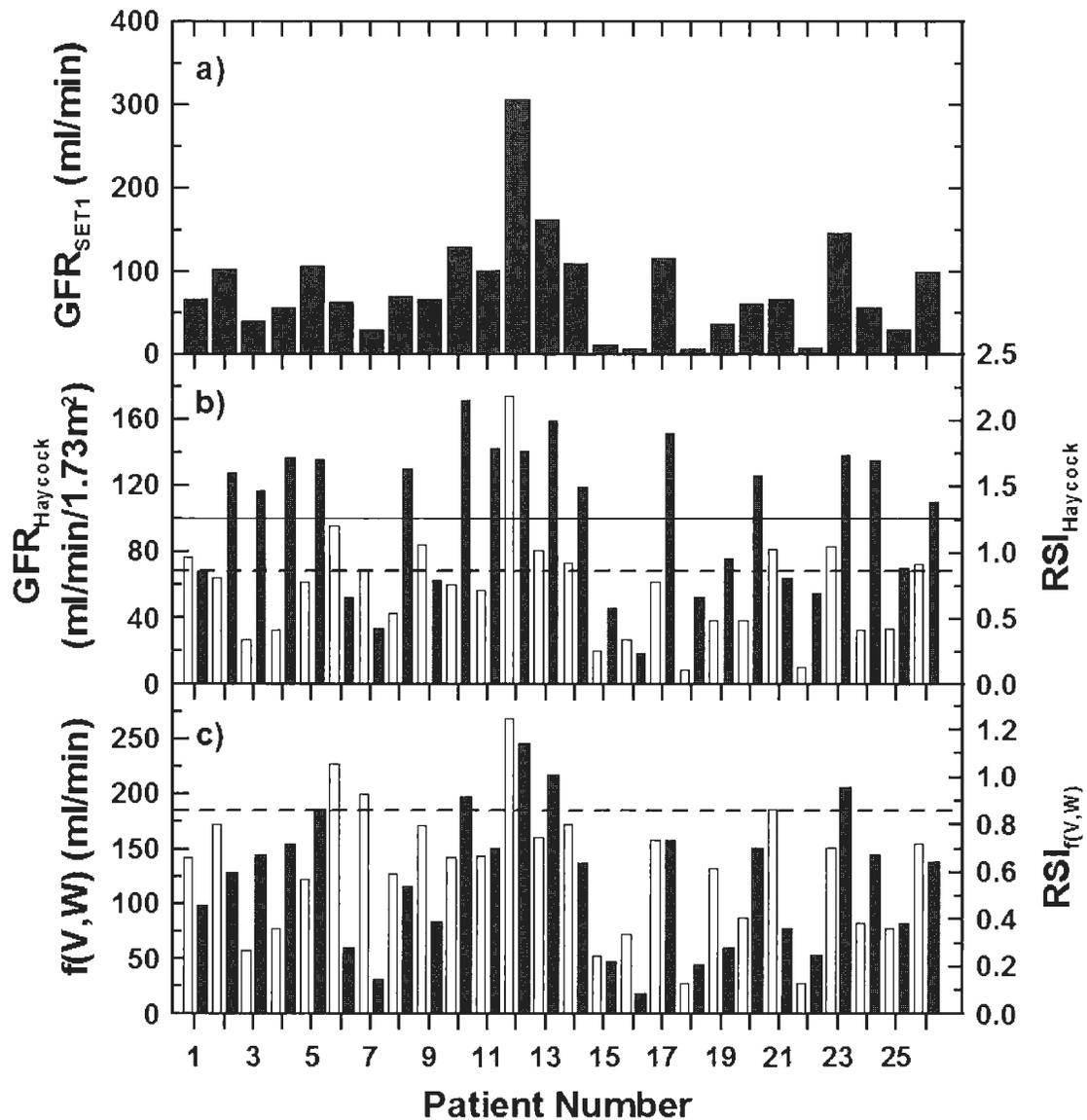


Figure 7.1: a) Estimated GFR_{SET1} results using the 2-sample $SET1$ model (ml/min). b) Estimated results of $GFR_{Haycock}$ (ml/min/1.73 m²) (dark bars) and $RSI_{Haycock}$ (white bars). The solid line indicates the Haycock criterion. c) Estimated results of $GFR_{f(Vol,W)}$ (ml/min) (dark bars) and $RSI_{f(Vol,W)}$ (white bars). b) and c) The dashed line represent the 85.89% RSI threshold.

having abnormal (or insufficient) kidney function. The BSA Haycock corrected *GFR* values identified 15 out of 26 patients as having normal *GFR* results (Figure 7.1 b)). From the RSI_{Haycock} analysis, 10 patients were estimated as having normal kidney function. The analysis based on the $f(\text{Vol}, W)$ corrected *RSI* values, $RSI_{f(\text{Vol}, W)}$, identified only four patients, Pt6, Pt7, Pt12 and Pt21, as having normal kidney functions. These four patients were also properly identified as having normal kidney function using the RSI_{Haycock} criterion.

After considering each patient's health information, the $RSI_{f(\text{Vol}, W)}$ results were more reliable in that they agreed better with the clinical findings from other medical examinations such as the blood serum level of creatinine or the transplant history, etc. Although the RSI_{Haycock} was more accurate than the BSA_{Haycock} value, it still misidentified the level of kidney function for six patients Pt1, Pt9, Pt13, Pt14, Pt23 and Pt26.

For example, Pt1 had only one kidney and the kidney function observed from the renal scan (performed on the same day) clearly indicated a slower than normal plasma clearance. Two previous studies of this patient (one was approximately 5 years old and the other one was two months prior to the kidney function test) showed an obstruction at the ureterovesical junction. However, the GFR_{Haycock} result for this patient was 81.2 ml/min/1.73 m² and the RSI_{Haycock} result was 96.1%. The $RSI_{f(\text{Vol}, W)}$ for this patient was 66.4% with the estimated $f(\text{Vol}, W)$ equal to 98.6 ml/min. Another example is patient Pt9 who has been diagnosed with kidney failure (and has been on dialysis ever since, and was waiting for a kidney transplant). The RSI_{Haycock} result for this patient was 106.3% while the $RSI_{f(\text{Vol}, W)}$ gave 79.5%. Patient Pt13 was found to have renal cysts in the left kidney with a size of 4.0 × 2.5 × 3.1 cm, indicating slightly abnormal kidney function. Another example is Pt26 who had a

mild renal asymmetry to the right, which means that the right kidney was not at the same level with the left kidney. Pt26 had a slightly slower perfusion observed from the renograph. Although these six patients did not have severe kidney diseases, the fact that the $RSI_{f(Vol,W)}$ results provide a better assessment of the patients' renal scans indicates the advantages of using $RSI_{f(Vol,W)}$ over $RSI_{Haycock}$. In addition, if the $RSI_{Haycock}$ results were used for these patients, it could have led to a potential delay in the early diagnosis of mild kidney diseases, therefore, under-diagnosing the patients. Lastly, Pt14 and Pt23 were also reported as having insufficient kidney functions due to the findings from other medical examinations.

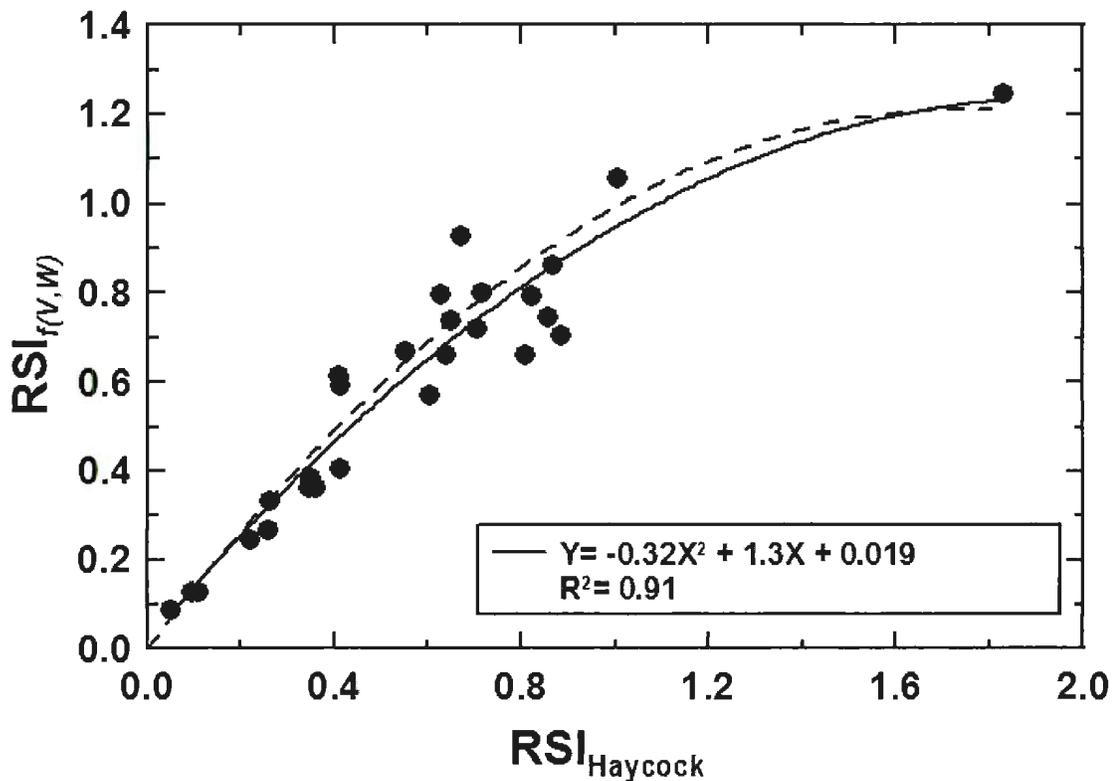


Figure 7.2: Correlation of $RSI_{Haycock}$ and $RSI_{f(Vol,W)}$ results using the linear regression. The solid line indicates the best fit using a second order polynomial from the 26 patients, and the dashed line indicates the fit from the initial study by Wesolowski et al. based on study of 133 children.

The correlation between the RSI_{Haycock} and $RSI_{f(Vol,W)}$ obtained using the linear regression is shown in Figure 7.2. The data were fitted with a second-order polynomial giving

$$RSI_{f(Vol,W)} = -0.32 \times RSI_{\text{Haycock}}^2 + 1.3 \times RSI_{\text{Haycock}} + 0.019 \quad (7.3)$$

with a R^2 value of 0.91 indicated by the solid line. The coefficients in this equation had two significant figures according to my calculation. This regression result was very close to that obtained in the initial study by Wesolowski et al. based on 133 children, which gave

$$RSI_{f(Vol,W)} = -0.4018 \times RSI_{\text{Haycock}}^2 + 1.3953 \times RSI_{\text{Haycock}} \quad (7.4)$$

with $R^2 = 0.9337$, shown as a dashed line in Figure 7.2 [132].

7.4 Conclusion

Using the complete medical information of 26 patients it has been possible to validate the use of the $RSI_{f(Vol,W)}$ criterion for adults. The replacement of $RSI_{f(Vol,W)}$ for the absolute GFR_{SET1} value or GFR_{Haycock} can effectively improve the accuracy for differentiating between normal and abnormal kidney function. However, as promising as this study was, a larger sample study could further validate the conclusion. The application of $RSI_{f(Vol,W)}$ should be recommended to institutions where 4 blood samples are not available for using the Tk-GV model.

Chapter 8

Bone Mineral Density

This chapter introduces a new criterion for analyzing the bone mineral density (BMD) changes from sequential BMD studies over time. The recommended criterion, called the total detectable difference (TDD), is calculated from the mean value of a half-normal distribution (HND) histogram of the standard deviations from the linear regressions of approximately 8,800 patients. The advantage of the TDD criterion is the consideration of long-term machine error, short-term patients' BMD variations, and long-term patients' BMD variations compared with the conventional criterion which only considers the short-term machine error.

8.1 Introduction

The history of using bone densitometry for measuring BMD dates back to the 1990s [134]. At the present, bone mineral density scan results are mainly reported in terms of: 1) the bone mineral density (g/cm^2) in the L2-L4 and L1-L4 sections of the lumbar spine, femoral neck, and total hip, 2) T-score, and 3) Z-score. T-score is calculated as the number of standard deviations (*SDs*) above or below the mean value of BMD

for a healthy young adult of the same gender and ethnicity. T-score is used to predict osteoporosis by the World Health Organization (WHO) [135]. If the T-score is higher than or equal to -1.0 (1 standard deviation below the reference BMD value of young adults, where the reference value was obtained from a large group of people in the United States in 1990 [135]), the patient's BMD is reported as normal with a 95% confidence interval. If the T-score is between -1.0 and -2.5, the patient is reported to have osteopenia (a precursor to osteoporosis) and if the T-score is below -2.5, the patient is reported to have osteoporosis. Z-score is defined as the number of standard deviations by which a patient's BMD differs from the reference BMD value for a comparable age, gender and ethnicity. The difference between the T-score and Z-score is whether or not the reference BMD value is age adjusted.

BMD results are important for evaluating the patients' bone health, especially for elderly people such as post-menopausal women, and men with prostate disease. In general, women (due to their decreasing estrogen levels) experience more rapid bone mineral density loss than men. As a result, an annual or bi-annual BMD test is usually recommended for women over the age of 50 years old. From the repeated BMD examinations over time, patients have a series of their BMD results. The difference among these sequential BMD results for the same patient are calculated and analyzed in order to determine whether the BMD changes are normal or not.

How to determine whether a BMD change is significant or not compared with the baseline (or expected BMD change) is critical for predicting the bone fracture risks of patients. The conventional criterion for analyzing BMD change, as provided by the WHO, is called the Least Significant Change (LSC). LSC is defined as 2.8 times the site-specific (herein, site refers to the health institution) precision error obtained on the same day [136]. The factor 2.8 is obtained from $1.96 \times \sqrt{2}$ where 1.96

indicates the 95% confidence level of LSC and $\sqrt{2}$ indicates the conversion between standard error (SE) and standard deviation (SD) with a sample size of $n=3$, given as $SE = SD/\sqrt{n-1}$ where the sample size of 3 is recommended from the guidelines of the WHO on how to calculate the LSC [135]. LSC indicates the machine error of the bone densitometer. Precision error is related only to the machine error, which means that it contains no information of the patients' natural BMD changes over time. The machine error is generally understood as the error caused by limitations of the instruments instead of an error caused by human operation (known as human error). In this study, the machine error is estimated using repeated studies of the same study group. The WHO recommends the machine error to be calculated as the standard deviation of the standard deviations of repeated BMD measurements on the same group of patients. For example, in order to achieve a 95% CI level, the machine error can be calculated using 15 patients, each of which had 3 repeated BMD studies taken on the same day, or using 30 patients, each of which had 2 repeated BMD studies taken on the same day. For each patient, there is a standard deviation of repeated BMD values (2 or 3 depending on the number of patients used). The machine error is based on the standard deviation of these 15 or 30 standard deviations. If the LSC value is not calculated for the bone densitometer machine, reference LSC values can be used, which are 0.028 g/cm² for the femur at the hip region, and between 0.033 and 0.039 g/cm² for the lumbar spine (these values were obtained from a reference group of age 20-40 from the United States.) [137, 138]. That means, for example, if the patient's BMD change for the femur is more than 0.028 g/cm² compared with the previous BMD study, he/she is considered to have significant bone mineral density change.

There are two problems regarding the use of the LSC criterion. Firstly, the refer-

ence LSC values do not include the patients' natural BMD variations. Bone mineral density varies due to many factors under normal conditions, such as genetic, hormonal, physical activity, and dietary factors [139, 140]. Despite all these factors, LSC accounts only for the same day machine error. Therefore using LSC tends to overestimate BMD changes. For example, if a woman has been nursing during the time interval of her BMD tests, her bone mineral density loss is likely to exceed the LSC value despite the fact that she is experiencing normal BMD change with respect to her age and health condition. Therefore, unless the BMD change is larger than what the normal range of this specific age group or gender allows, it cannot be labeled as a significant change. Secondly, despite that the LSC only accounts for the machine error, the LSC is calculated based on a relatively small study group (size 15 or 30 patients) which is insufficient to represent the criterion for the general population. The practice of using the LSC as a threshold for BMD change can lead to a significant misdiagnosed patients with abnormal BMD changes [136].

8.2 Literature Review

The development of the mathematical tools for analyzing the BMD results has experienced several stages. In 1994, the World Health Organization announced a protocol for assessing fracture risk of patients using the T-score to analyze individual BMD result. In 2004, the International Society for Clinical Densitometry (ISCD) recommended using serial BMD tests to assess a 10-year prediction of the probability of fracture risk [135, 141]. In 2005, the national guidelines from the ISCD and the WHO suggested interpreting BMD results by comparing with the previous BMD results from the same patient in addition to the use of T-score to predict the patients' fracture risk. Ever since then, the conventional T-score and Z-score analysis has been

gradually replaced by the fracture risk prediction based on BMD changes [142, 143]. The LSC calculated from the same-day machine error was recommended to define clinically significant BMD changes [136, 144]. The ISCD also recommended that each imaging center obtain their own LSC reference values for different biological sites for assessing BMD values. In the case where this is not possible, imaging centers can use the benchmark LSC values calculated from the ISCD as their reference LSC values. For example, the reference LSC values from the ISCD are approximately 0.020 g/cm^2 ($\approx 2\%$ of the BMD value for the spine or hip) for both the spine and the hip. This means that if a patient has a BMD change over 0.020 g/cm^2 compared with the previous BMD value, he/she is likely experiencing a significant BMD change. Suspicions of the validity of the clinical application of LSC have been reported since 2000. There have been studies showing the inaccuracy of the LSC values (mainly that the LSC values have been underestimated) and the statistical method used to calculate them [145]. Two major attempts for adjusting the LSC criterion are summarized below.

One attempt focused on determining the proper sample size required for calculating the LSC values as mentioned previously. The WHO recommended a sample size of 15 patients (each having 3 BMD tests taken on the same day) or 30 patients (each having 2 BMD tests taken on the same day) [135]. In 2006 W. D. Leslie from the Manitoba Research group pointed out that a larger sample size should be used for assessing the LSC value [146]. The Leslie group studied 1,420 BMD cases and concluded that the estimated LSC values from these 1,420 cases were significantly larger (approximately 140.1% times the reference LSC) than the reference LSC values. The same research group later showed that a long-term study (not the same day study as recommended by the WHO) of the GE DXA densitometer gave LSC

values for the lumbar spine and femur of at least 0.055 g/cm^2 , which for the lumbar spine, is approximately twice that of the reference LSC value. W. D. Leslie's group performed 10,000 Monte Carlo simulations on a phantom study in 2007 [147]. Their results showed that the small sample size (15 or 30) used for calculating the reference LSC value resulted in an up to 12.5% over detection of the BMD change for the spine or hip.

The second attempt aimed to modify the statistical method used for calculating the LSC. A bone-phantom study performed with 21 repeated BMD studies during one year (instead of doing repeated studies within one day as recommended by the WHO) showed that the obtained LSC value in consideration of long-term patient's BMD variation was on the order of 0.050 g/cm^2 . In 2008, W. D. Leslie's group employed a heteroscedastic regression to study the BMD changes, and determined that it was impossible to separate the long-term changes in the patients from the errors due to the equipment from their measurements [148]. They also proposed to use a larger sample size for improving the accuracy of the estimated LSC value.

Some studies have proposed different mathematical models for calculating the LSC value. For example, one study recommended using a 90% or 80% confidence interval (see definition in Section 3.4) for calculating the LSC value instead of the 95% confidence interval recommended by the WHO [149]. Some groups have recommended standardizing the LSC value with a correction factor from the cross-calibration of the BMD values obtained from different bone densitometers. For example, Ying Lu et al. derived a criterion called the Generalized Least Significant Change (GLSC) by cross-calibrating the BMD results obtained from bone densitometers manufactured by GE and Hologic [150]. The obtained GLSC value was approximately 2.6 times larger than the reference LSC value for the spine, and 3.6 times larger than the reference

value for the femur. There have been different attempts to analyze the sequential BMD results in order to obtain a LSC estimation. For example, in 2008 W. D. Leslie developed an “optimal decision criterion” using a Bayesian approach multi-variable regression analysis and obtained an LSC value of 0.041 g/cm^2 for the spine and 0.035 g/cm^2 for the femur [148].

In 2008, the WHO released the latest development for predicting the fracture risk from BMD studies, the fracture risk assessment tool (FRAX). FRAX became available in Canada in 2010 [151–153]. FRAX was developed in order to account for the many factors affecting bone health. The FRAX result, which is personalized to each patient, is obtained from a mathematical model integrating all possible risk factors associated with bone health at the femoral neck at the hip region. The FRAX tool predicts the 10-year fracture risk by assessing the patients health information such as gender, race, height, weight, previous fracture, smoking and alcohol history, the femoral neck BMD value, etc. The limitations of the FRAX tool is that it is recommended only for people aged 50 years old or more for estimating the probability of hip fracture and it does not predict the fracture risk for the other biological sites of the body such as the lumbar spine or forearm. Therefore, while using FRAX might improve the accuracy of predicting the fracture risk for the femoral neck at the hip region, it has not appropriately solved the following problem: how to compare two or more BMD studies obtained from the same patient over time.

8.3 Motivation

Despite the different approaches for adjusting the LSC, the definition of LSC criterion has never been questioned since its recommendation by the WHO [142,154]. The motivation of this study is to find a more appropriate criterion for analyzing BMD

changes. The proposed criterion, Total Detectable Difference (TDD), has been calculated by considering both short-term and long-term BMD changes with respect to the machine error and the patient's natural BMD variations using a patient database spanning 11 years. The hypothesis is that TDD is a more appropriate criterion for analyzing the BMD change compared with LSC. The findings from this study might lead to the modification of the BMD guidelines for interpreting BMD changes used worldwide.

8.4 Total Detectable Difference

It is necessary to clarify the difference between the *statistical significance* and the *clinical significance* before introducing the TDD criterion. Statistical significance can identify what is theoretically significant, however, the clinical significance has the difficulty of defining what is the smallest significant change that matters in clinical practice, for example, from the point of view of physicians in diagnosing certain diseases. LSC is calculated based on the statistical significance from the performance of the bone densitometer machine, while a more appropriate criterion should be invented for the clinical significance which accounts for overall factors that affect the clinical outcomes.

The Total Detectable Difference (TDD) criterion proposed was based on the histogram analysis of a group of patients' BMD results. Each patient had three sequential BMD results over the period of 10 years. An OLS regression was performed for each patient to obtain the standard deviation of the regression residual. All these standard deviations were plotted together as a histogram and then fitted by a half-normal distribution. The dispersion of the half-normal distribution is used to calculate the TDD value, which accounts for all the BMD variations from the machine and patients

over short- and long-term study (see more details in Section 9.3).

Figure 8.1 demonstrates the distribution of machine errors and patient's BMD variations for both the short-term and long-term based on the data obtained from local study and literatures. The numerical values for each category in Figure 8.1 were approximations. The short-term machine error is obtained from the study completed in 2006 based on a sample of 15 patients, each of which had 3 repeated studies of their femoral neck BMD over time performed in the department of nuclear medicine at the General Hospital in St. John's, NL. The value of the short-term machine error is approximately 0.010 g/cm^2 (shown as 20% in the gray bar in Figure 8.1). The long-term machine error was calculated as 0.010 g/cm^2 (shown as 19.9% in the darker bar in Figure 8.1) from a 10-year phantom study on the same densitometer machine (GE Lunar Prodigy). The long-term patients' BMD variation is approximately 0.020 g/cm^2 per year (shown as 60% in the lighter bar in Figure 8.1) from the reference [155]. Based on the long-term patients' BMD variation, the short-term patients' BMD variation was then estimated as $0.020/365 \approx 0.00006 \text{ g/cm}^2$ (shown as 0.1% in the white bar in Figure 8.1) from the same reference [155]. The consideration of all four variations gives a value of $0.010 + 0.010 + 0.00006 + 0.020 \approx 0.042 \text{ g/cm}^2$.

According to the WHO recommendation, the LSC values is determined as 2.8 times the short-term machine error, herein, $2.8 \times 0.010 = 0.028 \text{ g/cm}^2$. Based on Figure 8.1, using the LSC to represent the BMD expected change has a tendency of underestimating the total BMD detectable difference (0.042 g/cm^2) by approximately $(0.042 - 0.028)/0.042 \approx 30\%$. Therefore, it is important to understand the limitation of the LSC criterion. In the next chapter, the new TDD criterion is investigated with a study group of 8,800 patients.

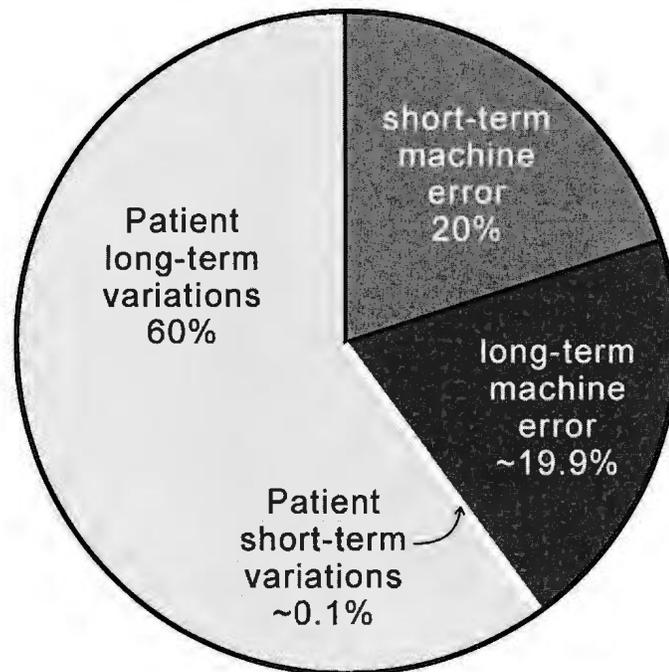


Figure 8.1: Demonstration of the composition of total BMD changes including short-term and long-term machine errors, as well as patients' BMD variations. The short-term and long-term machine errors were estimated from local study while the short-term and long-term patients' BMD variations were estimated from the reference [155].

Chapter 9

Data and Analysis Methods

This chapter introduces the study group and the analysis methods used to investigate the new TDD criterion. A discussion of how the half-normal distribution (HND) is generated from the normal distribution is also presented.

9.1 Data

This is a retrospective study which includes only secondary data usage without any clinical examinations. Therefore, since there was no radiation exposure required, only an ethical approval for secondary data usage was required. This study was approved by the human investigation committee in 2009 with document number HIC 10-57 (for the secondary data usage). The data were obtained from the department of nuclear medicine in both the General Hospital and the St. Clare's Mercy Hospital in St. John's, Newfoundland and Labrador. The data were exported from the GE Lunar densitometer (General Hospital) and the Hologic QDR 4500 mode densitometer (St. Clare's Mercy Hospital). A total of 8,800 patients with three sequential BMD studies performed from 1998 to 2009 were studied. Among these patients, approx-

imately 5,500 were from the General Hospital and approximately 3,300 were from the St. Clare's Mercy Hospital. These 5,500 patients were the primary data used for investigating the TDD criterion while the 3,300 patients were used for performing a comparison analysis between the GE and Hologic densitometers. Daily quality assurance and weekly calibration were performed on the GE Lunar Prodigy densitometer using an anthropomorphic spine phantom. An 11-year phantom study showed the precision of the GE Lunar Prodigy densitometer machine to be within an acceptable range with a coefficient of variance value less than or equal to 1.5%. Similarly, the Hologic QDR 4500 has been reported to have an error of less than 1.0%, also obtained through a phantom study. Calculations were performed using Mathematica version 6 software.

9.2 Analysis Methods

The patients studied in this work each had three BMD studies (taken in different years) within an 11-year period. A sample size of 3 BMD tests is the minimum requirement for performing a linear regression using the OLS algorithm. The residuals indicate the deviations between the observed BMD values and the predicted BMD values from the OLS linear regression. All the SD values were then used to create frequency distribution histograms to be fitted with the half-normal distribution (HND) function. The Normal Distribution (ND) has been used by other investigators for the study of BMD [97]. However, HND was used in this study instead of the ND because the observed distribution of non-negative SD values was heavily skewed from the ND indicating that an alternative distribution from the ND should be used.

9.2.1 Half-normal Distribution

For each patient, there was a standard deviation SD of the residuals from the OLS regression of three BMD values. Because of the small samples, the OLS regression is highly biased. For the BMD data from the GE Lunar Prodigy densitometer, there was a total of 5,500 calculated SD values (g/cm^2). A histogram was used to plot the probability density function of these SD values. Twenty successive bins ranging from $SD=0$ to $SD=1 \text{ g}/\text{cm}^2$ with an interval of $SD=0.005 \text{ g}/\text{cm}^2$ were used to map the SD values. The number of bins and the width of the bins were chosen to ensure the representation of the shape of the frequency distribution of all the SD values [156]. The frequency was calculated as the portion of the number of SD values per bin to the total number of SD values. Plotting the SD values using this method allows the data to be fit to a known distribution function. Therefore, the statistical properties of this distribution function can be used for describing the histogram. However, the obtained frequency distribution of the standard deviations did not follow a normal distribution. The common assumption is that the frequency should be a maximum at the middle bin and a minimum at the two end bins $SD=0$ and $SD=1 \text{ g}/\text{cm}^2$. Instead, the maximum frequency was at (or near) the first few bins and the frequency decreased as the SD values of the bins increased. In other words, the histogram of all the calculated SD values from the residuals was heavily skewed to the right. Therefore, non-normal distributions were used to fit the frequency histogram, such as the folded-normal distribution (FND), half-normal distribution (HND), the Raleigh distribution, etc. From this it was found that the FND and HND gave the best fit to the frequency histogram. The following table lists the probability density functions of the ND, HND, and FND.

Figure 9.1 shows a schematic representation of how a FND is generated from

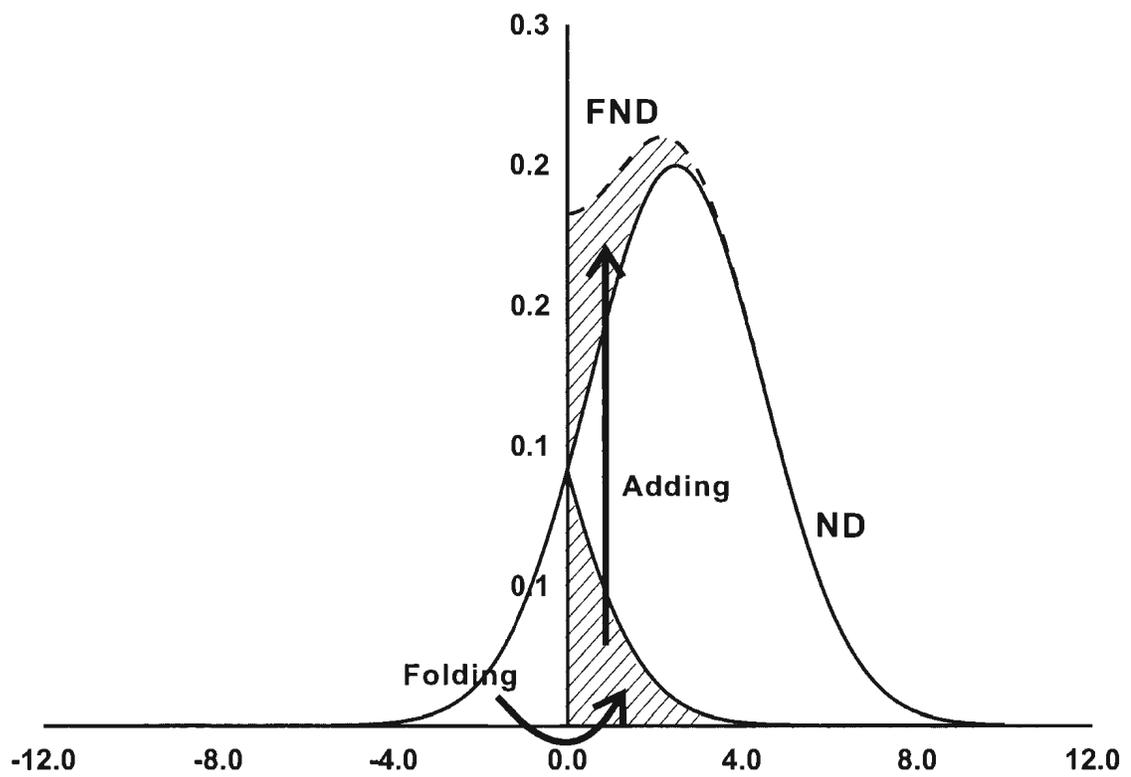


Figure 9.1: Demonstration of the generation of the folded-normal distribution (FND) from the normal distribution (ND).

Table 9.1: The probability density function (*PDF*) of the folded-normal distribution (FND), half-normal distribution (HND), and normal distribution (ND), where μ is the mean value and σ is the standard deviation.

Function	<i>PDF</i>
ND	$\frac{1}{\sqrt{2\pi}\sigma_{\text{ND}}} \exp\left(-\frac{(x - \mu_{\text{ND}})^2}{2\sigma_{\text{ND}}^2}\right)$
HND	$\frac{2}{\sqrt{2\pi}\sigma_{\text{HND}}} \exp\left(-\frac{x^2}{2\sigma_{\text{HND}}^2}\right)$
FND	$\frac{1}{\sqrt{2\pi}\sigma_{\text{FND}}} \left(\exp\left(-\frac{(x - \mu_{\text{FND}})^2}{2\sigma_{\text{FND}}^2}\right) + \exp\left(-\frac{(x + \mu_{\text{FND}})^2}{2\sigma_{\text{FND}}^2}\right) \right)$

folding a ND (bell-shape curve indicated by a solid line) and then adding the folded area (shaded) to the unfolded area of the ND. If the folding point is in the middle of the ND (e.g. approximately $x = 3$ in Figure 9.1), the FND is equivalent to a HND. From the *PDF* expression in Table 9.1, the HND is generated when the mean value $\mu_{\text{FND}} = 0$ in the FND. Using the HND requires the use of a conversion factor between the SD_{HND} and SD_{ND} , giving $SD_{\text{ND}} = \sqrt{\pi/2} \times SD_{\text{HND}}$.

Another concept used here for indicating how much of the ND is folded to generate the FND is called the cumulative distribution function (*CDF*). *CDF* describes the area under the *PDF* curve from $-\infty$ to the folding point x , given by $CDF(x) = \int_{-\infty}^x PDF(t)dt$ [97].

9.2.2 Small Sample Correction

A small sample bias is caused by an insufficient sample size that leads to non-representational estimated parameters in the study group. In this project, a minimum sample size of 3 was required for processing a linear regression. Therefore, it is necessary to add a bias correction. The sample size correction is $\sqrt{(n-1)/(n-p)} = \sqrt{2}$, where $n = 3$ is the number of samples and $p = 2$ is the number of parameters [157]. Considering both correction factors $\sqrt{\pi/2}$ (between the HND and ND) and $\sqrt{2}$, the overall correction was $\sqrt{\pi/2} \times \sqrt{2} = \sqrt{\pi}$.

9.3 TDD Calculation

A flow chart in Figure 9.2 lists the steps for calculating the TDD values from the patient's original BMD data. At the onset, each patient is subjected to three BMD studies over time which are labeled as $(\text{BMD}_1, \text{BMD}_2, \text{BMD}_3)$ taken at time (t_1, t_2, t_3) . An OLS regression is performed for each patient and the standard deviation of the regression residuals SD_i is calculated as the difference between the estimated BMD values and the observed BMD values where the estimated BMD values are obtained from the OLS regression. Each patient in the study group will generate one value for SD_i . The estimated SD_i values ranged from 0 to 1. The frequency histogram is then fitted with a half-normal distribution (HND) and the dispersion of the HND is calculated using the standard deviation of the HND. Using the total correction factor derived above the formula, for calculating TDD is given by:

$$\text{TDD} = 2\sqrt{\pi}\text{SD}_{\text{HND}} \quad (9.1)$$

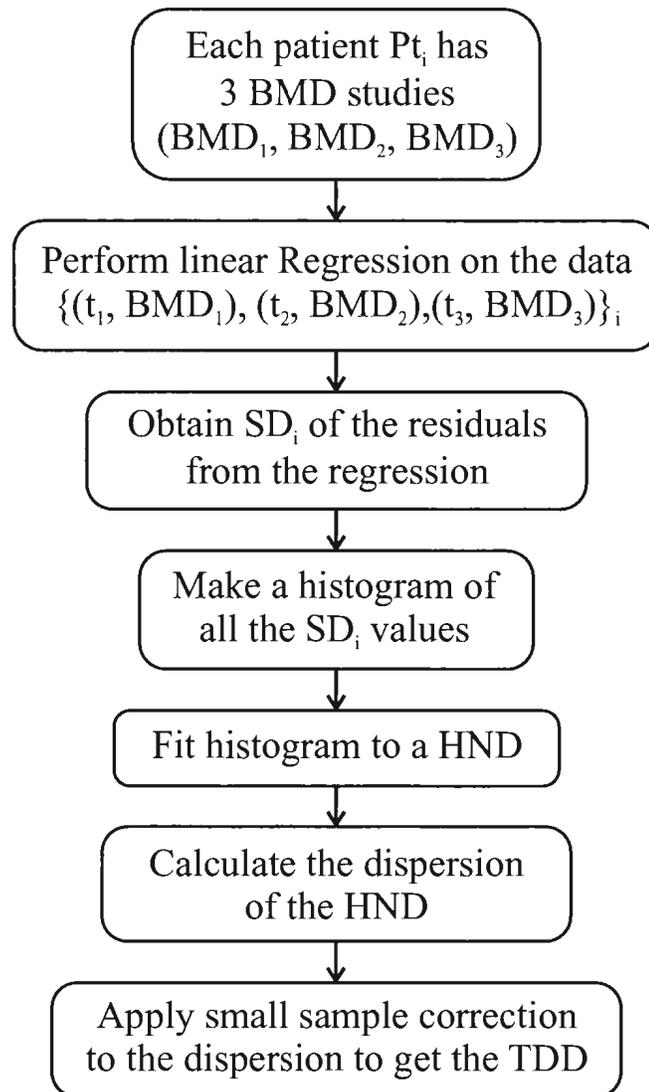


Figure 9.2: A flow chart indicating how the TDD value is calculated from the initial patient's dataset.

where SD_{HND} is the standard deviation of the half-normal distribution and the factor 2 indicates that the TDD value is estimated at a 95% CI. The analysis of the 8,800 patients is presented in the next chapter.

Chapter 10

Results and Discussion

This chapter lists the results and discussion of the performance of the TDD criterion compared with the LSC criterion. The investigation of the TDD values with respect to gender, age, examination time interval, different bone densitometer machine, and different database is given in the following sections.

In the section below we will present data obtained using the GE Lunar Prodigy densitometer at the General Hospital in St. John's, Newfoundland and Labrador (NL) with a sample size of approximately 5,500 patients. In Section 10.5 these values will be compared to the TDD values obtained from the 3,265 patients analyzed with the Hologic QDR 4500 densitometer. The TDD is calculated based on the dispersion of the half-normal distribution, and the 95% confidence interval (CI) is for the dispersion, instead of for all the SD s calculated for all patients. The presented TDD values were $TDD \pm 2 SE$ where the SE is the standard error, calculated as $SE = SD/\sqrt{n}$ with SD being the standard deviation, and n the sample size. Both the SD and SE are in units of g/cm^2 .

10.1 Results

This section gives the complete TDD results using both the HND and FND for four examination sites.

Table 10.1: TDD results (g/cm^2) of approximately 5,500 patients analyzed using the GE Lunar Prodigy densitometer at the General Hospital using the HND and FND. The R^{2*} (see the definition below) value is for both the HND and FND. Area indicates the folded area of the FND from the ND.

Site	Patients	R^{2*}	TDD_{HND}	TDD_{FND}	area
L2-L4	5,534	0.997	0.063 ± 0.002	0.063 ± 0.002	0.499
L1-L4	5,495	0.998	0.059 ± 0.002	0.059 ± 0.002	0.499
Femoral neck	5,483	0.995	0.047 ± 0.001	0.047 ± 0.001	0.499
Total hip	5,359	0.997	0.042 ± 0.001	0.042 ± 0.001	0.499

Table 10.1 shows the TDD values calculated for the lumbar spine L1-L4, lumbar spine L2-L4, femoral neck, and total hip, using the GE Lunar Prodigy densitometer. Both the HND and FND achieved similar R^{2*} values (R^{2*} is the adjusted R^2 for different sample sizes used for the OLS regression of the BMD values as function of the examination time) (>0.99) from fitting the frequency histograms[†] In all cases, a folded area of the FND (calculated from the CDF of the FND) close to 0.5 (a half folding) of the ND was required, indicating the equivalence of the FND to the HND. The sample size used for each examination site was very similar to each other (e.g. 5,483 for the femoral neck study and 5,534 for the lumbar spine L2-L4 study).

The TDD values from the HND and FND are shown with a standard error of $\text{TDD} \pm 2SE$, indicating a 95% CI for all four examination sites. The TDD values using the GE Lunar Prodigy densitometer were $0.063 \pm 0.002 \text{ g}/\text{cm}^2$ for the lumbar

[†]The equation is given $R^{2*} = 1 - (1 - R^2) \frac{n-1}{n-p-1}$ where n is the number of samples and p is the number of independent variables [158].

spine L2-L4, 0.059 ± 0.002 g/cm² for the lumbar spine L1-L4, 0.047 ± 0.001 g/cm² for the femoral neck and 0.042 ± 0.001 g/cm² for the total hip. The TDD values are different for different examination sites. Given that BMD values for the lumbar spine and hip are on the order of 1.000 g/cm², the TDD values are approximately 5% of the absolute BMD values, indicating that a BMD change of over 5% should be considered detectable.

The frequency histograms fitted with the HND are shown in Figure 10.1 a) for the lumbar spine L1-L4, Figure 10.1 b) for the lumbar spine L2-L4, Figure 10.2 a) for the femoral neck, and Figure 10.2 b) for the total hip. The solid dots in Figures 10.1 and 10.2 represent the original data, and the solid line indicates the fit using the HND. All four figures show good agreement between the observed frequency histograms and the HND fit, with R^{2*} values ≥ 0.995 .

10.2 Gender

This section compares the TDD results from females and males, in order to determine whether the TDD values are gender specific.

Table 10.2 shows the calculated TDD_{HND} and TDD_{FND} values from the 5,500 patients using a GE Lunar Prodigy densitometer. The sample size of the male group was approximately 6.8% that of the female group. The TDD_{HND} and TDD_{FND} values consistently showed no difference for the same examination site (both having high R^{2*} values, and the folded area for the FND calculated from the *CDF* was very close to 0.5), indicating that the FND is equivalent to the HND.

The TDD values for the female group were generally smaller than those for the male group. For example, the TDD_{HND} value for the lumbar spine L2-L4 from the

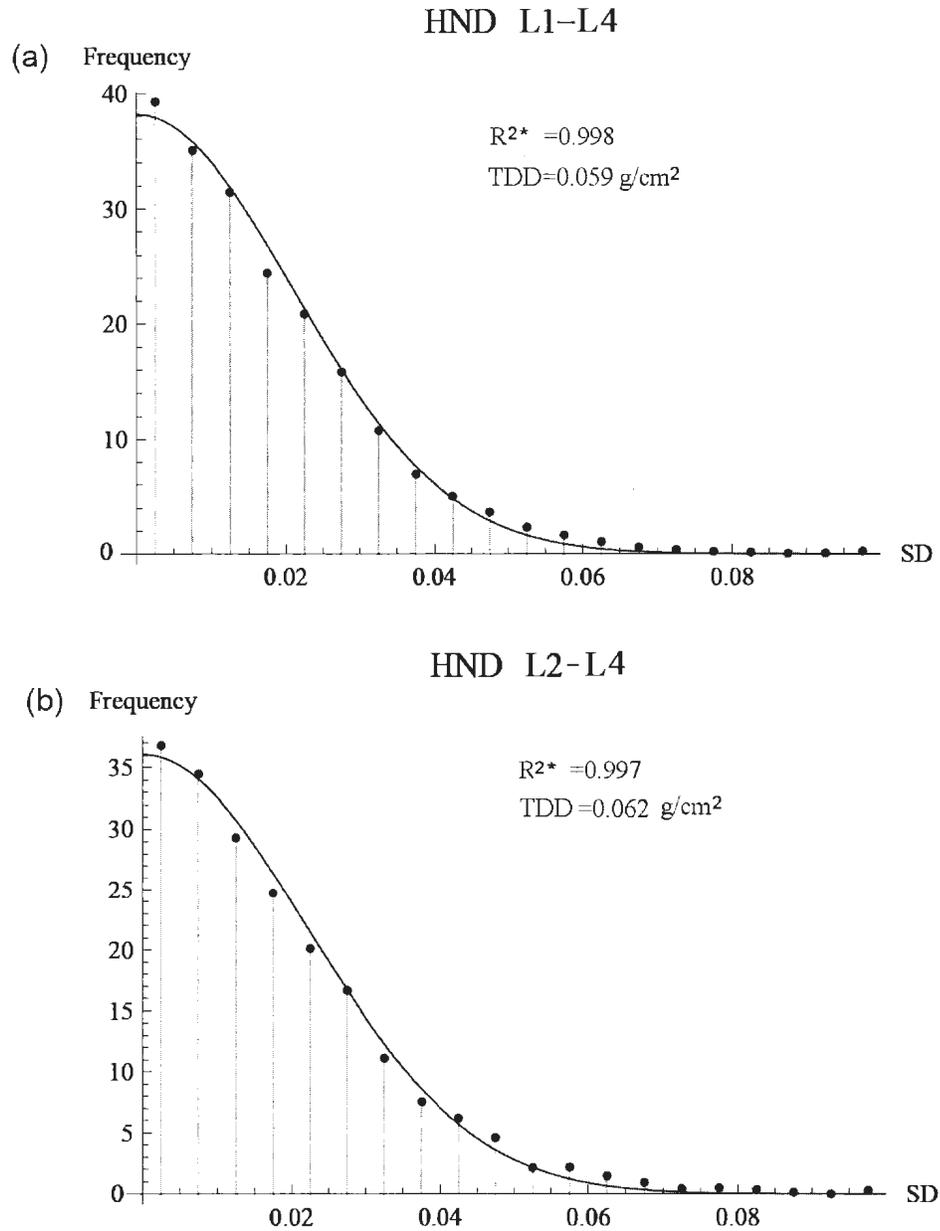


Figure 10.1: The HND fitting for patients with three sequential BMD studies at the examination site of (a) the lumbar spine L1-L4 and (b) the lumbar spine L2-L4, using a GE densitometer.

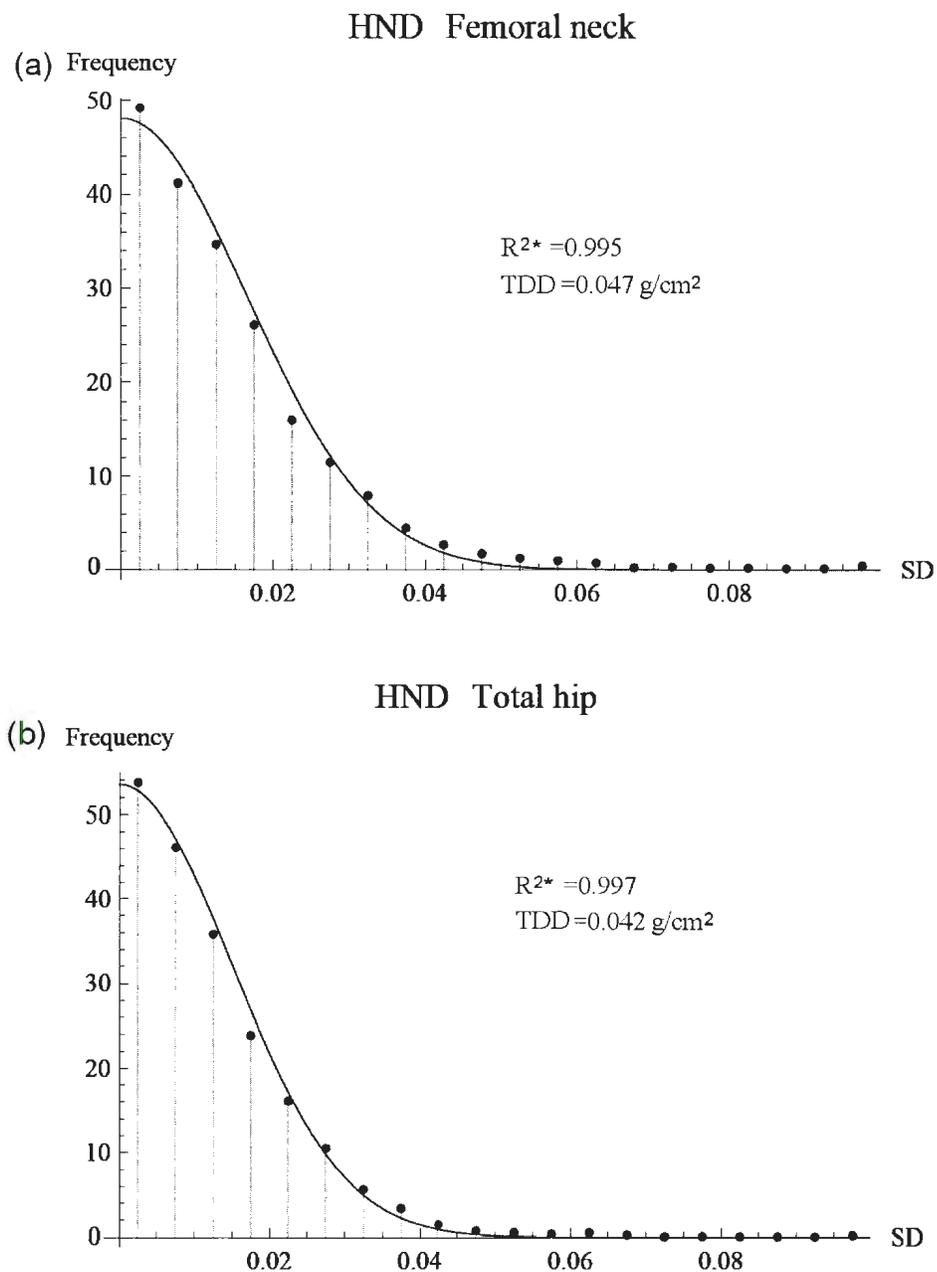


Figure 10.2: The HND fitting for patients with three sequential BMD studies at the examination site of (a) the femoral neck and (b) the total hip, using a GE densitometer.

Table 10.2: The TDD (g/cm^2) values of the female and male groups, for HND and FND, using a GE Lunar Prodigy densitometer.

Female	Patients	TDD _{HND}	TDD _{FND}
L2-L4	5,183	0.062 ± 0.002	0.062 ± 0.002
L1-L4	5,146	0.059 ± 0.002	0.059 ± 0.002
Femoral neck	5,138	0.047 ± 0.001	0.047 ± 0.001
Total hip	5,020	0.042 ± 0.001	0.042 ± 0.001
Male	Patients	TDD _{HND}	TDD _{FND}
L2-L4	350	0.066 ± 0.005	0.065 ± 0.005
L1-L4	348	0.061 ± 0.005	0.061 ± 0.005
Femoral neck	345	0.051 ± 0.003	0.050 ± 0.003
Total hip	339	0.042 ± 0.003	0.042 ± 0.003

female group was $0.062 \pm 0.002 \text{ g}/\text{cm}^2$ while it was $0.066 \pm 0.005 \text{ g}/\text{cm}^2$ for the male group. The TDD_{HND} value for the femoral neck from the female group was $0.047 \pm 0.001 \text{ g}/\text{cm}^2$ while it was $0.051 \pm 0.003 \text{ g}/\text{cm}^2$ for the male group. The results from the female groups were slightly smaller than those from the male groups for the same examination sites because women are more prone to bone mineral density loss than men of the same age. Women and men have different patterns of bone development. For women, there are two unique stages of rapid BMD changes (BMD loss): nursing and menopause. As an example, the results in Table 10.2 indicate that if a woman's BMD change for the lumbar spine L2-L4 over time is larger than $0.062 \pm 0.002 \text{ g}/\text{cm}^2$, she is likely to have experienced detectable BMD change. If a man's BMD change for the same region is over $0.066 \pm 0.005 \text{ g}/\text{cm}^2$, he is considered to have detectable BMD change.

An investigation with a larger sample of male patients would be useful to further validate the findings from the above discussion.

10.3 Age

This section aims to investigate whether the TDD values are affected by the patients' age. The age group of the 5,500 patients from the General Hospital involved in this study ranged from 20 to 103 years old.

Table 10.3: The TDD_{HND} results (g/cm²) of four age groups from the 5,500 patients, using a GE Lunar Prodigy densitometer.

Age(year)	TDD _{L2-L4}	TDD _{L1-L4}	TDD _{Femoral-neck}	TDD _{Total-hip}
<50	0.060 ± 0.004	0.057 ± 0.004	0.050 ± 0.004	0.044 ± 0.003
(50, 59)	0.063 ± 0.003	0.059 ± 0.003	0.047 ± 0.002	0.041 ± 0.002
(60, 69)	0.064 ± 0.003	0.061 ± 0.003	0.045 ± 0.002	0.042 ± 0.002
70<	0.061 ± 0.004	0.057 ± 0.004	0.050 ± 0.003	0.044 ± 0.003
All-age	0.062 ± 0.002	0.059 ± 0.002	0.047 ± 0.001	0.042 ± 0.001

Table 10.3 lists the TDD values from patients divided into four different age groups, age less than 50 yrs (denoted as <50), age between 50 and 59 yrs (denoted as (50, 59)), age between 60 and 69 yrs (denoted as (60, 69)), age greater than 70 yrs (denoted as 70 <) years old. Considering the 95 % CI, the TDD values from the different age groups were not significantly different. For the lumbar spine L2-L4, a minimum TDD value of 0.060 ± 0.004 g/cm² was obtained for the age group (<50) while a maximum TDD value of 0.064 ± 0.003 g/cm² was obtained for the age group (60, 69). Figure 10.3 shows how the TDD values changed for the four different age groups for each examination site. The fluctuations of each bar were less than 0.005 g/cm², see Table 10.3. Considering the 95 % CI of each TDD value for each examination site, the difference of the TDD values between each age group was not significant.

In order to avoid any potential averaging, the age group (50, 59) and (60, 69) were

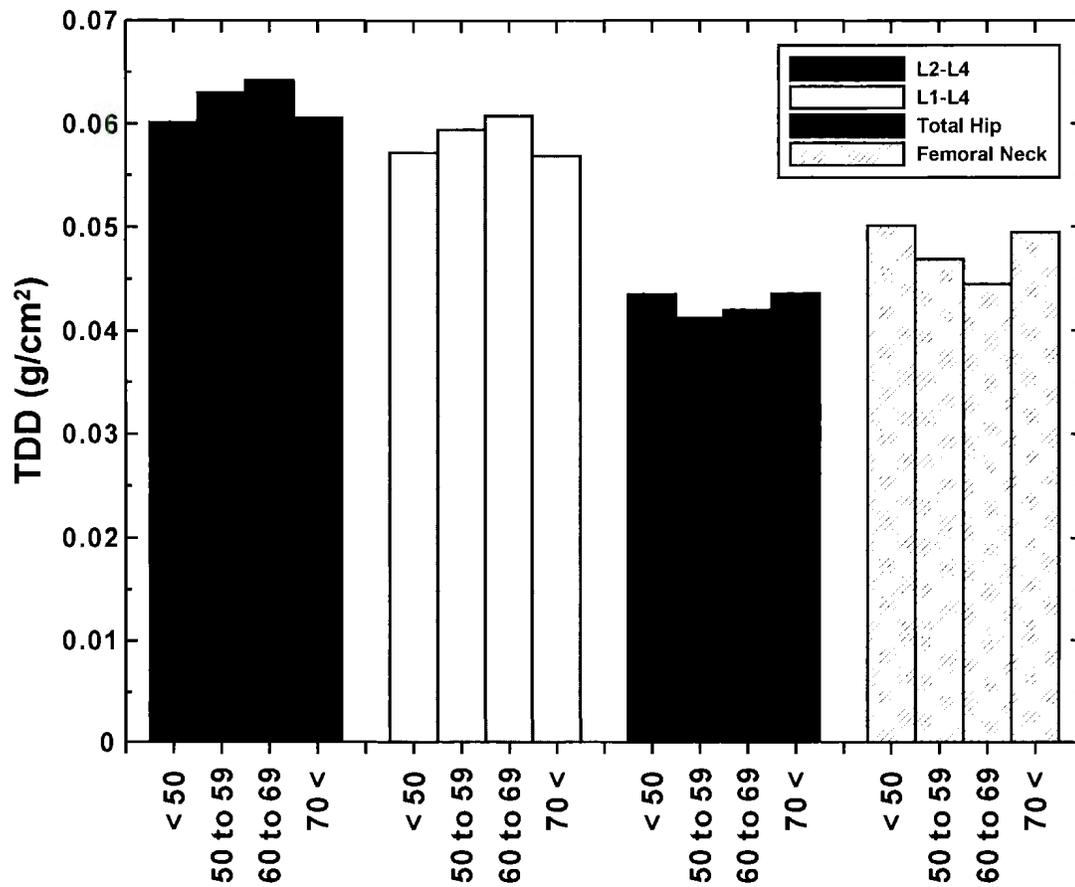


Figure 10.3: The TDD values (g/cm^2) for the four examination sites of the four age groups (<50, (50, 59), (60, 69) and 70<). The color scheme indicating each examination site is shown in the legend.

divided into 5 yrs groups (50, 54), (55, 59), (60, 64), (65, 69). The resulting TDD values with respect to these age groups were not significantly different from each other at a 95% CI. Therefore, the TDD values from the four age groups were not significantly different.

10.4 Examination Time

This section aims to investigate the effect of the examination time interval on the TDD values. The examination time interval, ΔTime , was calculated as the time difference (years) between the first BMD study and the third BMD study for each patient. The 5,500 patients were categorized into 10 different groups, with ΔTime ranging from 1 year to > 10 years.

Table 10.4: The TDD results (g/cm^2) using the HND for 10 examination time intervals from the 5,500 patient data, using a GE Lunar Prodigy densitometer.

Interval(year)	TDD _{L2-L4}	TDD _{L1-L4}	TDD _{Femoral-neck}	TDD _{Total-hip}
(1, 2)	0.051 \pm 0.018	0.052 \pm 0.018	0.044 \pm 0.015	0.037 \pm 0.012
(2, 3)	0.055 \pm 0.008	0.056 \pm 0.002	0.046 \pm 0.007	0.037 \pm 0.005
(3, 4)	0.058 \pm 0.006	0.057 \pm 0.006	0.041 \pm 0.005	0.038 \pm 0.004
(4, 5)	0.057 \pm 0.005	0.054 \pm 0.005	0.041 \pm 0.004	0.038 \pm 0.003
(5, 6)	0.061 \pm 0.005	0.059 \pm 0.005	0.046 \pm 0.004	0.043 \pm 0.003
(6, 7)	0.067 \pm 0.005	0.062 \pm 0.004	0.049 \pm 0.003	0.043 \pm 0.003
(7, 8)	0.066 \pm 0.004	0.062 \pm 0.004	0.050 \pm 0.003	0.042 \pm 0.003
(8, 9)	0.064 \pm 0.004	0.060 \pm 0.004	0.049 \pm 0.003	0.043 \pm 0.003
(9, 10)	0.062 \pm 0.005	0.058 \pm 0.004	0.047 \pm 0.003	0.044 \pm 0.003
10<	0.064 \pm 0.008	0.061 \pm 0.008	0.048 \pm 0.006	0.044 \pm 0.003
All-time	0.062 \pm 0.002	0.059 \pm 0.002	0.047 \pm 0.001	0.042 \pm 0.001

The TDD values calculated from the 10 time intervals are listed in Table 10.4 and shown in Figure 10.4. The TDD results from the 10 time intervals are more

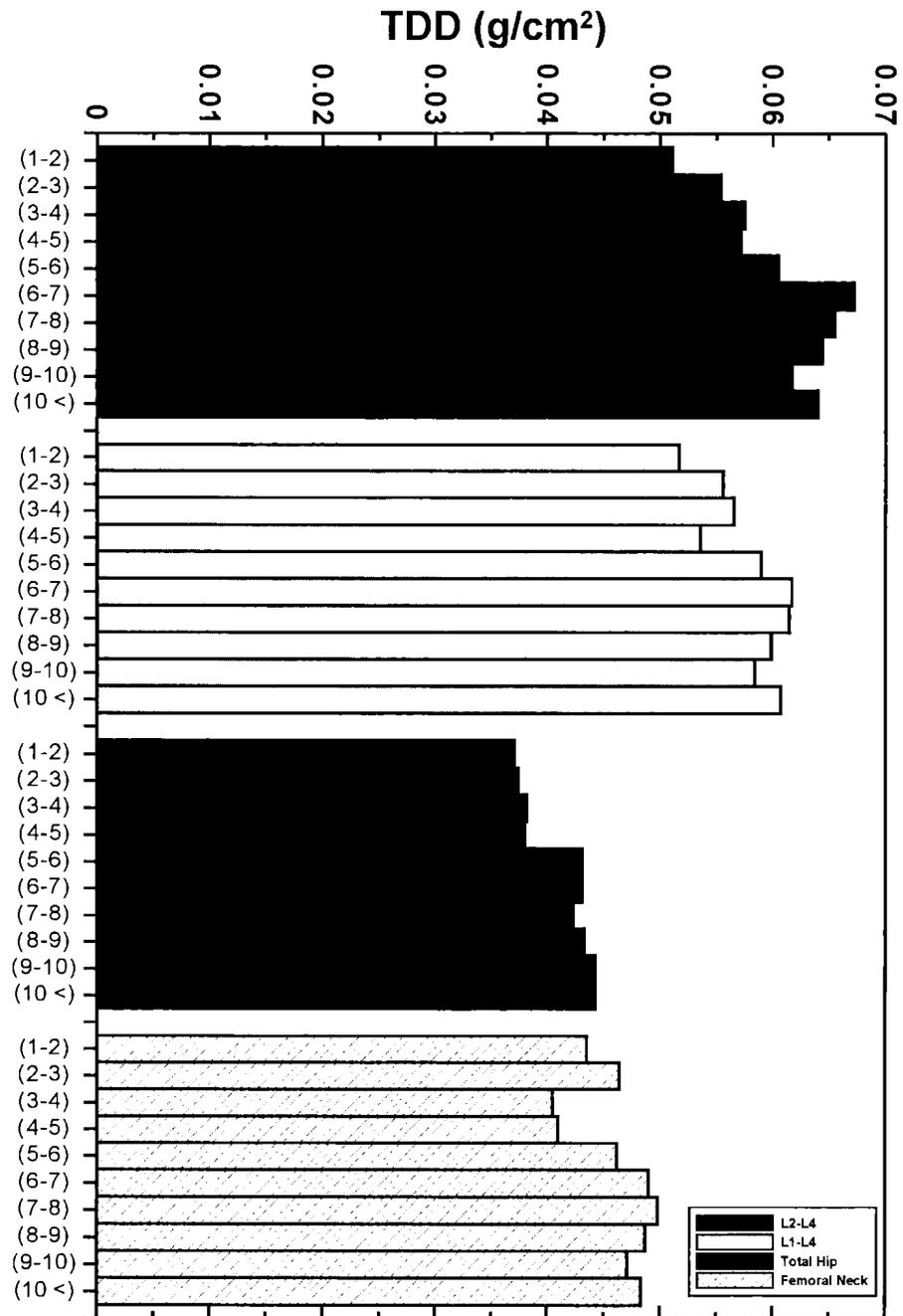


Figure 10.4: The TDD values from the 10 time intervals for the four examination sites: lumbar spine L2-L4, lumbar spine L1-L4, total hip and femoral neck. The color scheme indicating each examination site is shown in the legend.

complicated than those of the four age groups (see Figure 10.3). All the TDD values varied homoscedastically for each examination site. A change of TDD value over 11 years was between 0.009 g/cm^2 and 0.014 g/cm^2 for the spine, and 0.007 g/cm^2 and 0.009 g/cm^2 for the hip. The results for each examination site indicated an overall increase in the TDD values with increasing ΔTime . However, for the lumbar spine and femoral neck, the TDD values were observed to increase to a maximum value before decreasing slightly. For example, the TDD values for the lumbar spine L2-L4 increased from 0.051 g/cm^2 for $\Delta\text{Time} = (1, 2)$ to 0.066 g/cm^2 for $\Delta\text{Time} = (7, 8)$, then decreased to 0.062 g/cm^2 for $\Delta\text{Time} = (9, 10)$. A similar trend was found for the lumbar spine L1-L4, femoral neck and the total hip. The TDD values from the 10 time intervals were different, indicating that the TDD values were time-dependent. The longer the time interval between the BMD examinations, the larger the TDD values, indicating the larger the expected BMD changes. In order to determine the TDD values within a given examination time interval, a rate of change (e.g. TDD / year) would need to be known for assessing the BMD change rate, therefore providing a recommendation of the time interval between the BMD examinations for patients.

10.5 GE and Hologic

The 3,265 patient data collected using the Hologic QDR 4500 densitometer at the St. Clare's Mercy Hospital were used in this work for comparing with the TDD values using the GE Lunar Prodigy densitometer. The calculated TDD values from the GE and Hologic densitometers are shown in Table 10.5. The TDD values of the Hologic densitometer were approximately 19% less than those from the GE densitometer. This value agrees with a previous report that indicated the BMD values for the spine obtained using the Hologic densitometer were typically 11.7% lower than from the GE

Lunar Prodigy densitometer [159]. The difference in the TDD values obtained from the GE and Hologic densitometers is largely due to the different calibration standards and the mathematical algorithms used for choosing the regions of interest.

Table 10.5: The TDD values (g/cm^2) from the GE and Hologic densitometers in St. John's, NL.

Site	Patients	TDD _{GE}	Patients	TDD _{Hologic}
L2-L4	5,534	0.063 ± 0.002	3,265	0.049 ± 0.002
L1-L4	5,459	0.059 ± 0.002	3,265	0.046 ± 0.002
Femoral neck	5,483	0.047 ± 0.001	3,265	0.039 ± 0.001
Total hip	5,359	0.042 ± 0.001	3,265	0.036 ± 0.001

Figure 10.5 shows a comparison of the TDD values obtained from the GE (white bars) and Hologic (gray bars) densitometers in St. John's, NL. The TDD values for all four examination sites (the lumbar spine L2-L4, L1-L4, the femoral neck, and the total hip) from the GE densitometer were larger than those from the Hologic densitometer, see Table 10.5. These results are consistent with the statements made in Section 1.7 describing how the GE densitometer gives larger BMD values than the Hologic densitometer.

In summary, the TDD values obtained using the GE densitometer (from the General Hospital) were between 14% to 22% larger than those from the Hologic densitometer (from the St. Clare's Mercy Hospital), indicating the need for machine-specific TDD values to be calculated for each institution.

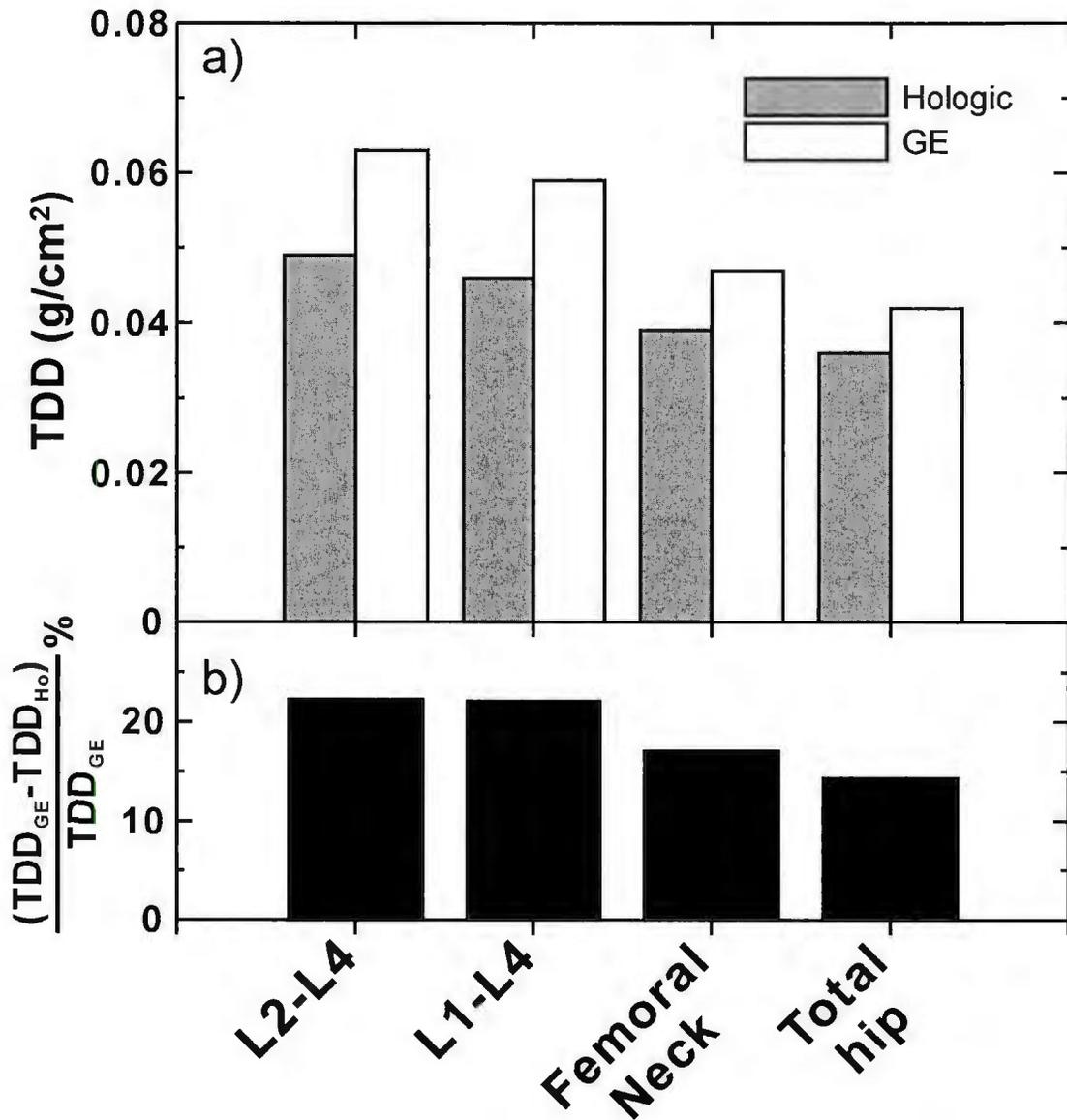


Figure 10.5: a) Comparison of the TDD values (g/cm^2) from the GE (white bar) and Hologic QDR 4500 densitometer (gray bar) for the lumbar spine L2-L4, L1-L4, femoral neck and total hip. b) Relative difference $(TDD_{GE} - TDD_{\text{Hologic}})/TDD_{GE}$ of the TDD values between the GE and Hologic densitometer.

10.6 CaMos data

The BMD studies from NL might not be consistent with those from the national dataset. In order to investigate the TDD values from other provinces, 648 BMD studies have been obtained from the Canadian multi-center osteoporosis study (CaMos). CaMos is an ongoing project for studying osteoporosis and bone health across Canada [160]. The CaMos program started in 1999 with 7 institutions in different provinces. Since the CaMos database contains patients' data from different institutions, it is interesting to compare the TDD values calculated from the CaMos BMD database to those calculated using the 8,800 patients from NL.

Table 10.6: TDD results (g/cm^2) from NL and the CaMos database using the GE Lunar densitometers, with number of patients in the parentheses.

Site	$\text{TDD}_{\text{GE-NL}}$	$\text{TDD}_{\text{GE-CaMos}}$
L1-L4	0.059 ± 0.002 (5,534)	0.047 ± 0.002 (389)
Femoral neck	0.047 ± 0.001 (5,483)	0.041 ± 0.002 (389)
Total hip	0.042 ± 0.001 (5,359)	0.041 ± 0.002 (389)

Table 10.7: TDD results (g/cm^2) from NL and the CaMos database using the Hologic QDR 4500 densitometers, with number of patients in the parentheses.

Site	$\text{TDD}_{\text{Hologic-NL}}$	$\text{TDD}_{\text{Hologic-CaMos}}$
L1-L4	0.049 ± 0.002 (3,265)	0.038 ± 0.002 (259)
Femoral neck	0.039 ± 0.002 (3,265)	0.032 ± 0.002 (259)
Total hip	0.036 ± 0.002 (3,265)	0.031 ± 0.002 (259)

The CaMos database includes BMD studies conducted using several different bone densitometers. The BMD data of the lumbar spine L2-L4 were not available for this study, therefore, the examination sites of interest for comparison in the following only include the lumbar spine L1-L4, femoral neck and the total hip. Since the

TDD values are machine specific, it is necessary to compare the TDD results based on different densitometers. Table 10.6 lists the TDD values calculated from BMD studies conducted using GE Lunar densitometers from the NL and CaMos database. The sample size of the NL database was approximately 1,400% times that of the CaMos database. The obtained TDD values from the GE Lunar densitometer in NL were significantly larger than those from the CaMos database, with an average difference of 20% for the lumbar spine L1-L4, and 13% for the femoral neck. The TDD values for the total hip were relatively close, with only a 2% difference between the NL and CaMos results.

Table 10.7 lists the TDD values calculated from BMD studies conducted using the Hologic QDR 4500 densitometers from the NL and CaMos database. The sample size of the NL database was approximately 1,200% that of the CaMos databases. The TDD values calculated from the BMD values from the Hologic QDR 4500 densitometer in NL were also significantly larger than those from the CaMos database, by approximately 22% for the lumbar spine L2-L4, 18% for the femoral neck, and 14% for the total hip. The TDD values of the lumbar spine L1-L4 from the CaMos database were consistently larger than those for the hip, which agrees with the preliminary findings from the NL database. Also, the TDD values from the GE Lunar densitometer were consistently larger than those from the Hologic densitometer using both the NL and CaMos database. However, due to the limited sample size in the CaMos data, it is difficult to compare the effects of gender, age or examination time interval within the CaMos database.

In summary, the TDD values from the NL and CaMos database were significant different for all but one (the TDD values for the total hip using the GE densitometer) of the examination sites.

10.7 Impact on Diagnosis

This study suggests a new criterion (TDD) for analyzing the BMD changes over time. In order to compare the two criteria (TDD and LSC) with respect to their impact on clinical diagnosis, a study was performed from a group of 9,379 patients, each of whom had two repeated BMD studies over a 11 year period.

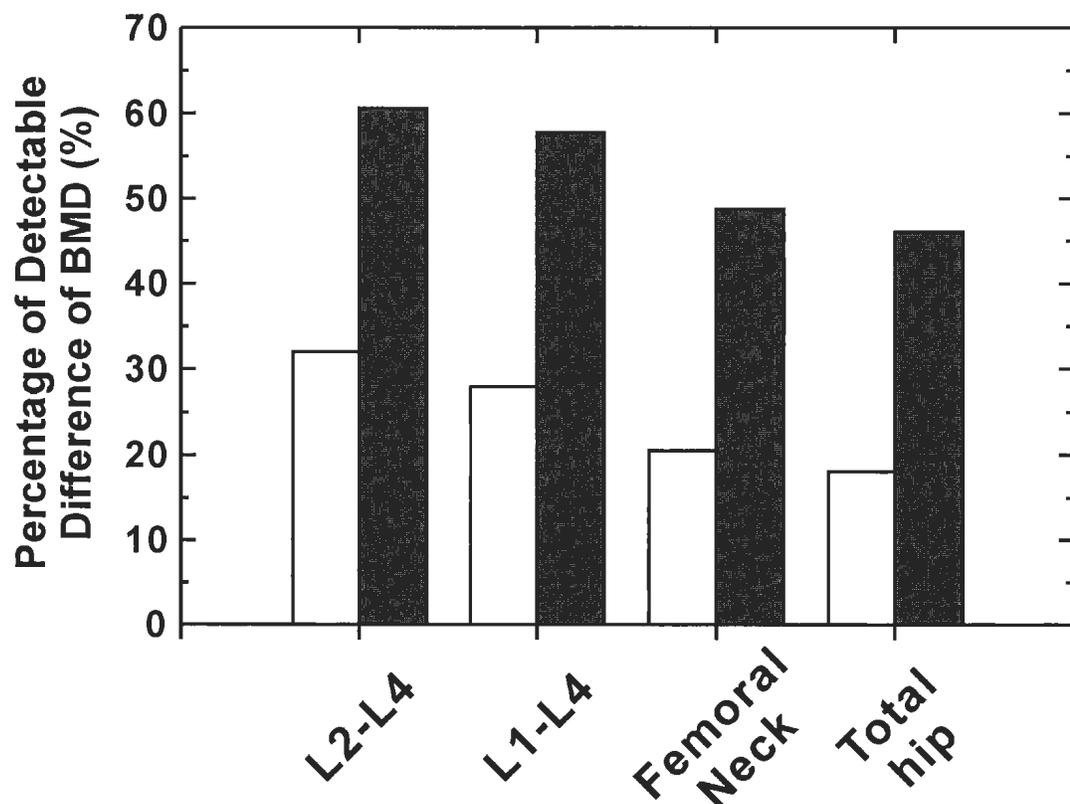


Figure 10.6: Percentages of the detectable BMD differences using the LSC (dark bars) and TDD (white bars) from the local GE Lunar densitometer.

Figure 10.6 lists the percentages of detectable BMD changes using the TDD and LSC based on these 9,379 patients. The assessment was obtained by comparing the two BMD studies for each patient using the TDD and LSC criterion. The reference

LSC values were calculated by Dr. Peter Hollett from the department of nuclear medicine at the General Hospital in 2006 using 15 patients (obtained through private communication), each of which had 3 repeated studies. For the lumbar spine L1-L4 and lumbar spine L2-L4, approximately 60% of the patients were diagnosed as having significant BMD changes using the LSC criterion, while 30% of the patients were assessed to have detectable BMD differences using the TDD criterion. Similarly, the percentage was approximately 50% for the femoral neck and total hip using the LSC, and 20% using the TDD criterion. The difference in the frequency of diagnosing detectable BMD changes using these two criteria is significant. As a result, using the LSC criterion might cause unnecessary medications or other health management to approximately 30% patients whose BMD changes were undetectable according to the TDD criterion.

10.8 Conclusion

This study has shown that the half-normal distribution accurately describes the distribution of SD values from the OLS linear regressed residuals from three sequential BMD studies. The TDD values were calculated from the scaled SD values from the HND. This TDD criterion can only be applied to patients with three sequential studies. For BMD studies with more than three sequential examinations, the folded-normal distribution should be used to find the TDD values. Females and males were found to have slightly different TDD values. The TDD values ranged from 0.042 to 0.063 g/cm² for the four examination sites. The 30% difference in detection frequency between using the LSC and using the TDD criteria indicates that there is a probability of 30% that the LSC misclassifies insignificant differences as significant. The TDD method has been reinforced as the method for detecting the total BMD differ-

ences over time by accounting for both machine and patient time-based variations. The results shown here have been presented at the annual meeting of the American Society of Bone and Mineral Research and the annual meeting of the Canadian Association of Nuclear Medicine. The abstract from this study was published in the Annual congress of the European associate of nuclear medicine. Two papers based on the above findings are currently in progress.

Acknowledgments

I wish to express my gratitude to Maureen Dunphy and Rick Scanlan for making patient data available, without which this work would not have been possible. I also wish to express my gratitude to Dr Christopher Kovacs for helping me access the CaMos data. This work was supported by grant HIC 10-57 from the Eastern Health Foundation.

Chapter 11

Conclusions

In this final chapter we provide a summary of the two research projects discussed in this thesis. Although both projects involved different topics, together they have shown how using the appropriate mathematical tools and methods of statistical analysis can be invaluable to assist in a scientific investigation of complex systems, solidifying the choice of title of this thesis as “Application of Physics and Mathematics in Clinical Nuclear Medicine”.

11.1 Plasma Clearance

Although others have proposed the use of the gamma variate function as a possible model for estimating GFR , it has never, to our knowledge, been used successfully for estimating plasma clearance [68]. Therefore our study of kidney function has made two significant contributions to the study of plasma clearance: 1) The Tikhonov regularization has been shown to provide a method of using the gamma variate model for estimating GFR by allowing a solution of the ill-posed inverse problem to be obtained, and 2) By performing a thorough comparison of the Tk-GV model with

the *SET1*, *SET2*, and OLS-GV models, it has been shown that the Tk-GV model best describes the plasma clearance in terms of: 1) the precision and accuracy, 2) the fewest non-physical *GFR* results, 3) the robustness to the choice of sample size and sampling times, and 4) agreement with the *GFR* results from the constant infusion of inulin without requiring further corrections. In the end, the application of the Tk-GV model in clinic has shown it to be the more feasible method for replacing the current two-sample *SET1* method. As a result, the Tk-GV model developed based on the findings from this work is already being used at the Roswell Hospital in Buffalo, New York.

In the case where four blood samples cannot be collected from the patient, the *SET1* model is often the method of choice for estimating *GFR* values. Unfortunately the *SET1* model has been shown to overestimate *GFR*. We have analyzed a criterion, $RSI_{f(Vol,W)}$, used to account for the overestimation of GFR_{SET1} by differentiating between patients with normal and abnormal *GFR* values [132]. Although this $RSI_{f(Vol,W)}$ correction criterion was initially developed for children, the analysis conducted here has shown, by cross comparison with other available clinical data, that this correction criterion is also applicable to adults. The importance of this finding is that it allows for a less invasive *GFR* test to be conducted in order to determine whether the patient is in need of further medical examination. As a result, this correction method can not only reduce the burden on hospital work loads but also decrease the number of false positives.

11.2 Bone Mineral Density

The conventional LSC criterion used for analyzing bone mineral density (BMD) changes has been shown to be insufficient for discriminating between significant and

insignificant changes in BMD [145–147]. This is because the LSC only considers changes in machine precision as a source of BMD variation. The Total Detectable Difference (TDD) has been proposed as a new criterion for determining significant changes in BMD by considering natural changes in a patients' short-term and long-term BMD, and the short-term and long-term errors introduced by densitometers used to measure BMD. By performing a frequency histogram analysis on the standard deviations of 8,800 patients, values of TDD have been obtained for the lumbar spine and the hip for both General Electric and Hologic densitometers. The calculated TDD values were found to be indifferent of age, but dependent on gender and the time interval between BMD examinations. A comparison of detectable BMD changes indicated that the LSC criterion has the potential of over-diagnosing patients as having detectable BMD change by approximately 30%. As a result, the replacement of the LSC by the TDD will significantly increase the accuracy of determining changes in BMD.

11.3 Future Directions

Similar to the kidney, the liver function can be estimated from a hepatic plasma clearance study using the radiopharmaceutical $^{99m}\text{Tc-HEPI-DA}$. Such a study is usually performed by taking several blood samples within 90 min of the injection, and then calculating the liver function based on the concentration-time curve produced from radiation measurements of the blood samples. The estimation of the liver function is commonly based on a *SETs* model. Since the Tk-GV model has been shown to improve the analysis of plasma clearance over the *SETs* models, it would be of interest to see if it could have the same impact on the study of liver function.

Besides the liver, kidneys are the second most important organ for eliminating

toxic chemotherapy agents from the body. Patients with kidney insufficiency have a higher risk of developing kidney dysfunction or acute kidney failure after chemotherapy due to the delayed secretion or metabolism of the chemo-agents. Therefore, it is important to evaluate the kidney function pre- and post- treatment. However, not every cancer center considers the kidney function when calculating the dose used in chemotherapy. Part of the reason is that the conventional methods of estimating kidney function are not sufficiently accurate, especially for patients with very low kidney function, patients that are morbidly obese, or patients who suffer from malnourishment.

Being able to perform animal studies has already been my passion in expending my research. It is interesting to apply the Tk-GV model and the constant infusion of inulin method to a group of animals with much better control of the injection and blood or urine collection. This will allow for a direct comparison of the results from the Tk-GV model and the gold standard for measuring kidney function. It will also be interesting to perform image studies on these animals now that the local General Hospital is planning to purchase MicroPET scan for performing small animal PET imaging. The application of the Tk-GV method can potentially improve the curve fitting of the renogram.

I have been awarded a Canadian Institutes of Health Research (CIHR) post-doctoral fellowship to start in April 2012 to investigate: 1) the appropriateness of BSA scaling in calculating the chemotherapy dosage, and 2) the application of the Tk-GV model to cancer patients for monitoring their kidney function before and after treatment. The initial phase of this research will compare the BSA and $f(Vol, W)$ scaling methods with respect to treatment outcomes and side effects. This work will focus on patient groups whose chemotherapy treatments can significantly influence

their kidney function. In the case where $f(Vol, W)$ is found to be insufficient in providing an accurate GFR assessment for chemotherapy patients, the Tk-GV model will be used before and after the treatment, in the hope of allowing the early detection of side effects on the kidneys due to the treatment. The Tk-GV model has a potential of providing two significant contributions in this area: 1) the early detection of patients experiencing insufficient kidney function so that the appropriate dosage can be determined for these patients, and 2) allow accurate monitoring of patients experiencing mild kidney dysfunction or acute kidney dysfunction and allow for accurate treatment to be assigned.

In a comparison of BMD studies conducted elsewhere in Canada it has been shown that the TDD values for Newfoundland and Labrador (NL) were, on average, 20% larger for the lumbar spine and 13% larger for the femoral neck than the national average. As a result, what would be considered as natural BMD changes in NL would be considered as significant BMD changes in rest of Canada. It would be of great interest to broaden our initial investigation of the TDD criterion by considering additional factors such as 1) geographic location, 2) nutrition level such as vitamin D intake, calcium level, diet, etc, and 3) sunshine exposure, exercise level, etc. It may also be of interest to investigate the effect of different seasons on BMD levels. For example, if patients experience greater BMD changes in summer than in winter, it may be necessary to have separate TDD criteria for each season.

Once a sufficient understanding of the TDD is obtained based on all significant parameters, it would be of great use to develop a program which would accept all the relevant information of the patient and provide, as output, the TDD value. This program would differ from FRAX developed by the WHO in that it would consider all examination sites, a larger age group than patients over the age of 50, and pro-

vide information on whether the observed changes in BMD are significant based on previous tests.

Bibliography

- [1] J A Sorenson and M E Phelps. *Physics in nuclear medicine*. Grune and Straton Inc, 1987.
- [2] K S Krane. *Introductory Nuclear Physics*. Wiley, 1987.
- [3] R E Henkin and et al. *Nuclear Medicine*. Grune and Straton Inc, First edition 1996.
- [4] W G Frankle, M Slifstein, P S Talbot, and M Laruelle. A review on brain-imaging applications of SPECT. *Int Rev Neurobiol*, 67:385–440, 2005.
- [5] M E Phelps. *PET: physics, instrumentation, and scanners*. Springer, 2006.
- [6] P C Lauterbur. Image formation by induced local interactions: examples of employing nuclear magnetic resonance. *Nature*, 242:190–191, 1973.
- [7] N J Hangiandreou. Physics tutorial for residents: topics in US. *Radiographics*, 23(4):101–109, 2003.
- [8] J M Mcalister. *Radionuclide techniques in medicine*. CUP Archive, 1979.
- [9] M Skugor. *Thyroid Disorders (A Cleveland Clinic Guide)*. Cleveland Clinic Press, 2006.

- [10] S Raymond, C Moses, and C Moyer. *Modern Physics*. Thomson, 2005.
- [11] P F Sharp, H G Gemmell, and A D Murray. *Practical Nuclear Medicine*. Springer, 2005.
- [12] N N Greenwood and A Earnshaw. *Chemistry of the elements*. Oxford: Butterworth-Heinemann, 1997.
- [13] N E Holden. *CRC Handbook of Chemistry and Physics (85th ed)*, volume section 11. CRC Press, 2004.
- [14] T C Yen, W P Chen, S L Chang, R S Liu, S H Yeh, and C Y Lin. Technetium-99m-DMSA renal SPECT in diagnosing and monitoring pediatric acute pyelonephritis. *J Nucl Med*, 37:1349–1353, 1996.
- [15] A Gonzále, L Jover, L I Maria, J Martin-Comin, and Puchal R. Evaluation of obstructed kidneys by discriminant analysis of 99mTc-MAG3 renograms. *Nuklearmedizin*, 33(6):244–247, 1994.
- [16] R F Fleay. ^{99m}Tc-labeled EDTA for renal scanning. *Australas Radiol*, 12:265–267, 1968.
- [17] R P Parker, P H S Smith, and D M Taylor. *Basic science of nuclear medicine*. Churchill Livingstong, 1978.
- [18] J R Hart. *Ethylenediaminetetraacetic acid and related chelating agents, Ullmann's encyclopedia of industrial chemistry*. Wiley-VCH, 2005.
- [19] G F Knoll. *Radiation Detection and Measurement*. John Wiley and Sons, 1999.

- [20] AAPM Task Group Members. *The Selection, Use, Calibration, and Quality Assurance of Radionuclide Calibrators Used in Nuclear Medicine*. American Association of Physicists in Medicine, 2012.
- [21] L S Graham. *Nuclear Medicine: Self-Study Program II: Instrumentation*. Society of Nuclear Medicine, 1996.
- [22] S Vallabhajosula. *Molecular Imaging: Radiopharmaceuticals for PET and SPECT*. Springer, 2009.
- [23] M W Benjamin. The aapm/rsna physics tutorial for residents, physics of SPECT. *Radiographics*, 16:173–183, 1996.
- [24] G L Zeng. Image reconstruction - a tutorial. *Comput Med Imaging Graph*, 25(2):97–103, 2001.
- [25] C Cobelli, D Foster, and G Toffolo. *Tracer Kinetics in biomedical research: from data to model*, volume 1. Springer, 2000.
- [26] D G Steele and C A Bramblett. *The anatomy and biology of the human skeleton*. Texas A and M University Press, 1988.
- [27] B L Specker and E Schoenau. Quantitative bone analysis in children: current methods and recommendations. *J Pediatr*, 146(6):726–731, 2005.
- [28] J A Sorenson and J R Cameron. A reliable in vivo measurement of bone mineral content. *J Bone Joint Surg*, 49(A):481–497, 1967.
- [29] G M Blake and I Fogalman. Technical principles of dual energy x-ray absorptiometry. *Seminars in Nuclear medicine*, 27(3):210–228, 1997.

- [30] R Morita, H Orimo, I Yamamoto, M Fukunaga, and et al. Some problems of dual-energy X-ray absorptiometry in the clinical use. *Osteoporosis Int*, 3(Suppl. 1):87–90, 1993.
- [31] B Fan, Y Lu, G T Fuerst, and J Shepherd. Does standardized bmd still remove differences between hologic and ge-lunar state-of-the-art dxa systems? *Osteoporos Int*, 21(7):1227–1236, 2010.
- [32] N A Pocock, P N Sambrook, T Nguyen, P Kelly, J Freund, and J A Eisman. Assessment of spinal and femoral bone density by dual x-ray absorptiometry: comparison of lunar and hologic instruments. *J Bone Miner Res*, 7(9):1081–1084, 1992.
- [33] IAEA Report Group. Dual energy x ray absorptiometry for bone mineral density and bone body composition assessment, IAEA human health series 15, 2011.
- [34] L A Stevens, J Coresh, T Greene, and A S Levey. Assessing kidney function—measured and estimated glomerular filtration rate. *N Engl J Med*, 354:2473–2483, 2006.
- [35] J E Hall. *Textbook of Medical Physiology (Guyton Physiology)*. Saunders, 2000.
- [36] J Brchner-Mortensen, H Rickers, and I Balslev. Renal function and body composition before and after intestinal bypass operation in obese patients. *Scand J Clin Lab Invest*, 40(8):695–702, 1980.
- [37] G B Haycock, G J Schwartz, and D H Wisotsky. Geometric method for measuring body surface area: A height-weight formula validated in infants, children and adults. *J Pediatr*, 93(1):62–66, 1978.

- [38] A M Peters, R Gordonand, and R Sixt. Normalization of glomerular filtration rate in children: body surface area, body weight or extracellular fluid volume? *J Nucl Med*, 35:438–444, 1994.
- [39] P Delanaye, R P Radermecker, M Rorive, G Depas, and J M Krzesinski. Indexing glomerular filtration rate for body surface area in obese patients is misleading: concept and example. *Nephrol Dial Transplant*, 20(10):2024–2028, 2005.
- [40] R D Lindeman, J Tobin, and NM Shock. Longitudinal studies on the rate of decline in renal function with age. *J Am Geriatr Soc*, 33(4):278–285, 1985.
- [41] A Levin, B Hemmelgarn, B Culleton, and et al. Guidelines for the management of chronic kidney disease. *Can Med Assoc J*, 179(11):1154–1162, 2008.
- [42] National kidney foundation. KDORI clinical practice guidelines for chronic kidney disease: evaluation, classification and stratification. *Am J Kidney Dis, Suppl1*, 39:1–266, 2002.
- [43] C T Jurkovitz, Y Qiu, C Wang, D T Gilbertson, and W W Brown. The Kidney Early Evaluation Program (KEEP): Program design and demographic characteristics of the population. *Am J Kidney Dis*, 51(4 Suppl 2):3–12, 2008.
- [44] S Stevens, L A anmd Li, C Wang, C Huang, B N Becker, and et al. Prevalence of CKD and comorbid illness in elderly patients in the united states: Results from the kidney early evaluation program (KEEP). *Am J Kidney Dis*, 55(3 Suppl 2):23–33, 2010.
- [45] B M Curtis, B J Barrett, O Djurdjev, J Singer, A Levin, and on behalf of the Can-Care Investigator Group. Evaluation and treatment of CKD patients

- before and at their first nephrologist encounter in Canada. *Am J Kidney Dis*, 50(5):733–742, 2007.
- [46] C Edelstein. *Biomarkers of kidney disease*. Elsevier Inc, 2010.
- [47] D S Siegel, L McBride, E Bilotti, N Lendvai, J Gonsky, T Berges, D Schillen, A McNeill, L Schmidt, and K H Van Hoven. Inaccuracies in 24-hour urine testing for monoclonal gammopathies: Serum free light chain analysis provides a more accurate measure of light chain burden than urine protein electrophoresis. *Laboratory Medicine*, 40(6):341–344, 2009.
- [48] M H Ferguson, O Olbrich, J S Robson, and C P Stewart. The use of inulin clearance as a measure of glomerular filtration. *Experimental Physiology*, 35(4):251–279, 1950.
- [49] R Chandra and J L Barron. Anaphylactic reaction to intravenous sinistrin (inutest). *Ann Clin Biochem*, 39 Pt1:76, 2002.
- [50] G J Weir. Radioisotopic determination of glomerular filtration rate and estimated renal plasma flow. *J Natl Med Assoc*, 64(4):338–341, 1972.
- [51] M Rehling, M Moller, B Thamdrup, J Lund, and J Trap-Jensen. Simultaneous measurement of renal clearance and plasma clearance of ^{99m}Tc-labeled diethylenetriaminepenta-acetate, ⁵¹Cr-labeled ethylenediaminetetraacetate and inulin in man. *Clin Sci*, 66:613–619, 1984.
- [52] G A Rose. Measurement of glomerular filtration rate by, inulin clearance without urine collection. *British Medical Journal*, 2(5649):91–93, 1969.

- [53] B R Cole, J Giangiacomo, J R Ingelfinger, and A M Robson. Measurement of renal function without urine collection of renal function without urine collection. *N Engl J Med*, 287(4):1109–1114, 1972.
- [54] C V Guldener, ROB Gans, and P M T Wee. Constant infusion clearance is an inappropriate method for accurate assessment of an impaired glomerular filtration rate. *Nephrol Dial Transplant*, 10(1):47–51, 1995.
- [55] T A Chowdhury, P H Dyer, W A Bartlett, E S Legge, S M Durbin, A H Barnett, and S C Bain. Glomerular filtration rate determination in diabetic patients using iohexol clearance comparison of single and multiple plasma sampling methods. *Clin Chim Acta*, 277(2):153–158, 1999.
- [56] M Walser. Assessing renal function from creatinine measurements in adults with chronic renal failure. *Am J Kidney Dis*, 32(1):23–31, 1998.
- [57] A S Levey, J P Bosch, J B Lewis, T Greene, N Rogers, and D Roth. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. modification of diet in renal disease study group. *Ann Intern Med*, 130(6):461–70, 1999.
- [58] D W Cockcroft and M H Gault. Prediction of creatinine clearance from serum creatinine. *Nephron*, 16(1):31–41, 1976.
- [59] T H Mathew, D W Johnson, G R Jones, and et al. Chronic kidney disease and automatic reporting of estimated glomerular filtration rate: revised recommendations. *Med J Aust*, 187(8):459–63, 2007.

- [60] C Chantler and T M Barratt. Estimation of glomerular filtration rate from plasma clearance of 51-chromium edetic acid. *Arch Dis Child*, 47(254):613–617, 1972.
- [61] A E Moore, S J Park-Holohan, G M Blake, and I Fogelman. Conventional measurements of GFR using 51Cr-EDTA overestimate true renal clearance by 10 percent. *Eur J Nucl Med Mol Imaging*, 30(1):4–8, 2003.
- [62] G J Schwartz and et al. New equations to estimate GFR in children with CKD. *J AM Soc Nephrol*, 20(3):629–637, 2009.
- [63] G Pathuri, K Sahoo, V Awasthi, and H Gali. Renogram comparison of p-[(18)f]fluorohippurate with o-[(125)i]iodohippurate and [(99m)tc]mag3 in normal rats. *Nucl Med Commun*, 32(10):908–912, 2011.
- [64] R Weissleder and U Mahmood. Molecular imaging. *Radiology*, 219 (2):316–333, 2001.
- [65] B L kathel. Radioisotope renography as a renal function test in the newborn. *Arch Dis Child*, 48(247):314–320, 1971.
- [66] A M Davison. *Oxford textbook of clinical nephrology*, volume 1. Oxford University Press, 1988.
- [67] B G Katzung. *Basic and Clinical Pharmacology*. McGraw-Hill Companies, 2007.
- [68] M E Wise. Negative power functions of time in pharmacokinetics and their implications. *J Pharmacokinet Biopharm*, 13(3):309–346, 1985.

- [69] A Piepsz, P Colarinha, I Gordon, K Hahn, P Olivier, R Sixt, and V J Van. Guidelines for glomerular filtration rate determination in children. *Eur J Nucl Med*, 28(3):31–36, 2001.
- [70] C D Russell, A T Taylor, and E V Dubovsky. A bayesian regression model for plasma clearance. *J Nucl Med*, 43(6):762–766, 2002.
- [71] C D Russell, P G Bischoff, F N Kontzen, K L Rowell, M V Yester, L K Lloyd, W N Tauxe, and E V Dubovsky. Measurement of glomerular filtration rate, single injection plasma clearance method without urine collection. *J Nucl Med*, 26(11):1243–1247, 1985.
- [72] C D Russell, P G Bischoff, F N Kontzen, K L Rowell, M V Yester, L K Lloyd, W N Tauxe, and E V Dubovsky. Estimation of extracellularfluid volume from plasma clearance on technetium-99m dtpa by a single-injection,two-sample method. *J Nucl Med*, 29(2):255–258, 1988.
- [73] C D Russell. A bayesian 3-compartment model for 99mTc-MAG3 clearance. *J Nucl Med*, 44(8):1357–1361, 2003.
- [74] A S Perkinson, C J Evans, M T Burniston, and S W Smye. The effect of improved modeling of plasma clearance in paediatric patients with expanded body spaces on estimation of the glomerular filtration rate. *Physiol Meas*, 31(2):182–193, 2010.
- [75] M Boroujerdi. *Pharmacokinetics: Principles and Applications*. McGraw-Hill Medical Publishing Division, 2001.

- [76] G Koren, C Barker, G Goresky, D Bohn, G Kent, J Klein, S M MacLeod, and W D Biggar. The influence of hypothermia on the disposition of fentanyl-human and animal studies. *Eur J Clin Pharmacol*, 32(4):373–376, 1987.
- [77] R S Soman, H Zahir, and F Akhlaghi. Development and validation of an hplc-uv method for determination of iohexol in human plasma. *J of Chromatography*, 816(1):339–343, 2005.
- [78] C Kollmannsberger, M Schittenhelm, F Honecker, J Tillner, D Weber, K Oechsle, L Kanz, and C Bokemeyer. A phase i study of the humanized monoclonal anti-epidermal growth factor receptor (egfr) antibody emd 72000 (matuzumab) in combination with paclitaxel in patients with egfr-positive advanced non-small-cell lung cancer (nslc). *Annals of Oncology*, 17(6):1007–1013, 2006.
- [79] L A Sapirstein, D G Vidt, M J Mandel, and G Hanusek. Volumes of distribution and clearances of intravenously injected creatinine in the dog. *Am J Physiol*, 181(2):330–336, 1955.
- [80] K Florijn, J Barendregt, E Lentjes, D W Van, W Prodjosudjadi, J Van Saase, L Van Es, and P Chang. Glomerular filtration rate measurement by single-shot injection of inulin. *Kidney Int*, 46(1):252–252, 1994.
- [81] K ITOH, S Tsushima, E Tsukamoto, and N Tamaki. Accuracy of plasma sample methods for determination of glomerular filtration rate with ^{99m}Tc -dtpa. *Ann Nucl Med*, 16(1):39–44, 2002.
- [82] C D Russell, P G Bischoff, F Kontzen, K L Rowell, M V Yester, and et al. Measurement of glomerular filtration rate using ^{99m}Tc -DTPA and the gamma camera: a comparison of methods. *Eur J Nucl Med*, 10(11-12):519–521, 1985.

- [83] C D Russell, E V Dubovsky, and A T Taylor. Prediction of urinary excretion of technetium-99m-mag3. *J Nucl Med*, 39(7):1257–1259, 1998.
- [84] C D Sadeleer, K V Larer, B Geroges, A Piepsz, and H R Ham. Influence of time interval and number of blood samples on the error in renal clearance determination using a mono-exponential model: a monte carlo simulation. *Nucl Med Commun*, 21(8):747–745, 2000.
- [85] C D Russell. Optimum sample times for single-injection, multisample renal clearance methods. *Nucl Med Commun*, 34(10):1761–1765, 1993.
- [86] A W Murray, M A Gannon, M C Barnfield, and M L Waller. Optimized robust plasma sampling for glomerular filtration rate studies. *Nucl Med Commun*, 33(9):995–1001, 2012.
- [87] A Silvers, R S Swenson, J W Farquhar, and Reaven. Derivation of a three compartment model describing disappearance of plasma insulin - 131-I in man. *J Clinical Investigation*, 48(8):1461–1469, 1969.
- [88] K Zierler. A critique of compartmental analysis. *Annu Rev Biophys Bioeng*, 10:531–562, 1981.
- [89] M D Hughson, W E Samuel Hoy, and J F Bertram. Glomerular volume and clinicopathologic features related to disease severity in renal biopsies of African Americans and whites in the southeastern united states. *Arch Pathol Lab Med*, 131(11):1665–1672, 2007.
- [90] Y Maaravi, M Bursztyn, R Hammerman-rozenberg, and J Stessman. Glomerular filtration rate estimation and mortality in an elderly population. *Q J Med*, 100(7):441–449, 2007.

- [91] J F M Wetzels, L A L M Kiemeney, D W Swinkels, N L Willems, and H D Heijer. Age- and gender-specific reference values of estimated GFR in Caucasians: The nijmegen biomedical study. *Kidney Int*, 72:632–637, 2007.
- [92] I Sharkey, A V Boddy, H Wallace, J Mycroft, R Hollis, and S Picton. Body surface area estimation in children using weight alone: application in pediatric oncology. *Br J Cancer*, 85(1):23–28, 2001.
- [93] M Sawyer and M J Ratain. Body surface area as a determinant of pharmacokinetics and drug dosing. *Invest New Drugs*, 19(2):171–177, 2001.
- [94] C R Rao. *Linear statistical inference and its applications*. New York: John Wiley and Sons, 1973.
- [95] C D Russell, A T Taylor, and E V Dubovsky. A bayesian regression model for plasma clearance. *J Nucl Med*, 43(6):762–766, 2002.
- [96] P R Schloerb. Total body water distribution of creatinine and urea in nephrectomized dogs. *Am J Physiol*, 199:661–665, 1960.
- [97] R V Hogg and A T Craig. *Introduction to mathematical statistics*. New York, Macmillan, 1978.
- [98] O Heymansa, J Fissetea, P Vico, S Blacher, D Masset, and F Brouers. Is fractal geometry useful in medicine and biomedical sciences? *Medical Hypothesis*, 54(3):360–366, 2000.
- [99] G A Losa, D Merlini, T F Nonnenmacher, and E R Weibel. *Fractals in Biology and Medicine*. A Birkhäuser Basel product, 2000.

- [100] G T Tucker, P R Jackson, G C A Storey, and D W Holt. Amiodarone disposition: Polyexponential, power and gamma functions. *Eur J Clin Pharmacol*, 26(5):655–656, 1984.
- [101] H K Thompson, F Frank Starmer, R E Whalen, and H D McIntosh. Indicator transit time considered as a gamma variate. *Circ Res*, 14(6):502–515, 1963.
- [102] R O Binswanger, H Rösler, U Noelpp, L Matter, and M Haertel. The bedside determination of extravascular lung water. *Eur J Nucl Med*, 3:109–114, 1978.
- [103] G Galli, C L Maini, M Salvatori, and F Andreasi. A practical approach to the hepatobiliary kinetics of 99m Tc-HIDA. clinical validation of the method and a preliminary report on its use for parametric imaging. *Eur J Nucl Med*, 8(7):292–298, 1983.
- [104] M J Avram, T K Henthorn, D A Spyker, T C Krejcie, P M Lloyd, and et al. Recirculatory pharmacokinetic model of the uptake, distribution, and bioavailability of prochlorperazine administered as a thermally generated aerosol in a single breath to dogs. *Drug Metab Dispos*, 35(2):262–267, 2007.
- [105] S R Golish, J D Hove, H R Schelbert, and S S Gambhir. A fast nonlinear method for parametric imaging of myocardial perfusion by dynamic ^{13}N -Ammonia PET. *J Nucl Med*, 42(6):924–931, 2001.
- [106] H J Michaely, H Kramer, N Oesingmann, K P Lodemann, M F Reiser, and S O Schoenberg. Semiquantitative assessment of first-pass renal perfusion at 1.5 T: Comparison of 2D saturation recovery sequences with and without parallel imaging. *AJR*, 188(4):919–926, 2007.

- [107] H Schwilden, J Honerkamp, and C Elster. Pharmacokinetic model identification and parameter estimation as an ill-posed problem. *Eur J Clin Pharmacol*, 45(6):545–550, 1993.
- [108] A N Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics*, 4:1035–1038, 1963.
- [109] A N Tikhonov and V Y Arsenin. *Solution of Ill-posed Problems*. Washington: Winston and Sons, 1977.
- [110] D L Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J Assoc Comput t*, 9(1):84–97, 1962.
- [111] A E Hoerl. Application of ridge analysis to regression problems, chemical engineering progress. *Chem Eng Prog*, 58:54–59, 1964.
- [112] P R Bevington and D K Robinson. *Data reduction and error analysis for the physical sciences*. McGraw-Hill Book Co., 2002.
- [113] L Khinkis, W Krzyzanski, W Jusko, and W Greco. D-optimal designs for parameter estimation for indirect pharmacodynamic response models. *J Pharmacokinetic Pharmacodyn*, 36(6):523–539, 2009.
- [114] N Hilliard. *Correspondence Continuing Education Courses for Nuclear Pharmacists and Nuclear Medicine Professionals*. University of New Mexico, 1991.
- [115] J A M Santos, A G Dias, and A L Bastos. Characterization of geometric uncertainties of a dose calibrator during measurement of 90y activity. *J Nucl Med Technol*, 39(2):125–130, 2011.

- [116] S Barai, S Gambhir, N Prasad, R K Sharma, M Ora, A Kumar, A Gupta, D S Parasar, and B Suneetha. Levels of gfr and protein-induced hyperfiltration in kidney donors: A single-center experience in india. *Am J Kidney Dis*, 51(3):407–414, 2008.
- [117] D Lee and L C Chang. Development of the pipetting error sensor. *Sensors and Actuators B: Chemical*, 119(1):150–158, 2006.
- [118] C A Wesolowski, R C Puetter, L Ling, and P S Babyn. Tikhonov adaptively regularized gamma variate fitting to assess plasma clearance of inert renal markers. *J Pharmacokinet Pharmacodyn*, 37(5):435–474, 2010.
- [119] L Ingber. Simulated annealing: Practice versus theory. *Mathl Comput Modelling*, 18:29, 1993.
- [120] W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. *Numerical recipes: the art of scientific computing*. Cambridge University Press, 2007.
- [121] J A Nelder and R Mead. A simplex method for function minimization. *Comput J*, 7(4):308–313, 1965.
- [122] S H Brooks. A discussion of random methods for seeking maximum. *Oper Res*, 6(2):244–251, 1958.
- [123] L L Cam. Maximum likelihood - an introduction. *ISI reviews*, 58:153–171, 1990.
- [124] R S Murray. *Probability and statistics*. McGraw-Hill Book Co., 1980.
- [125] W Hamilton, J Moore, J Kinsman, and R Spurling. Further analysis of the injection method, and of changes in hemodynamics under physiological and pathological conditions. *Am J Physiol*, 99:534–551, 1932.

- [126] S S Huang, A P Sharma, A Yasin, R M Lindsay, W F Clark, and G Filler. Hyperfiltration affects accuracy of creatinine egfr measurement. *CJASN*, 6(2):274–280, 2011.
- [127] I Helal, G M Fick-Brosnahan, B Reed-Gitomer, and R W Schrier. Glomerular hyperfiltration: definitions, mechanisms and clinical implications. *Nat Rev Nephrol*, 8(5):293–300, 2012.
- [128] A Chagnac, T Weinstein, A Korzets, and et al. Glomerular hemodynamics in severe obesity. *Am J Physiol Renal Physiol*, 278(5):817–822, 2000.
- [129] B Rippe and B Haraldsson. Transport of macromolecules across microvascular walls: the two-pore theory. *Physiol Rev*, 74(1):163–219, 1994.
- [130] J S Fleming, M A Zivanovic, G M Blake, M Burniston, and P S Cosgriff. Guidelines for the measurement of glomerular filtration rate using plasma sampling. *Nucl Med Commun*, 25(8):759–769, 2004.
- [131] C A Wesolowski, L Ling, E Xirouchakis, M T Burniston, R C Puetter, P S Babyn, I G Giamalis, and A K Burroughs. Validation of tikhonov adaptively regularized gamma variate fitting with 24-hour plasma clearance in cirrhotic patients with ascites. *Eur J Nucl Med Mol Imaging*, 38(12):2247–2256, 2011.
- [132] C A Wesolowski, P S Babynb, and R C Puetter. An improved method for determining renal sufficiency using volume of distribution and weight from bolus ^{99m}Tc -DTPA, two blood sample, paediatric data. *Nucl Med Commun*, 27(12):963–970, 2006.

- [133] C A Wesolowski, P S Babynb, and R C Puetter. A critical comparison of renal sufficiency index from 99mtc-dtpa, slope-intercept, pediatric data. *Nucl Med Commun*, 48(Suppl2):294–306, 2007.
- [134] R Morita. The history and present status of bone densitometry. *Rinsho Hoshasen*, 35(1):1–6, 1990.
- [135] Report of a WHO Scientific Group. Prevention and management of osteoporosis: report of a WHO scientific group. *WHO technical report*, 2000.
- [136] C C Glüer, G Blake, Y Lu, B A Blunt, M Jergas, and H K Genat. Accurate assessment of precision errors: How to measure the reproducibility of bone mineral techniques. *Osteoporosis Int*, 5(4):262–270, 1995.
- [137] S L Bonnick, C C Johnston, M Kleerekoper, R Lindsay, P Miller, L Sherwood, and E Siris. Importance of precision error in bone density measurement. *J Clin Densitom*, 4(2):105–110, 2001.
- [138] S L Bonnick and L A Lewis. The precision of PA spine, dual femur and single femur bone density studies on the GE Lunar Prodigy, a DXA fan-array device. *J Clin Densitom*, 5:Suppl 48, 2002.
- [139] S L Hui, C W Slemenda, and C C Johnston. Age and bone mass as predictors of fracture in a prospective study. *J Clin Invest*, 81(6):1804–1809, 1988.
- [140] H Kröger, M Lunt, J Reeve, J Dequeker, J E Adams, J C Birkenhager, and et al. Bone density reduction in various measurement sites in men and women with osteoporotic fractures of spine and hip. The European quantization of osteoporosis study. *Calcif Tissue Int*, 64(3):191–199, 1999.

- [141] J A Kanis, D Black, C Cooper, P Dargent, B Dawson-Hughes, C D Laet, P Delmas, J Eisman, O Johnell, and et al. A new approach to the development of assessment guidelines for osteoporosis. *Osteoporosis Int*, 13(7):527–536, 2002.
- [142] L Lenchik, G M Kiebzak, and B A Blunt. What is the role of serial bone mineral density measurements in patient management? *J Clin Densitom*, 5(Suppl 3):29–38, 2002.
- [143] A F Olson. Osteoporosis detection: is BMD testing the future? *J Prim Health Care*, 32(6):20–27, 2007.
- [144] The writing group for the ISCD Position Development conference. Technical standardization for dual-energy X-ray absorptiometry. *J Clin Densitom*, 7(1):27–36, 2004.
- [145] J C Wong and Griffiths M R. Precision of bone densitometry measurements: When is change true change and does it vary across bone density value? *Australas Radiol*, 47(3):236–239, 2003.
- [146] W D Leslie. The importance of spectrum bias on bone density monitoring in clinical practice. *Bone*, 39(2):361–368, 2006.
- [147] A Moayyeri, M Sadatsafavi, and W D Leslie. Sample size requirement for bone density precision assessment and effect on patient categorization: A Monte Carlo simulation study. *Bone*, 41(4):679–684, 2007.
- [148] M Sadatsafavi, A Moayyeri, L Wang, and W D Leslie. Heteroscedastic regression analysis of factors affecting BMD monitoring. *J Bone Miner Res*, 23(11):1842–1849, 2008.

- [149] T V Nguyen, N A Pocock, and J A Eisman. Interpretation of bone mineral density measurement and its change. *J Clinic Densi*, 3(2):107–119, 2000.
- [150] Y Lu, T Fuerst, S Hui, and H K Genant. Standardization of bone mineral density at femoral neck, trochanter and wards triangle. *Osteoporos Int*, 12(6):438–444, 2001.
- [151] A Papaioannou, S Morin, M Angela, Cheung, and et al. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary, 2010.
- [152] J A Kanis, O Johnell, H Johansson, A Oden, P Delmas, J Eisman, and et al. Prior clinical vertebral fractures are a particularly strong predictor of hip fracture: a meta-analysis. *Osteoporosis Int*, 17(Suppl 3):365, 2006.
- [153] C D Laet, A Odén, Johansson h, O Johnell, and et al. The impact of the use of multiple risk factors on case finding strategies: a mathematical framework. *Osteoporosis Int*, 16(3):313–318, 2005.
- [154] M Sadatsafavi, A Moayyeri, L Wang, and W D Leslie. Optimal decision criterion for detecting change in bone mineral density during serial monitoring: a bayesian approach. *Osteoporosis Int*, 19(11):1589–1596, 2008.
- [155] K A McGarry and D P Kiel. Postmenopausal osteoporosis. strategies for preventing bone loss, avoiding fracture. *Postgrad Med*, 108(3):79–82, 2000.
- [156] W N Venables and B D Ripley. *Modern applied statistics*. Springer, 2002.
- [157] J Gurland and R C Tripathi. A simple approximation for unbiased estimation of the standard deviation. *J Am Stat*, 25(4):30–32, 1971.

- [158] A Milton, S A Irene, and etc. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1965.
- [159] B Fan, Y Lu, H Genant, T Fuerst, and J Sheperd. Does standardized BMD still remove differences between Hologic and GE-Lunar state-of-the-art DXA systems? *Osteoporosis Int*, 21:1227–36, 2009.
- [160] N Kreiger, A Tenenhouse, L Joseph, T Mackenzie, S Poliquin, J P Brown, J C Prior, and R S Rittmaster. The Canadian Multicentre Osteoporosis Study (CaMos): Background, rationale, methods. *Can J Aging*, 18(3):376–387, 1999.

Appendix A

A.1 Clinical Protocol for Measuring *GFR*

A detailed protocol is given for preparing the administered dose of radiopharmaceutical, collecting and processing the blood samples from the patients.

Collect the Samples

1. Record patient's height and mass.
2. Prepare the patient's dose using the ^{99m}Tc from $^{99}\text{Mo}/^{99m}\text{Tc}$ generator and measure the activity with the dose calibrator. The required dose for a patient undergoing a kidney function test (*GFR*) and renal scan on the same day is twice that of a patient undergoing only a kidney function test.
3. Prepare the standard dose equal to approximately 1/10 of the patient's dose.
4. Administer the dose to the patient using an angiocatheter and flush the bolus with at least 10 ml of a saline solution for children and 20 ml for adults. Record the starting time.
5. Measure the radioactivity of all materials/objects (syringe used for the injection, gloves and alcohol swabs, etc) that have been in contact with the injected dose and record it as the patient's residual.

6. Four blood samples (6 ml each) to be drawn (by needle or angiocatheter) at 10, 30, 120, and 240 min. Blood is drawn using a 21G needle and one dark green topped tube (6 mL, heparin coated) from the opposite arm from where the injection was given. The angiocatheter, if used, should be filled with an anticoagulant or preservative free saline and in each case the first 5 ml of blood (3 ml for children) is to be discarded before the true samples are collected to avoid dilution of the sample concentration. Record each sampling time to the hour, minute, second.

Process the Samples

1. The standard solution is poured in a 500 ml volumetric flask and diluted by filling the remainder of the flask with water (tap water is sufficient) up to the 500 ml mark. Cover the flask and gently shake it until fully mixed. Measure and record the radioactivity left in the syringe used for dispensing the standard dose and record the time.
2. All blood samples must be centrifuged for 10 min at 3,750 rpm until the plasma is separated from the blood. If hemolysis occurs, centrifuge the plasma for an additional 5 minutes.
3. Label 17 counting 5 ml tubes: 2 tubes for background (BKG), 3 tubes for the standard dose (STD) and 3 tubes for each of the 4 plasma samples.
4. Deposit 200 μ l of the STD and plasma samples in the corresponding tubes using an air displacement pipette. The pipette tip should be changed every time a new sample is drawn.
5. Add 1 ml (1000 μ l) of water to each tube using a larger pipette. (There is no need to change the tip between these samples).

6. The radioactivity of each sample is then measured in the single-well counter for 1 minute. If a pediatric study is performed with a low injection dose and standard dose, the samples should each be measured for 2 minutes to obtain better counting statistics. The measurement time for each sample should be recorded in order to perform a decay correction.
7. The obtained values (radioactivity of 17 tubes, patient's information and exam date) are entered into the computer program for calculations.

A.2 Mathematical Formulae

This section aims to supplement Section 2.6.4 by listing additional mathematical formulae which are too long to include within the text of this thesis. Equation 2.33 lists a series of standard deviation terms which are part of the Tikhonov regularization. These standard deviations are listed below:

$$\begin{aligned}
 S_{\alpha}^2 &= \frac{1}{N-1} \sum_i^n (\alpha_i - \bar{\alpha})^2 \\
 S_{\beta}^2 &= \frac{1}{N-1} \sum_i^n (\beta_i - \bar{\beta})^2 \\
 S_K^2 &= \frac{1}{N-1} \sum_i^n (K_i - \bar{K})^2 \\
 S_{\alpha\beta}^2 &= \frac{1}{N-1} \sum_i^n (\alpha_i - \bar{\alpha})(\beta_i - \bar{\beta}) \\
 S_{\alpha K}^2 &= \frac{1}{N-1} \sum_i^n (\alpha_i - \bar{\alpha})(K_i - \bar{K}) \\
 S_{\beta K}^2 &= \frac{1}{N-1} \sum_i^n (\beta_i - \bar{\beta})(K_i - \bar{K}), \tag{A.1}
 \end{aligned}$$

where $\bar{\alpha}$, $\bar{\beta}$ and \bar{K} are the average values of α , β and K . The parameters α , β , K , ERG and ERV are functions of the shrinkage factor λ used in the Tk-GV model and are obtained through a series of mathematical operations and as a result cannot be shown here in a simple analytical form. Instead I will show the equations for α , β , K , ERG and ERV below after the patient information has been applied. For more specific information on the exact equations for α , β , K , ERG and ERV or for the Mathematica code used to perform the calculations in this work, the reader is recommended to contact Dr. Wesolowski directly.

$$\alpha = \frac{(0.062 + \lambda)(1.5 + \lambda)}{0.14 + 2.0\lambda + \lambda^2} \quad (\text{A.2})$$

$$\beta = \frac{0.00019 + 0.0056\lambda}{0.14 + 2.0\lambda + \lambda^2} \quad (\text{A.3})$$

$$K = \exp\left(-\frac{5.6(0.068 + \lambda)(1.5 + \lambda)}{0.14 + 2.0\lambda + \lambda^2}\right) \quad (\text{A.4})$$

$$\begin{aligned}
ERG &= \lambda^3((0.200(\lambda + 0.0045)(\lambda + 2.54)(\lambda + 2.72)(\lambda^2 - 0.080\lambda + 0.054) \\
&\quad (\lambda^2 + 0.54\lambda + 0.089))/(0.95\lambda + 0.0323)^2/(\lambda^2 + 2.00\lambda + 0.14)^4 \\
&\quad + 0.16(\lambda + 0.27)(\lambda + 2.69)(\lambda^2 + 0.15\lambda + 0.096) \\
&\quad \log(0.0056\lambda + 0.000191)/(\lambda^2 + 2.00\lambda + 0.14) \\
&\quad + 0.035(\lambda + 0.0045)(\lambda^2 + 0.21\lambda + 0.130)(\lambda^2 + 0.34\lambda + 0.092) \\
&\quad \log(0.0056\lambda + 0.00019)/(\lambda^2 + 2.00\lambda + 0.143)^2 \\
&\quad + (-0.16\lambda^2(\lambda + 0.272)(\lambda + 2.69) - 0.070(\lambda + 0.0045) \\
&\quad (\lambda^2 + 0.21\lambda + 0.13)(\lambda^2 + 0.34\lambda + 0.092)) \\
&\quad \log(0.0056\lambda + 0.00019)/(\lambda^2 + 2.00\lambda + 0.14) \\
&\quad \psi((\lambda + 0.062)(\lambda + 1.50)/(\lambda^2 + 2.00\lambda + 0.14)) \\
&\quad + 0.035(\lambda + 0.0045)(\lambda^2 + 0.21\lambda + 0.13)(\lambda^2 + 0.34\lambda + 0.092) \\
&\quad \psi((\lambda + 0.062)(\lambda + 1.50)/(\lambda^2 + 2.00\lambda + 0.143))^2 \tag{A.5}
\end{aligned}$$

$$\begin{aligned}
ERV = & 1/((0.034 + \lambda)^2)(0.062 + \lambda)^2(1.5 + \lambda)^2((0.14 + 2.0\lambda + \lambda^2)^4)) \\
& \times (1.1\lambda^3(1.78(0.007 + \lambda)(0.063 + \lambda)(0.36 + \lambda)(1.49 + \lambda) \\
& (1.64 + \lambda)(0.003 - 0.067\lambda + \lambda^2)(0.16 + 0.22\lambda + \lambda^2) \\
& + (-8.8 \times 10^{-8} + 0.092\lambda^4 + 0.43\lambda^5 + 1.15\lambda^6 + 2.22\lambda^7 + 1.82\lambda^8 + 0.50\lambda^9) \\
& \log(0.0056\lambda + 0.000191)/(\lambda^2 + 2.00\lambda + 0.143) \\
& + 0.035\lambda(0.0082 + \lambda)(0.219 + \lambda)(1.49 + \lambda)(1.51 + \lambda)(0.074 + 0.128\lambda + \lambda^2) \\
& (0.149 + 0.326\lambda + \lambda^2) \log(0.0056\lambda + 0.00019)/(\lambda^2 + 2.00\lambda + 0.14)^2 \\
& + (8.8 \times 10^{-8} - 0.092\lambda^4 - 0.43\lambda^5 - 1.15\lambda^6 - 2.22\lambda^7 - 1.82\lambda^8 - 0.50\lambda^9 - 0.070\lambda \\
& (0.0082 + \lambda)(0.219 + \lambda)(1.49 + \lambda)(1.51 + \lambda)(0.074 + 0.128\lambda + \lambda^2) \\
& (0.15 + 0.33\lambda + \lambda^2) \log(0.0056\lambda + 0.00019)/(\lambda^2 + 2.00\lambda + 0.14)) \\
& \psi((\lambda + 0.062)(\lambda + 1.50)/(\lambda^2 + 2.00\lambda + 0.143)) \\
& + 0.035\lambda(0.0082 + \lambda)(0.219 + \lambda)(1.49 + \lambda)(1.51 + \lambda)(0.074 + 0.128\lambda + \lambda^2) \\
& (0.149 + 0.326\lambda + \lambda^2)\psi((\lambda + 0.062)(\lambda + 1.50)/(\lambda^2 + 2.00\lambda + 0.14))^2) \quad (A.6)
\end{aligned}$$

where the symbol ψ stands for the digamma function which is given as

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)} \quad (A.7)$$

where $\Gamma(x)$ is a gamma function of the variable x and $\Gamma'(x)$ is the first derivative of the gamma function as [158],

$$\Gamma'(x) = \frac{d\Gamma(x)}{dx}. \quad (A.8)$$

The minimization of the *ERG* value is used for finding the shrinkage factor λ .

