ZÁPADOČESKÁ
UNIVERZITA
V PLZNI

FAKULTA APLIKOVANÝCH VĚD

DISERTAČNÍ PRÁCE

PLZEŇ, 2003

Ing. Petr Gebouský

ZÁPADOČESKÁ
UNIVERZITA
V PLZNI

Fakulta aplikovaných věd

DISERTAČNÍ PRÁCE

k získání akademického titulu doktor
v oboru

Kybernetika

Ing. Petr Gebouský

Bayesian Quantitative Lymphoscintigraphy of Upper Limbs

Školitel: Doc. Ing. Jiří Cendelín, CSc.
Datum státní doktorské zkoušky: 30.11.2001
Datum odevzdání práce: 19.12.2003

V Plzni, 2003

Abstract

The dissertation focuses on the design and verification of a new methodology of *quantitative lymphoscintigraphy of upper limbs*. Lymphoscintigraphy seems to be the only clinically and economically adequate method for diagnostics of latent lymphedema. Even partial statistics from Czech Republic indicate medical, social and economic need for such a methodology. Its potential for examination of upper limbs is, however, inhibited by the lack of a reliable, routinely applicable, quantitative evaluation. It is caused by availability of at most three of scan extremities. This lack of data follows from patient state, economic and time factors. Use of Bayesian methodology, that consistently supplements the sparse data by a rich prior knowledge, appears to be the only viable remedy.

The proposed quantification is based on simplified Bayesian modelling of the radiotracer accumulation dynamics on the pre-specified regions of interest (ROI) on the limb. The patient-specific cascade model is constructed via estimation of both limb local and common model parameters. The parametric and predictive inference respecting both types of parameters together with computationally feasible solutions are elaborated.

Bayesian decision theory in conjunction with the model is then employed to make inference useful to physicians. The whole time activity curve (TAC) — scintigraphic response on ROI in time — is estimated and estimate uncertainty characterized. TAC forms the quantification basis and serves for construction of quantitative indicators that are expected to characterize lymphedema stages. This conjecture will be inspected after collecting sufficient amount of patient data. At present, the dissertation offers an efficient quantitative comparison of patient limbs so that suspicion on bad state of one of them can be reliably tested. The comparison uses 3 available data pairs and thus solves the problem, which would be hopeless without a Bayesian treatment.

The limited number of routine inspections makes influence of measurement times significant. Their innovative optimization is solved in the thesis, too. It selects the best combination of sampling times present in experimental set of data.

The proposed methodology is verified from several perspectives on the real scintigraphic data and its favourably outcomes compared on experimental set of data with conclusions of physicians.

Абстракт

Тема диссертации предполагает создание и последующую тестирование эффективности функционирования новой методологии *количественной лимфосцинтиграфии верхних конечностей*.

По всей видимости, лимфосцинтиграфия является единственным клинически и экономически подходящим методом для диагностики скрытой (латентной) формы лимфедемы. Даже неполные статистические данные по Чешской республике указывают на острую необходимость подобной методологии, мотивированную медицинскими, социальными и экономическими причинами.

Однако, потенциальные возможности данного метода исследования верхних конечностей существенно ограничены отсутствием надёжной и удобной в применении методики качественной оценки. Это обусловлено отсутствием необходимого количества данных, что связано с состоянием конкретного пациента, а также временными и экономическими факторами. Использование методологии Байеса, позволяющей дополнять разбросанные данные используя априорную информацию, представляется единственно возможным решением проблемы. Разработанная количественная оценка основывается на упрощённом байесовском моделировании динамики аккумулирования радиоэлемента в предварительно заданной области (ROI) верхней конечности. Индивидуальная модель пациента конструируется при помощи совместного оценивания локальных и общих параметров модели. В диссертации были разработаны численно релизуемые решения для параметрической и предсказывающей моделей, учитывающих оба типа параметров.

Использование теория принятия решений Байеса позволило сделать полученные выводы пригодными для врачей. Проводится оценка всей кривой активности (TAC) — временной сцинтиграфический отклик в ROI — и характеризуется оценка неопределённости. TAC создаёт основу для количественной оценки и служит для создания количественных индикаторов, предположительно характеризующих различные стадии лимфедемы. Данное предположение может быть проверено только при наличии достаточного количества данных. В диссертации предлагается эффективное количественное сравнение верхих конечностей пациента, позволяющее достоверную проверку подозрения о нездоровом состоянии одной из конечностей. Именно благодаря использованию байесовского подхода, предлагаемый способ решения даёт достоверную оценку всего на основе сравнения трёх имеющихся в наличии пар данных.

Ограничение количества возможных измерений у пациента усиляет значимость времени проведения измерений. Усовершенствованная оптимизация, предлагаемая в данной диссертации, производит выбор наилучшей комбинации моментов измерений из имеющегося набора экспериментальных данных.

Различные аспекты предложенной методологии были проверены на реальных сцинтиграфических данных. Также было проведено сравнение результатов, полученных на экспериментальных данных, с заключениями врачей.

Preface

This thesis was prepared during my study at Faculty of Applied Sciences, University of West Bohemia in Pilsen, Czech Republic. It is based on the research work carried out in the Adaptive Systems Department, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

First of all, I would like to thank to my supervisor at the University of West Bohemia, Doc. Ing. Jiří Cendelín, CSc. At the same time, I would like to express my deepest gratitude to Ing. Miroslav Kárný, DrSc., head of the Department of Adaptive Systems for his patient guiding through my scientific research, profound discussions and creative ideas.

My thanks also belong all the members of the Clinics of Nuclear Medicine, Faculty Hospital Motol in Prague for providing the data and helpful comments. I am very grateful to MUDr. Hana Křížová for her patient and invaluable support in medical aspects of the solved problem.

I acknowledge as well my colleagues for their support and encouragement during my stay in the Department.

This work was partially supported by the grants:

- AV ČR S1075102,
- IGA MZ ČR NC7601–3/2003.

Pilsen, December 2003 Petr Gebouský

Contents

List of Tables

List of Figures

Symbols and Notations

Mathematical Symbols

$x,\ X,\ \boldsymbol{X}$	scalar quantity, vector, matrix
x^*	set of possible values of the quantity x
$\mathring{x}$	cardinality of the set x^*
$\boldsymbol{X}^T,\ X^T$	matrix $\boldsymbol{X}$/vector X transposition
$\equiv$	equivalence
$\propto$	proportional sign
$\Gamma(\cdot)$	Euler gamma function
$\cdot!$	factorial
$\psi(\cdot)$	digamma function ($\psi(\cdot) = \frac{d}{d\cdot} \log \Gamma(\cdot)$)
$\|\| \cdot \|\|$	Euclidean norm
$\boldsymbol{I}_n$	identity matrix of dimension n
$\mathrm{tr}(\cdot)$	matrix trace

$f(\cdot),\ f(\cdot\|\cdot)$	*p(d)f*/ conditional *p(d)f*
$\mathrm{Pn}(x\|\lambda)$	Poisson *pf*
$\mathrm{Un}(x\|\alpha,\beta)$	continuous uniform *pdf*
$\mathrm{Ga}(x\|\alpha,\beta)$	gamma *pdf*
$\mathrm{Pg}(x\|\alpha,\beta,n)$	Poisson-gamma *pf*
$\mathrm{N}(X\|M,\boldsymbol{\Lambda})$	multivariate normal *pdf* (univariate version $\mathrm{N}(x\|\mu,\lambda)$)
$\mathrm{St}(X\|M,\boldsymbol{\Lambda},\alpha)$	multivariate Student *pdf* (univariate version $\mathrm{St}(x\|\mu,\lambda,\alpha)$)
$E[x],\ E[X]$	expected (mean) value of x and X
$Cov[X],\ V[x]$	covariance (matrix) of X, variance of scalar x
$\sigma(x)$	standard deviation of x
$\mathcal{D}\left(f\|\|\hat{f}\right)$	Kullback-Leibler distance of *p(d)f*s f and $\hat{f}$

$\mathcal{Z}(\cdot)$	loss function
Θ	parameters of parametrised model
$t,\ \tau$	discrete time $\tau(t) = t\Delta$ (Δ sampling period)
$\mathcal{I}$	time interval
x_t	quantity x at the discrete time t, x used here for inner quantity
u_t	(scalar) system input at time t
y_t	(scalar) system output at time t
δ_t	unit impulse ($\delta_0 = 1$, $\delta_t = 0$ for $t > 0$)
$\mathcal{T} = (t_0, \ldots, t_n, \ldots)$	vector of observation sampling times
$Y = (y_{t_0}, \ldots, y_{t_n}, \ldots)$	system output at sampling times $\mathcal{T}$

Physical Units

min	minute (60 s)
Bq	becquerel [s^{-1}], unit of activity, number of nuclear changes per 1 s

Abbreviations

BCRL	breast-cancer-related-lymphedema
CDP	complex decongestive physiotherapy
p(d)f	probability (density) function
ROC	receiver operating characteristics
ROI	region of interest
TAC	time activity curve

Chapter 1

Introduction

Lymphoscintigraphy is a sensitive diagnostic technique in nuclear medicine. One of its principal applications is the investigation of upper limb lymphedema. The potential of quantitative scintigraphy as the promising tool for the recognition of early stages of lymphedema is however inhibited by the lack of reliable routinely applicable evaluation. This work concerns with the decisive part of the project that aims at creating and verifying such methodology of *quantitative lymphoscintigraphy of upper limbs*. It is based on the current, routinely performed, morphologically oriented measurements. The quantification should help to physicians in the assessment of stages of lymphedema and in evaluation of treatment success.

The methodology of quantitative lymphoscintigraphy is a practically important example of problems having little data with a vague expert knowledge and necessity to make inference only on their base. The proposed quantification involves Bayesian modelling of the radiotracer accumulation dynamic in the limb combining information from data, simplified theoretical modelling and expert knowledge. Bayesian decision theory in conjunction with the model is then employed to make inferences useful to physicians.

The objective of this Chapter is to give a brief overview of the addressed problem. Still missing satisfactory solution has motivated the current work and resulted into the formulation objectives of this thesis.

1.1 Motivation

Lymphedema, swelling caused by defects of the lymphatic system arising mostly after the complex therapy of breast carcinoma, is still an underestimated indicator. Its late stages can be hardly cured. They decrease substantially working abilities of the patient and often lead to a complete disability. Patients are often improperly cured due to the wrong diagnosis that explains problems as venal insufficiency, various rheumatic disease, allergic swelling, entesopathy etc. Often, they are not cured at all. This calls for an efficient and reliable diagnostic method that allows us to recognize safely *early* stages of lymphedema. Inspection by scintigraphy is potentially the method search for.

From the surgery viewpoint, diagnostics of *early* stages of the secondary lymphedema is of a great importance as the increasing number of solid tumors calls for increasing number of surgery lymphadectomy and radiotherapy of regional lymph nodes. Consequently, the number of secondary lymphedema increases. For instance, 42% women after a complex therapy of breast carcinoma exhibit lymphedema of collateral upper limb [85], often with a many-years delay after therapy. More than 10% of lymphedema manifest themselves even three years after surgery

and/or radiotherapy. Treatment of malignant breast tumor concerns yearly 4000 women considering Czech Republic only. This indicates socio-economic significance of a decrease of the number of unrecorded and untreated lymphedema of collateral upper limb. Recognition of the disease during the clinically silent phase, so called latent lymphedema is, however, decisive for the success of treatment. Early recognition and therapy of latent lymphedema lead with a high probability to full recovery of the lymphatic system. A substantial decrease of expenses spent for a difficult treatment of advanced stages of lymphedema with their complications can be expected. Lymphoscintigraphy, radionuclide based lymphography, seems to be the only clinically and economically adequate method for the diagnostics of latent stages of lymphedema [10, 32, 87].

Lymphoscintigraphy is non-invasive, sensitive method suitable for judging of the state of lymphatic system of upper limbs, in which radionuclides are used to image regional lymph drainage system. It provides structural data at the injection depot, along the extremity, and over the axillary lymph nodes. Furthermore, it provides functional information on diffusion of radiolabelled protein or colloid particles within the lymphatic system. Its potential for examination of upper limbs is, however, inhibited by the lack of a reliable *quantitative* evaluation when upper limbs are examined. Reliable *routinely applicable* quantitative scintigraphic evaluation is still missing. The main reason is number of measurements *limited* both by the time-capacity of the gamma camera and by the ability of a patient to undergo a series of measurement in long time intervals. Consequently, the number of images taken can hardly be larger than 2–3. This is very limiting condition for making inference even some traditional inferences of important physiological indicators are completely unreliable.

It is known that modelling and careful data processing is the only way how to counteract lack of measured data on complex phenomena. Considering the limited amount of data corrupted by the uncertainties, we rely on Bayesian data processing [5] as the only well established tool that is able to combine consistently information from data, theoretical modelling and expert knowledge. The power of the Bayesian paradigm is evident in inference problems based on a few measurements only. If at least some expert knowledge exists then more information can be included into the whole model so the uncertainty of the results can be reduced. It is just the case of quantitative scintigraphy of upper limbs. The amount of available scintigraphic data is strongly limited. At the same time a team of experienced physicians and physicists, whose expert knowledge can be incorporated into the model, is accessible. Consequently, modelling of the accumulation of the radiotracer within the limbs drifted by lymphatic flow at scintigraphic examination combined together with Bayesian paradigm and decision theory have been proposed for the scintigraphy quantification.

The scintigraphy quantification is the particular task to be solved. Though, similar problems with the limited amount of data are often met not only in the medicine branch thus the development of methodology may show the direction how to solve them.

1.2 State of the Art

Despite the fact that many diagnostic techniques have been proposed, anamnesis and careful clinical inspection form a natural diagnostic basis for the evaluation of lymphedema. The current *mainstream* of investigation is lymphoscintigraphy.

Judging of the lymphedema and therapy effect by measurements of the limb perimeter has a restricted value only. It provides information on a volume decrease of the liquid in the given region in the case of clinically advanced lymphedema. The inspection do not, however, provide information on presence or absence of functional disorder of the lymphatic system. Imaging techniques such as ultrasonography, computed tomography and magnetic resonance provide gross anatomic

information rather than a functional assessment of lymph drainage [13, 17]. Lymphoscintigraphy even suits much more for diagnostics of lymphedema than also conventional lymphangiography that provides information on morphology of lymphatic vessels. A lower cost, non-invasiveness, lower toxicity and lower radiation load are other advantages, comparing to roentgen lymphangiography. Thus, both the proper diagnosis of latent lymphedema and evaluation of the therapy effect require reliable evaluation of diagnostic and post-therapeutic lymphoscintigraphy.

The *qualitative* lymphoscintigraphy also called static qualitative scintigraphic inspection analyses the scintigrams for visualization of lymph vessels and nodes, dilatation of vessels, existence of collateral vessels and dermal backflow. It means it allows the characterization lymphatic morphology. The reported sensitivity of the morphologic evaluation is 71% [92]. However it turned out there, that this evaluation is sufficient for late disease stages but diagnostics for the critical early latent stage is *poorly* supported with the results depending enormously on the skill of the inspecting expert.

Consequently, another lymphoscintigraphic examination is necessary mainly for improving of diagnostics of lymphedema in its latent stage. Therefore there have been efforts to *quantify* scintigraphy. That is to find such quantitative parameters from scans for assessment of the lymphatic system and the stage of the disease. Both names, quantitative and semi-quantitative scintigraphy, are used for it in various papers. The reason why some authors use the term semi-quantitative is to emphasize that the scintigraphy is unable to quantify lymphatic flow rate. By contrast, the quantitative scintigraphy is used herein to distinguish it from visual qualitative evaluation. A few works are devoted to lymphoscintigraphy quantification up to now. There are mainly publications from eighties [11, 12, 29, 50, 92]. Employed methods differ by radiopharmaceutical used, locality administration and acquisition times. Majority of them found the quantitative evaluation more *accurate* in detection of incipient lymphedema. For instance, the sensitivity of 91% and specifity of 100% were found for 190 extremities at 115 patients [29] when the results are evaluated simultaneously morphologically and semi-quantitatively. Though, majority of them concerns examination of lower limbs. None of them gave the absolutely satisfying solution for the upper limbs. This is due to the fact that dynamics of drainage rate of upper limbs differs substantially from that of lower limbs.

An extensive group of studies correlate quantitative parameters acquired directly from a few scans with clinical findings of the disease stages, specifically the arrival times and late storage activities in pre-specified regions of the forearm, upper arm and axilla [36, 62, 71, 89, 92]. The experiments with the arrival times showed the promising results on the lower limbs but the similar results for upper limbs were not useful [92]. The arrival times alone were discarded by some authors for the evaluation supposing lymphedema results from decreased flow. The arrival time is a measure of velocity [62] but a reduced velocity may still be associated with normal (or even increased) rate of flow. The reason is that the flow rate is also function of the cross-sectional area and it is known that lymphatic vessels dilate after proximal obstructions. Storage activity in regional lymph nodes has been shown to be a useful diagnostic tool in cases of unexplained edema [92] and has been reported to be the most sensitive marker of lymphatic failure in breast cancer-related lymphedema [46, 59, 88]. In study [30], 100% of patients with arm swelling showed reduced storage activity in the ipsilateral axilla compared with that in the contralateral axilla at 2 years after surgery. However, patients in the study had between 5 and 35 axillary lymph nodes removed during surgery and had subsequent irradiation of the area, such that no meaningful comparison could be made.

The more elaborate examination of upper limbs is addressed e.g. in [38, 89, 91]. Authors of [91] use so called transport-index based differentiating nine stages of four parameters reflecting lymphedema. The method is time consuming and potentially strongly biased by subjective errors of the evaluating medical doctor. Authors of [89] performed dynamic scintigraphy in axilla region

for 45 minutes; they acquire late images with subsequent quantifications in ROIs of forearm, arm and liver 5 hours after radiopharmaceutical administration. These procedures, as well as methods using acquisition time of 120 minutes [92] or performing evaluation on the basis of 4 whole-body scintigrams in time intervals from 2 minutes to 2 hours after the administration [38], are not routinely applicable. It motivates search for new ways of diagnostic [70].

Depot clearance studies (e.g. [12, 58, 72]) are the only inspections employing the *dynamic* models. They examine the rate at which radiolabelled colloid is cleared from the initial site of injection. As the tracer clears from depot in exponential fashion it is simple to compute its removal rate. Least-square method is usually used for it. This parameter has been considered to be the best parameter for assessing lymphatic flow but many studies found it to be of a little use in the evaluation of lymphedema [30, 46, 59]. The explanation for this may be recent evidence suggesting the edema is the result of increased lymph formation rather than decreased flow [84].

In summary, reliable quantitative evaluation applicable under routine conditions of a few measurements is still *missing*. Chance for the resolution may be the studying the dynamic properties of lymphatic system at scintigraphic inspection at the whole limb. If the dynamics at the injection site seemed as the important feature then similar characteristic from the rest of limb can bring additional useful information. The dynamics in the injection site is a function only of the flow rate and local diffusion, but the rest of the limb reflects also other characteristics of lymphatic system including lymph formation. Dynamics of local tracer accumulation can serve for more detailed evaluation. It can reveal e.g. the location of tapering or blocking of lymphatic flow. Hence, it is necessary to exploit data from the whole limb. In this case, the available data are accumulated activities in predefined regions of interest (ROI). To evaluate the dynamics of lymphatic system at scintigraphy inspection requires to employ dynamic models describing the accumulation of the radiotracer.

Surprisingly, to the best author's knowledge, there has not been yet any published attempt employing modelling of the accumulation dynamics of the radiotracer in the limb during the lymphoscintigraphy except the mentioned modelling of depot clearance. The present work is an attempt to fill the gap believing it can improve the diagnosis of lymphedema.

In general, there are two possible basic methods how to derive such model. The physically based approach in which model is obtained by application of physical principles and laws. Alternatively, the data based approach where the model is obtained directly from experimental data using system identification [53, 81] and time-series analysis [7, 8].

Detailed modelling using the physically based approach is almost impossible here. The modelling of the liquid systems is rather complicated in general. Moreover, the behaviour on various limbs is quite variable and the physics of the lymphatic process in the limb is not understood enough to specify mathematical model of radiotracer diffusion uniquely. Even if such model would be constructed, that should be a version of partial differential equations, then it would be very complicated with many unknown parameters. Therefore it could not be practically used thereafter. The only possibility, then, is to adopt data based approach with a sufficiently flexible but relatively simple black-box models with a few of unknown parameters [71].

Already at the beginning it is necessary to exclude simple deterministic models as the suitable way of description. The reason is that data available for the processing can be "corrupted" with the uncertainty due to variation of measurement conditions but surely they are corrupted with uncertainty caused by the physical nature of the measurement process — counting particles of ionizing radiation in some fixed time interval. Using then the traditional deterministic formulae, the data uncertainties are cumulated in the uncertainty of the results. The uncertainty of the results obtained in this way is unknown. A systematic way how to incorporate also the uncertainties into the model, it means to make the whole description consistent, is to employ stochastic modelling [40, 41, 48].

Various well-established techniques how to obtain suitable stochastic model from the data has been already proposed and the research carries on. Publications concerning system identification and time series analysis is devoted directly to this problem [2, 8, 20, 31, 53, 68, 79, 81]. Though, majority of techniques assume sufficient amount of data and offer a limited ability how to incorporate further information (besides the model structure) into the final model. Consequently, limited number of data *prevents* their usage here. The Bayesian processing seems to be the only viable option.

The main reason to use Bayesian approach is that it allows to collect and handle all available information while the other techniques give user less freedom to express the prior knowledge and combine it consistently with data. This is important in the considered inference problem where the lack of data has to be compensated by careful employment of additional information. Since the unknown parameters of black-box models are actually not estimated in Bayesian approach but the whole a posteriori probability distribution for them calculated, problems like biasedness become mostly irrelevant. It is different from majority estimation methods where mostly point estimate is obtained and its uncertainty is expressed asymptotically. Such point estimate is not in many cases representative enough. Besides, in Bayesian framework there are no technical computational troubles even if the number of data is smaller than the number of unknown parameters. The next important reason for employing Bayesian approach is that theory of Bayesian decision making under uncertainty [5, 6, 47, 73] offers a consistent framework for solving various decision problems over the used model. The model alone is important outcome but its main aim is to serve as means for real decision making purposes.

This thesis comprises the proposal of quantitative scintigraphy methodology but it is also practical example demonstrating the strength of Bayesian approach to decision making. It can be considered as the case study that attempts to solve various partial tasks and problems that have appeared important or useful within scintigraphy employing the general Bayesian theory.

1.3 Objectives of the Thesis

The general aim of this work is to design methodology of quantitative scintigraphy, which helps to improve the assessment of lymphedema especially in its critical latent stage. At the same time, the demands of the methodology on data should stay restricted to be practically realizable, i.e. to get along with a limited number of measurements.

The emphasis is given on the simplified modelling of the accumulation dynamics of the radiotracer in regionally predefined parts. It forms the core of the proposed solution of the problem. Other important part of the thesis comprises employment of the proposed model for partial decision tasks useful within scintigraphy evaluation. In particular, the work solves the following specific subtasks:

Modelling of accumulation dynamics of tracer in the ROIs Design of suitable model is the necessary prerequisite for the success of the whole quantification. Bayesian approach using available data and expert knowledge is employed here.

Reconstruction of time activity curve on individual ROIs The scintigraphic response of the lymphatic system on individual ROIs, called *time activity curve*, has been the initial principal motivation for the modelling. This whole curve or some points on it form the basis of quantitative evaluations. Besides it gives further useful visual information. Therefore the whole response is reconstructed from its few measured points.

Selection of appropriate sampling times The amount of available data is a priori limited due to the routinely applicable evaluation. Consequently, it is obvious that the used mea-

surement times have a strong influence on the model estimate and consequently on results of whole quantification. Since the measurement times are not a priori fixed, their optimization is possible. A suitable manner of optimization is searched for.

Comparison of patient's limbs One of general objectives of quantitative scintigraphy is to compare the patient's limbs and decide if they differ or not. It is a very useful diagnostic aid, since, often it is known that one limb is healthy and it can act as a control one for evaluation of the other limb. The way how to treat this decision problem is proposed.

Computation of promising quantitative parameters It is expected that some quantitative parameter or a combination of them are suitable for the disease assessment. Employing the proposed model various quantities can be computed. It will be demonstrated on some quantities traditionally used in other nuclear medicine problems.

Verification In order to evaluate the proposed methodology, it has to be verified on the real scitigraphic data and compared with the clinical findings and qualitative scintigraphy where it is possible.

1.4 Layout of the Thesis

This thesis is divided in 8 Chapters and organized as follows.

The Chapter 1, you are just reading, gives a general introduction to problem treated in this thesis. It provides an overview of the current state of the art in the field scintigraphy quantification, it shows what is missing and stresses what has motivated this work. The direction of the proposed solution employing Bayesian framework together with specification of the objectives of the thesis follows.

In order to understand the function of the lymphatic system, complication of lymphedema and its evaluation, the Chapter 2 gives their description. It contains the specification of scintigraphic data available for the designed methodology of scintigraphy quantification. Also, other techniques of investigation of lymphatic system used for the verification of the proposed scintigraphic methodology are described.

The Chapter 3 provides basic ideas, statements and relations from the theory of Bayesian decision making under uncertainty. It contains only those parts of the global theory that are directly employed in the thesis.

The Chapters 4, 5, 6 and 7 present main contribution of the thesis. Chapter 4 is focused on the proposal probabilistic model that describes scintigraphic visualization of lymphatic system behaviour on the limbs of the patients. It contains general decomposition of the global observation model along the limbs of patients through suitable parametrisation together with the design of concrete models describing scintigraphic response on individual ROIs.

The parametric and predictive inference for all the considered models are expressed in the Chapter 5. It treats the partial difficulties with the computation of those inferences important to our purposes.

The Chapter 6 is focused on solution of all partial subtasks connected with scintigraphy quantification.

In Chapter 7 the proposed methodology is verified on the set of various experiments over the real scintigraphic data and the results compared with the conclusions of physicians.

Chapter 8 summarises the results of the thesis and points out to the open problems which have to be solved in the near future.

1.5 Methods

The approach is based on a systematic employment of the Bayesian theory [5, 6, 47, 73]. It is not used only for the determination of the model but the general theoretical framework is used for solving various decision problems. The whole work can be considered as the case study demonstrating the applicability of Bayesian theory to particular real-life problem.

A good numerical practice, standard and tailored approximation methods are used. The resulting algorithms are coded in *MATLAB*.

Chapter 2

Lymphedema and Lymphoscintigraphy

The aim of this Chapter is to give a description of medical and physiological background of lymphedema and its scintigraphic evaluation. This should help to understand the essence of the addressed diagnostic problem.

The basic information about the structure and function of the lymphatic system together with complication of lymphedema are outlined in the Section 2.1. The medical terms used should serve as technicians introduction into specialized language of medical articles. It covers also classification for lymphedema and its clinical evaluation.

The Section 2.2 is devoted to the scintigraphic inspection. The standard imaging procedure at the Clinics of Nuclear Medicine of the Faculty Hospital Motol is described together with its recent morphologic semi-quantitative evaluation used. At the end, the specification of scintigraphic data available for Bayesian analysis is made.

2.1 Upper Limb Lymphedema

In 1908, Handley described morbidity associated with a brawny swelling of the arm in patients with carcinoma of the breast [34]. Halsted, in 1921, described the same condition as a complication of radical mastectomy [33]. Lymphedema has been described as "a progressive pathologic state or condition characterized by chronic inflammatory fibromatosis and hypertrophy of the hypodermal and dermal connective tissues" [54]. Although now rarely seen in untreated breast carcinoma, it remains a common complication of breast cancer treatment (see Figure 2.1). Those affected have uncomfortable, unsightly and sometimes functionally impaired arm prone to repeated episodes of superficial infection. There exists the extremely rare but potentially fatal possibility of secondary lymphangiosarcoma [**?**].

Lymphedema is defined as the accumulation of an interstitial fluid in abnormally large amounts [22]. There are main two categories of lymphedema — primary and secondary. Primary lymphedema sometimes called "idiopathic" (of an unknown origin) is a genetic abnormality causing an insufficiency in lymphatic drainage. It can be present at birth, or more commonly occur at puberty or later in life. Secondary lymphedema is due to an external influence which reduces the drainage of the once normal lymphatics. Secondary lymphedema most often affects limbs. Lymphedema decreases mobility, causes embarrassment, can lead to depression, and causes a general worsening of the patient's life and health. If the lymphedema is severe, especially if more than one limb is involved, the patient is excessively heavy.

Figure 2.1: Lymphedema affected limbs.

2.1.1 Aetiology and Associated Factors

Underlying pathophysiology of breast-cancer-related-lymphedema (BCRL)is still not clearly understood. BCRL has long be considered a result of reduced lymph flow. The accepted aetiology for this condition is that dissection of the axilla interrupts the lymphatic drainage pathways from the arm, thus reducing the capacity for lymph drainage. The resulting inability to clear protein from the interstitial fluid leads to an increased protein concentration, with a reduction in the colloid osmotic pressure gradient, which acts against capillary filtration. The volume of fluid leaving the capillaries therefore increases, with edema developing until the increased hydrostatic pressure in the interstitium creates a new equilibrium, a "high protein" edema. The increased level of protein both attracts fluid into the tissues and also provokes inflammation and fibrosis within them [1]. Gradually it leads to the characteristic appearance of the brawny limb. The circulation of cellular elements such as lymphocytes and macrophages is hindered and their function may be suppressed [74]. Evidence supporting this comes from lymphangiography experiments, where radioactive tracers injected into peripheral lymphatics takes longer to reach the axilla in the swollen arm than in the normal arm. It also follows from lymphoscintigraphy studies, where removal rate of radioactive tracers from lymphedematous tissue is significantly reduced compared to the normal, contralateral arm.

Recent research indicates that the pathophysiology is more complex than a simple axillary lymphatic obstruction as a result of the cancer treatment. Stanton et al. [83] have shown that, although the clearance rate of radioactive substances from the swollen tissue in the affected arm is significantly lower than of the contralateral arm, in the non-swollen areas of the affected limb the drainage of fluid is faster than in the normal arm. One explanation for this finding is that the lymphatics in the swollen area are unable to cope with the increased resistance to flow and have become failing vessels. Consequently, a higher interstitial pressure is needed to drive the fluid along the lymphatic vessels. Fluid formation appears to be increased in the non-swollen parts of the affected arm.

Many risk factors have been cited leading to lymphedema following treatment of breast cancer. Types of treatment, both surgery and radiotherapy, are the most consistent correlates as they alter lymphatic pathways by clearance or damage of lymphatic nodes. Postoperative trauma and infection has often been put forward as a major correlative factor [60].

2.1.2 Lymphatic Physiology and Anatomy

The lymphatic system has two basic roles within the body. First, lymphoid tissue is involved in the production of cells and antibodies, destruction of red cells and other immunological tasks. Second, there is circulation of lymph, which serves to carry away from the tissues protein molecules and other large structures including particulate matter [49].

Lymphatic vessels drain the interstitial fluid, which is itself derived from the filtration and diffusion of water and molecules from blood capillaries. The forces acting upon water in this exchange are the hydrostatic pressure gradient from capillary to interstitium, and the colloid osmotic pressure gradient exerted by the plasma proteins in the other direction [69]. Two components contribute to the net flow rate in the lymphatics — lymph formation and lymph propulsion. The first describes fluid transport from the interstitium into the initial lymphatics, while the second refers to the systematic forces, which drive lymph from the initial capillaries into the collecting vessels, through the nodes and ducts, and eventually back to the blood. These two components are coupled.

Lymphatic capillaries in the connective tissue form a dense anastomosis plexus, draining into lymph-collecting vessels. These initially rely on external compression, principally secondary to striated muscle contraction, to generate flow, multiple valves ensuring unidirectional movement of lymph. In the larger, more proximal vessels, intrinsic smooth muscle within the vessel generates intermittent rhythmic contraction of the walls of the lymph vessels, the magnitude of which is apparently determined by the supply of lymph, behaviour akin to the chronotropic and ionotropic actions of the heart [56]. Superficial and deep systems are divided by the fascial layer, with a little communication between them. All collecting vessels pass through lymph nodules and nodes. The nodes receive lymph from valved afferent vessels. The lymph is then filtered through the node and passes out through a single efferent vessel, which also contains valves to prevent backflow. The nodes produce and contain lymphocytes and macrophages to phagocytose irregular cells. The trunks are the largest vessels that drain lymph from the final set of nodes into the ducts. Lymph fluid moves through the lymphatics ducts as a result of inspiration — as the diaphragm descends, the intra-abdominal pressure increases as the intra-thoracic pressure decreases. This creates gradient in the thoracic duct and encourages the lymph flow upwards. The lymph then flows back into the main circulation via the left and right subclavian veins.

In the arm, superficial drainage lymphatics from the hand and forearm converge to form radial, median and ulnar bundles. These come together above the elbow to create medial upper arm bundle, which also receives drainage from the medial part of the upper arm. This bundle, in the form of three or four proximal trunks, pierces the clavipectoral fascia to enter the axilla. The lateral part of the upper arm drains via the smaller lateral upper arm bundle, while the deep system, carrying lymph from the muscular compartments, follows the brachial artery to the axilla. The deep lymphatic vessels and medial upper arm bundle drain into the central and axillary vein groups, with efferent vessels then passing to the subclavicular group. From here, the subclavian lymphatic trunk emerges to drain into the venous system at the jugulosubclavian confluence. The lateral upper arm bundle runs in the deltopectoral groove before draining directly into the subclavicular node group [69].

2.1.3 Clinical Symptoms and Classification for Lymphedema

The typical clinical symptoms of the lymphedema are the swelling of the affected extremity, edema in the late stage, normal or pale skin colour, positive Stemmer's sign (skin over toes can barely be lifted), mounds of swelling on dorsum of foot or hand, toes swollen like sausages. A degree of arm swelling in the early postoperative period is commonly observed and tends to settle spontaneously within a matter of weeks [57]. Lymphedema may, however, develop months or

years later. Onset may be gradual, or rapid. Patients occasionally identify a precipitating factor, such as a minor infection following cut or graze, or a greater than usual degree of exercise involving the arm [93]. The swelling is initially soft, pits on pressure and may reduce with elevation. With time and recurrent superficial infection the characteristic changes of the "brawny" arm are seen; the edema is now non-pitting and accompanied by skin changes. Occasionally, this situation progresses to unmistakable "elephantiasis" [21].

Two classifications for grading of upper limb lymphedema exist. Three grades classification produced by International Society for Lymphology [69] and newer Schmidtke classification [63] preferred in this work:

Grade I: stasis without clinical symptoms of edema, only non-specific indications (sensed pressure, tension, increased fatigue of limb) are observed;

Grade II: reversible (localized) edema with evening effusion of the limb;

Grade III: tough irreversible lymphedema restricting motility and thickening the skin;

Grade IV: elephantiasis with secondary skin changes.

Unfortunately clinical staging is based on subjective criteria and is therefore subject to significant observer biases.

It is known that the swelling tends to be non-uniform along the limb. The proximal forearm and distal upper arm are commonly affected, but the hand is frequently spared, or may swell only if certain physical tasks are performed. This is the reason why the *employed clinical staging evaluates the parts of the limb separately.*

2.1.4 Treatment

Primary lymphedema and grades III and IV of secondary lymphedema are largely incurable. In the absence of a cure, precautions and prevention are emphasized. The treatment of established lymphedema varies from doing nothing at all to pursuing a host of aggressive surgical procedures, as was particularly the case in the past.

At present, conservative measures form the mainstream of management, with surgery reserved for resistant cases. The principles of conservative treatment are hygiene, elastic compression hosiery in combination with a gentle centripetal massage, exercise, the use of pneumatic compression devices and elevation. They are covered in the program of complex decongestive physiotherapy (CDP) [39]. The place of pharmacological therapy is still unclear. The use of drugs centres on attempts to encourage the breakdown of protein deposited within a limb. Though the employment of a suitable benzopyrone group of agents has serious side-effects. The used surgery techniques for the cases resistant to CDP include the debulking procedures and liposuction.

2.2 Lymphoscintigraphy

Indirect lymphoscintigraphy provides non-invasive, cost effective, highly sensitive method which is able to assess functional and morphological abnormalities of the lymphatic system on the peripherals, consequently upper limb lymphedema as well.

Technique is very simple, requiring the subcutaneous injection of a small volume of the radiopharmaceutical with external scintillation detection. The patient does not need to be hospitalized, besides there are unlimited possibilities for repetition. It readily demonstrates the location of major nodal groups and lymphatic drainage patterns at a low price and with low toxicity and radiation exposure to the patient. It provides adequate information to confirm the diagnosis of lymphedema.

An inherent limitation of lymphoscintigraphy is its poor visualization of the deep lymphatic system. Although the obtained images are poorer than those obtained by direct lymphangiography the indirect method has several advantages. The lymphangiography with oily contrast medium is technically more difficult and is uncomfortable for patient. It has been shown to damage the lymphatic vessels further and it is no longer in a common use. Besides scintigraphy provides functional information on movement of labelled protein within the lymphatic system.

Nuclear assessment of lymphatic flow can be traced back to the early 1950's. The original studies utilized Gold-198 colloid as the radiopharmaceutical. At present, ^{99m}Tc labelled colloidal particles are used. In order to make this method efficient, it is necessary to use solution with the particles of the diameter less then 50 nm for which the permeability of lymphatics is maximal. Then colloids are carried by lymphatic flow from the interstitial space through the limb and, finally, to the liver. Larger particles remain mostly at the injection site. The scintigraphic images visualize the accumulation and flow of solutions with colloids through the lymphatic system of the limbs. Its time evolution reflects the state of a patient's lymphatic system.

The scintigraphic inspections vary little bit in different centres. They differ by radiopharmaceutical used, locality of administration and acquisition parameters. Standard examination method applied at the Clinics of Nuclear Medicine of the Faculty Hospital Motol, the source of our data, is described bellow.

2.2.1 Scintigraphic Imaging

In a standardized inspection, 20 MBq of ^{99m}Tc-labelled sulphur colloid (Lymphoscint) in the smallest manageable volume (0.1–0.2 ml) is injected subcutaneously to the 1st and 4th digital web space of both hands. The aim is to follow lymphatic drainage of medial and lateral bundles of surface lymphatic system of upper limbs. The subcutaneous injection is preferred to intracutaneous one as it has been found better for lymphedema evaluation [71].

Images are obtained with the large-field-of-view gamma camera Sopha DXT with a low energy high resolution collimator LEHR (peak 140 keV, W 20%). Each image is accumulated over one minute in the acquisition matrix 64x64. Immediately after the injection, the initial image positioned to the site of injection is acquired. Then the muscular exercise commences to stimulate lymphatic flow (flexion and extension in the wrist are performed for 30 minutes). Next images of the whole arm are acquired in semi-flexion with a led shielding on hand and wrist, 30 and 180 minutes after administration (see Figure 2.2). In some cases, the examination is complemented by additional images. Markers on all images indicate elbows, wrist and shoulders. For the purpose of quantitative evaluation experiments, the number of acquired images has been increased. The majority patient data in experimental set includes 3–4 images besides the initial scan.

2.2.2 Qualitative Evaluation

For the healthy limb early images should demonstrate the expected major lymphatic vessels. Several axillary nodes are visualised later. Sometimes the accumulation in 1 to 3 cubit nodules is observed. Rarely, nodules at supra-clavicular region appear and a manifestation of pectoral nodule has been also referred.

Abnormal findings include dermal backflow, absent or faintly visualized regional lymph nodes, presence of collateral lymphatic vessels and no or barely and delayed visualized lymphatic vessels [38]. No axillary lymph nodes visualisation and retention of the radiopharmaceutical in the locality of the administration correlates with severe lymphedema in clinical grading and poor results of CDP. For patients without axilla externation and radiotherapy, the missing visualization of lymphatic nodules indicates the serious lymphedema. A few lymph nodes revealed in the axilla

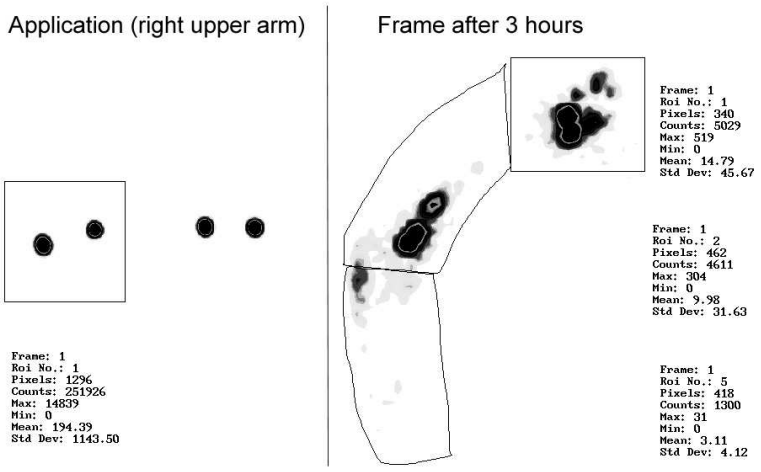


Figure 2.2: Typical scintigraphic images of upper arm with definition of ROIs. The left part is the initial image obtained over the injection sites of both limbs. The right one is the image of the whole arm where also the individual ROIs are defined.

with marked outflow to the subclavicular lymph nodes and delayed transport of the radioactive material correlates with moderate lymphedema in clinical grading and slightly better effects of CDP. Direct although delayed lymph outflow to the supraclavicular lymph nodes shows patients with clinically mild lymphedema achieving the best results in CDP. At the mild and latent forms, discrete dermal backflow can occur, mostly at the distal part of the forearm. Rarely, increased repletion of lymphatic vessels or collaterals is visualized.

Due to the discrete character of the findings the employed qualitative evaluation is done for the parts of the limb separately.

2.2.3 Arrangement for Quantitative Evaluation

The reliable quantitative evaluation is still missing in spite of many attempts done. Their overview is given in the Section 1.2. The arrangement of scintigraphic data used for the designed methodology follows.

The raw (non-reduced) images present the maximal information that can be gained from the scintigraphic imaging. For the purpose of proposed quantitative evaluation, the information from the whole images is reduced into the integral counts reflecting the accumulated activity on the pre-specified regions of interest (ROI).

The ROIs are drown around the axillary region and around the the forearm and upper arm (see Figure 2.2). The accumulated activities on the ROIs are aggregated as the total integral counts over each ROI corrected for physical decay of radionuclide (the physical half-life of Technetium is $T_p = 6.023$ hours). Also the total counts from the initial image reflecting administered activity are stored. It allows us to calibrate the data to the injected amount.

The available data for the Bayesian data analysis are the sequences of the counts reflecting the accumulated activities on the ROIs of axilla, upper arm and forearm in several times and the the initial "injected" activity count from the injection site. For the verification of the proposed methodology the clinical findings, Section 2.1.3, and outcomes of the qualitative scintigraphic evaluation, Section 2.2.2, are also available.

Chapter 3

Bayesian Decision Making

Theory of statistical decision-making under uncertainty [5, 90] provides a unified logical structure which can be used for solving various problems ranging from parameter estimation, prediction, pattern recognition, learning, testing of hypothesis etc. up to feedback control. At a general level, the theory helps the decision maker to select one (optimal) of the available options (decisions), when the uncertainty has to be taken into account. These options relate to a part of the real world, to a system, which is to be described or influenced.

The Chapter starts with formal introduction of a decision making under uncertainty problem in the Section 3.1. The Section 3.2 is then devoted to the used Bayesian solution called Bayesian theory. Its aim is not to give full overview of the theory (it can be found e.g. in [6]) but to extract the basic principles, statements and relations needed. The technical mathematical background is suppressed for clarity reasons.

3.1 Decision Making Under Uncertainty

We spend a considerable proportion of our lives in a state of uncertainty. This uncertainty may relate the past situations, where direct knowledge or evidence is not available, or to present or future developments which are not yet completed. Whatever the circumstances, all states may be described in the same way: namely, as an individual feeling of incomplete knowledge in relation to a considered situation.

Many feelings of uncertainty are rather insubstantial and we do not seek to analyse them. On the other hand, we all regularly encounter uncertain situations in which we at least aspire to behave "rationally" in some sense. This might be because we face the direct practical problem of choosing from among a set of possible actions, where each involves a range of uncertain consequences and we are concerned to avoid making an "illogical" choice. Alternatively, we might be called upon to summarise our beliefs about uncertain aspects of the situation, bearing in mind it will be used subsequently as the basis for choosing an action. More specifically, we might regard the summary itself, i.e. the choice of particular mode of representing beliefs, as being a form of action to which certain criteria of "rationality" might be directly applied.

To choose the best among a set of actions would, in principle, be immediate if we had perfect information about the consequences to which they would lead. The presented theory solves the decision problems for which such information is not available, and we must take *uncertainty* into account as a major feature of the problem.

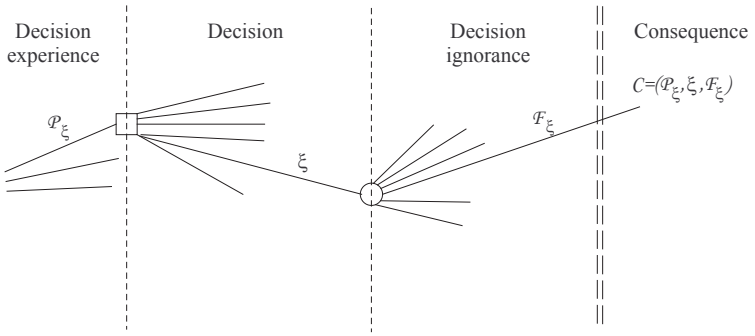


Figure 3.1: Decision tree.

The general structure of decision problem over real system involving choices under uncertainty is determined by five basic elements:

(i) decision $\xi \in \xi^*$ which is directly selected by the decision maker from the set of available actions (options);

(ii) (decision) experience $\mathcal{P}_\xi \in \mathcal{P}_\xi^*$ contains knowledge available to the decision maker for the decision ξ;

(iii) (decision) ignorance $\mathcal{F}_\xi \in \mathcal{F}_\xi^*$ specifies knowledge unavailable to the the decision maker for the decision ξ, i.e. the uncertainty expressing both incomplete knowledge and randomness. $\mathcal{F}_\xi^*$ contains uncertain events describing uncertain outcomes for each pair $(\mathcal{P}_\xi, \xi)$.

(iv) consequences $\mathcal{C}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$ corresponding to each realisation $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$ of system behaviour $\mathcal{Q}^*$;

(v) the relation $\leq$, which expresses the decision maker's preferences between pairs of actions, so that $\xi_1 \leq \xi_2$ signifies that ξ_1 is no preferred to ξ_2

The idea is as follows. Suppose we choose an action ξ having available information $\mathcal{P}_\xi$; then one and only one of the set of uncertain events $\mathcal{F}_\xi^*$ occurs and leads to the consequence $\mathcal{C}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$.

The decision problem can be presented schematically by means of a decision tree as shown in Figure 3.1. The square represents a decision node, where the choice of an action is required. The circle represents an uncertainty node, where the outcome is beyond our (total) control. Following the experience available, choice of an action and the occurrence of particular event the branch leads to the corresponding consequence.

The described problem presents general static decision task where the single decision is done. Of course, many practical problems involve sequential considerations but, as shown in [6], they reduce, essentially, to repeated analysis based on the above structure (through backward induction). Besides this work gets along with the framework of static decision task.

An individual's perception of the state of uncertainty resulting from the choice of any particular ξ is very much dependent on the information currently available $\mathcal{P}_\xi$. The representation in Figure 3.1 only captures the structure of the decision problem as perceived at a particular point in time. Preferences (v) about the uncertain scenarios resulting from the choice of actions depend on *attitudes to the consequences involved and assessment of the uncertainties attached to*

the corresponding events. While the former are independent of the state of information concerning relevant events (consequences are defined on the space $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$), the latter are clearly subject to change as the new information is acquired and it may well change overall preferences among the various courses of action.

3.2 Bayesian Approach

The Bayesian theory of decision making under uncertainty is derived on the axiomatic basis, with intuitive operational appeal, applied over the general structure of decision problem described in the Section 3.1 (see e.g. [6]). It uses the dual concept of probability , providing the quantitative numerical measure of the uncertainty attached to each event, and utility (loss) , providing numerical measure of value for consequences.

The probability in Bayesian view is not interpreted in terms of limits of relative frequencies but more generally as a *subjective* degree of belief of a rationally and consistently reasoning person (designer) which is used to describe quantitatively considered uncertainty.

Full description of the general static decision making in the Bayesian framework is given hereafter in the Section 3.2.2 followed in Section 3.2.3 by the description of Bayesian learning — technique which modifies the degrees of belief attached to the uncertain events according the new acquired information. The important preliminaries precede them in Section 3.2.1. Section 3.2.4 summarizes elements occurring in the decision tasks with practical instances from problem of lymphoscintigraphy quantification. The particular decision tasks employing the general framework are listed in the Section 3.2.5 employing the general framework of static decision task. At the end, the advantages and disadvantages of Bayesian methodology are pointed out in the Section 3.2.6 and the used particular probability distributions are listed in the Section 3.2.7.

3.2.1 Preliminaries

In Bayesian view random means uncertain. Any quantity and event are therefore random and uncertainty has a probability structure. This Section presents basic conventions and rules of the probability calculus for the random quantities respected in the whole work.

Quantity is a mapping with a numerical range.

Realization is a value of the quantity. The quantity and its realization are not distinguish, as usual. The proper meaning is determined by the context.

f is letter reserved for both probability functions (*pf*) of discrete quantities and probability density functions (*pdf*) of quantities of continuous type. The meaning of the *p(d)f* is given through the identifier of its argument. Implicitly all the general relations are defined for the quantities of the continuous type. One has only to keep in mind that the integration has to be replaced by regular summation whenever the argument is discrete.

Calculus with *pdf*s For generic possibly multivariate random quantities $(\alpha, \beta, \gamma) \in (\alpha, \beta, \gamma)^*$ and arbitrary function $\mathcal{Z}(\alpha, \beta, \gamma)$ it holds:

Non-negativity $f(\alpha, \beta|\,\gamma)\,,\ f(\alpha|\,\beta, \gamma)\,,\ f(\beta|\,\alpha, \gamma)\,,\ f(\beta|\,\gamma)\,, \geq 0.$

Normalization $\int_{\alpha,\beta^*} f(\alpha, \beta|\,\gamma)\ d\alpha d\beta = \int_{\alpha^*} f(\alpha|\,\beta, \gamma)\ d\alpha = \int_{\beta^*} f(\beta|\,\alpha, \gamma)\ d\beta = 1.$

Chain rule $f(\alpha, \beta|\,\gamma) = f(\alpha|\,\beta, \gamma)\, f(\beta|\,\gamma) = f(\beta|\,\alpha, \gamma)\, f(\alpha|\,\gamma).$

Marginalization $f(\beta|\,\gamma) = \int_{\alpha^*} f(\alpha, \beta|\,\gamma)\ d\alpha,\ f(\alpha|\,\gamma) = \int_{\beta^*} f(\alpha, \beta|\,\gamma)\ d\beta.$

Bayes rule

$$f(\beta|\,\alpha,\gamma) = \frac{f(\alpha|\,\beta,\gamma)\,f(\beta|\,\gamma)}{f(\alpha|\,\gamma)} \propto f(\alpha|\,\beta,\gamma)\,f(\beta|\,\gamma)\,. \tag{3.1}$$

The proportionality sign, $\propto$ means that the factor independent of β and uniquely determined by the normalization is not explicitly written in the equality represented.

Conditional independence of quantities α and β under the condition γ

$$f(\alpha,\beta|\,\gamma) = f(\alpha|\,\gamma)\,f(\beta|\,\gamma) \Leftrightarrow f(\alpha|\,\beta,\gamma) = f(\alpha|\,\gamma) \text{ or } f(\beta|\,\alpha,\gamma) = f(\beta|\,\gamma) \tag{3.2}$$

Expectation of function $\mathcal{Z}(\alpha,\beta,\gamma)$ under the condition γ

$$E\left[Z(\alpha,\beta,\gamma)|\,\gamma\right] = \int_{\alpha,\beta^*} Z(\alpha,\beta,\gamma) f(\alpha,\beta|\,\gamma)\,d\alpha d\beta \tag{3.3}$$

***Pdf* of transformed quantity** $\beta \equiv T(\alpha)$ If T is bijection (one-to-one mapping) of multivariate quantity α, $T:\ \alpha^* \to \beta^* \equiv T(\alpha^*)$, with finite continuous partial derivatives on α^*, $J_{ij}(\alpha) \equiv \partial T_i(\alpha)/\partial\alpha_j$, then

$$f(T(\alpha))\,|J(\alpha)| = f(\alpha) \tag{3.4}$$

where $|\cdot|$ denotes determinant .

3.2.2 Static Decision Task

The static decision making deals with the design and use of a single decision rule $\mathcal{R}$. The optimal one from the group of admissible decision rules is searched for. Under term admissible it is understood that such rules are causal, i.e. $\mathcal{R}:\ \mathcal{P}_\xi^* \to \xi^*$, and meet the "technological" restrictions.

The selection is based on the principle that the consequences $\mathcal{C}(\mathcal{Q})$ can be quantified, i.e. that the degree of the achievement of the designer's aim can be expressed numerically. *Utility* function $\mathcal{U}:\ \mathcal{Q}^* \to \Re$ (alternatively *loss* function $\mathcal{Z}:\ \mathcal{Q}^* \to \Re$) measures the quality of consequences $\mathcal{C}(\mathcal{Q})$. The greater the value of $\mathcal{U}(\mathcal{Q})$ (alternatively smaller value of $\mathcal{Z}(\mathcal{Q})$) the better. They give the preferences among the consequences, though they are insufficient for ordering of decision rules (options).

The overall numerical measure of value has to be attached to decision rule (option), which depends both on uncertain events and consequences to which these events lead in order to select the optimal rule. The only decision criterion compatible with the axiom system says [6]: The optimal decision rule $\mathcal{R}^o$

$$\mathcal{R}^o(\mathcal{P}_\xi) \equiv \xi^o(\mathcal{P}_\xi),\ \forall \mathcal{P}_\xi \in \mathcal{P}_\xi^*$$

is such admissible decision rule $\mathcal{R}$ which minimises the expected loss function

$$\xi^o(\mathcal{P}_\xi) \in \arg\min_{\xi\in\xi^*} E\left[\mathcal{Z}\left(\mathcal{P}_\xi,\,\xi,\,\mathcal{F}_\xi\right)|\,\mathcal{P}_\xi,\,\xi\right]. \tag{3.5}$$

Such decision rule is constructed value-wise by assigning to each $\mathcal{P}_\xi \in \mathcal{P}_\xi^*$ a minimising argument $\xi^o(\mathcal{P}_\xi)$ in (3.5). If there is more absolutely minimising arguments $\xi^o(\mathcal{P}_\xi)$ (3.5) implies no preferences. If the utility function is used instead of the loss function then the expected utility is maximised.

As the uncertainty has probability structure in the Bayesian view, the loss is considered as a "random quantity", contingent on the occurrence of uncertain events. Then the expectation (3.5) is identical with mathematical expectation (3.3).

All the particular static decision tasks can be solved within this general framework. The specific instances differ only in:

- Decomposition of the behaviour $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi) \equiv$(experience, decision,ignorance).
- decision rules $\mathcal{R} \in \mathcal{R}^*$ determined by the domain $\mathcal{P}_\xi^*$ and range ξ^*
- loss function $\mathcal{Z}: \ \mathcal{Q}^* \equiv (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)^* \to \Re$

These elements determine the outer probabilistic model $f(\mathcal{F}_\xi | \mathcal{P}_\xi, \xi)$ needed for the minimisation (3.5) where the minimised expectation is computed according (3.3) as

$$E[\mathcal{Z}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi) | \mathcal{P}_\xi, \xi] \equiv \int_{\mathcal{F}_\xi^*} \mathcal{Z}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi) f(\mathcal{F}_\xi | \mathcal{P}_\xi, \xi) d\mathcal{F}_\xi. \tag{3.6}$$

The *pdf* $f(\mathcal{F}_\xi | \mathcal{P}_\xi, \xi)$ represents a model in the form suitable for the design of decision rule by which the designer describes the real system with all the considered uncertainties. Construction of such model is designer's business.

Adopted natural conditions of decision making

In general, the ignorance part $\mathcal{F}_\xi$ in the outer model necessary for minimization 3.6 depends on the potential decision rule expressed by *pdf* $f(\mathcal{F}_\xi | \mathcal{P}_\xi, \xi)$. It means that it is influenced by it. However, there exist variety of decision tasks, where the designed decision rule has no influence on considered ignorance. Thus decision ξ is superfluous in the conditioning, i.e.

$$f(\mathcal{F}_\xi | \mathcal{P}_\xi, \xi) = f(\mathcal{F}_\xi | \mathcal{P}_\xi). \tag{3.7}$$

It catches the adopted version of natural conditions of decision making [47]. Although the decision is not present in condition of outer model (3.7) its optimization is needed as it is still included in the loss function $\mathcal{Z}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$.

3.2.3 Learning

The previous Section 3.2.2 concerns the design of the decision rule in particular point of time where the fixed experience (information) $\mathcal{P}_\xi$ is at disposal. This Section deals with the question, how the decision task changes if the new information is acquired and how the overall model is modified by it. It corresponds to the case where the designer needs to revise his decision in the light of new information. Such information is mostly relevant piece of evidence or data Y obtained on the real system.

New data Y causes that the old decomposition of the behaviour, $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$, is naturally modified to the new one, $\mathcal{Q} = (\tilde{\mathcal{P}}_\xi, \xi, \tilde{\mathcal{F}}_\xi)$, where

$$\tilde{\mathcal{P}}_\xi = (\mathcal{P}_\xi, Y), \quad \tilde{\mathcal{F}}_\xi = \mathcal{F}_\xi \setminus Y.$$

Data Y originally included in the ignorance $\mathcal{F}_\xi$ moves to the experience $\tilde{\mathcal{P}}_\xi$ of the new decomposition. It is the only change in the framework of the static decision task which implies that the expectation $E\left[\mathcal{Z}\left(\tilde{\mathcal{P}}_\xi, \xi, \tilde{\mathcal{F}}_\xi\right) | Y, \mathcal{P}_\xi, \xi\right]$ is minimised, so the outer model $f\left(\tilde{\mathcal{F}}_\xi | Y, \mathcal{P}_\xi, \xi\right)$ is needed.

Update of the old model $f\left(\mathcal{F}_\xi|\,\mathcal{P}_\xi,\,\xi\right)$ to the new version $f\left(\tilde{\mathcal{F}}_\xi|\,Y,\mathcal{P}_\xi,\xi\right)$ captures the *learning* through the new data Y. Bayes' theorem (3.1) offers the straight way how to realize this learning process:

$$f\left(\tilde{\mathcal{F}}_\xi|\,Y,\mathcal{P}_\xi,\xi\right)=\frac{f\left(\tilde{\mathcal{F}}_\xi,Y|\,\mathcal{P}_\xi,\xi\right)}{f\left(Y|\,\mathcal{P}_\xi,\xi\right)}=\frac{f\left(Y|\,\tilde{\mathcal{F}}_\xi,\mathcal{P}_\xi,\xi\right)f\left(\tilde{\mathcal{F}}_\xi|\,\mathcal{P}_\xi,\xi\right)}{f\left(Y|\,\mathcal{P}_\xi,\xi\right)}. \tag{3.8}$$

The *pdf* $f\left(\tilde{\mathcal{F}}_\xi|\,\mathcal{P}_\xi,\,\xi\right)$ represents *prior* (to data) model of unknown part $\tilde{\mathcal{F}}_\xi$, that allows us to introduce information based on an expert knowledge. In order to make the Bayesian learning about the part $\tilde{\mathcal{F}}_\xi$ possible, a model that relates it with the observed data Y, $f\left(Y|\,\tilde{\mathcal{F}}_\xi,\mathcal{P}_\xi,\xi\right)$ has to be determined. The *pdf* $f\left(Y|\,\mathcal{P}_\xi,\xi\right)$ is the relevant outer model of the observed quantities from the Bayesian viewpoint. The considered ignorance part $\tilde{\mathcal{F}}_\xi$ contains those ignorance of our interest. Generally, it can contain quantities that are never observed directly and, in spite of this, we want to describe them. In summary, the Bayesian learning combines three information sources: (i) prior model of the ignorance $\tilde{\mathcal{F}}_\xi$, $f\left(\tilde{\mathcal{F}}_\xi|\,\mathcal{P}_\xi,\,\xi\right)$, (ii) observed data Y, and (iii) model that relates observations with ignorance part, $f\left(Y|\,\tilde{\mathcal{F}}_\xi,\mathcal{P}_\xi,\xi\right)$. In other words it corrects (i) the prior model of the ignorance part, by means of (ii) data and (iii) model relating the ignorance and data into the posterior model of ignorance.

In order to avoid cumbersome notation, explicit notational reference to the initial state of experience $\mathcal{P}_\xi$ is commonly omitted, e.g. $f\left(\tilde{\mathcal{F}}_\xi|\,\mathcal{P}_\xi,\,\xi\right)\equiv f\left(\tilde{\mathcal{F}}_\xi|\,\xi\right)$.

Parametric modelling

The *pdf* $f\left(Y|\,\mathcal{P}_\xi,\xi\right)$ is an important outer model describing the observed quantities Y on the real system, that is always needed. Though, its exact construction in this form for non-trivial cases is not realizable. The common and popular way how to overcome this problem is to employ parametric modelling.

If a suitable parametrised model is constructed

$$f\left(Y|\,\Theta,\mathcal{P}_\xi,\xi\right)$$

that is known up to unknown parameters Θ, then the outer model $f\left(Y|\,\mathcal{P}_\xi,\xi\right)$ can be simply gained as:

$$f\left(Y|\,\mathcal{P}_\xi,\xi\right)=\int_{\Theta^*}f\left(Y|\,\Theta,\mathcal{P}_\xi,\xi\right)f\left(\Theta|\,\mathcal{P}_\xi,\xi\right)d\Theta, \tag{3.9}$$

where $f\left(\Theta|\,\mathcal{P}_\xi,\xi\right)$ is *prior pdf* of the parameter Θ, that allows us to introduce prior information what is known in advance about parameters by analysis or experience. Thus, the joint-distribution model $f\left(Y,\Theta|\,\mathcal{P}_\xi,\xi\right)$ (and consequently $f\left(Y|\,\mathcal{P}_\xi,\xi\right)$ too) is determined by the pair of of parametrised model $f\left(Y|\,\mathcal{P}_\xi,\xi\right)$ and prior model of the parameter Θ, $f\left(\Theta|\,\mathcal{P}_\xi,\xi\right)$.

Parameter Θ is a part of ignorance, i.e. $\Theta\in\tilde{\mathcal{F}}_\xi$, as it is not directly observable. So the parametrised model $f\left(Y|\,\Theta,\mathcal{P}_\xi,\xi\right)$ belongs to the group of models $f\left(Y|\,\tilde{\mathcal{F}}_\xi,\mathcal{P}_\xi,\xi\right)$ and prior *pdf* $f\left(\Theta|\,\mathcal{P}_\xi,\xi\right)$ to $f\left(\tilde{\mathcal{F}}_\xi|\,\mathcal{P}_\xi,\xi\right)$. Consequently, according (3.8) the Bayesian generalized estimate of parameter Θ is

$$f\left(\Theta|\,Y,\mathcal{P}_\xi,\xi\right)\propto f\left(Y|\,\Theta,\mathcal{P}_\xi,\xi\right)f\left(\Theta|\,\mathcal{P}_\xi,\xi\right).$$

If the posterior *pdf* $f\left(\Theta|\,Y,\mathcal{P}_\xi,\xi\right)$ has the same functional form as the posterior *pdf* $f\left(\Theta|\,\mathcal{P}_\xi,\xi\right)$, then the prior *pdf* is called *conjugate* or self-reproducing with respect to model $f\left(Y|\,\Theta,\mathcal{P}_\xi,\xi\right)$, see e.g. [5].

3.2.4 Construction Elements

The aim of this Section is to summarize elements occurring in all decision tasks. Their choice and use represent typically specific decision subtasks that have to be harmonised with the final aim considered. All the elements are commented and supplemented by the information from the concrete problem of quantitative lymphoscintigraphy.

Data

Data Y connect the artificial world of evaluations with reality. Their information content is crucial for the success of the decision making that use them. They belong to the decision experience $Y \subset \mathcal{P}_\xi$. If the designer has an opportunity to influence the information content of data he should do that by a proper experimental design [6, 47].

Generally the observable data can be split into two groups by definition of system influence. System *input* is understood as directly manipulable "decision" that influences the system behaviour while the system *output* is an observable quantity that informs the decision maker about the behaviour of the system. This decomposition is relative according the solved problem.

By the definition, the injection of radiotracer to the limb is the only system *input*, but it is not directly measured here. All the scintigraphic data belongs to system output, even the initial image measuring the injected amount of radiotracer. The scintigraphic imaging is already standardized procedure, so there are not many possibilities to change any conditions. The only possibility how to influence the information content of measured data is to select and change their measurements times.

The obtained data can realize aspects that are a priori uninteresting for the problem at hands. Thus, a wise and commonly used practice is to suppress uninteresting details in data before further data processing [66, 64] . Such pre-processing makes pre-projection to a simplified world. However, it influences the quality of the final decision, so it should be done carefully.

Aggregation of counts from the pixels of scintigraphic images within the ROIs expresses the overall activity in the individual ROIs. Remotion of radionuclide decay is done to suppress its role in the accumulation dynamics. The normalization of the scintigraphic data with respect to injected amount has an important role, as it makes the data from various limbs directly comparable.

Decision

At the beginning of solving decision making tasks there is a formulation of the addressed problem. The major role has the definition what should be decided. It means to name all elements of the decision ξ. The restrictions arising from various considerations (e.g. economic or safety, but also causality restrictions) have to be considered in the design.

Within the designed scintigraphy quantification, the final major decisions are assessment of the disease stages and quantitative comparison of limbs. Reconstruction of time activity curves, optimization of sampling times, etc. are partial minor but important decision tasks to make the scintigraphy quantification maximally useful.

Model

Generally, model $f\left(\mathcal{F}_\xi \middle| \mathcal{P}_\xi, \xi\right)$ in (3.6) is necessary for the design of the decision rule, see Section 3.2.2. To make learning from data Y possible, i.e. to revise model necessary for decision according the new data, it is replaced by the pair of models $f\left(\tilde{\mathcal{F}}_\xi \middle| \mathcal{P}_\xi, \xi\right)$ and $f\left(Y \middle| \tilde{\mathcal{F}}_\xi, \mathcal{P}_\xi, \xi\right)$ which have to be determined, see Section 3.2.3.

In the case, where the outer model of observed quantities is constructed by means of parametric modelling (3.9), the considered parameters Θ belong to ignorance part $\tilde{\mathcal{F}}_\xi$ and the pair of models $f(Y|\Theta, \mathcal{P}_\xi, \xi)$ and $f(\Theta|\mathcal{P}_\xi, \xi)$ is necessary. The former defines traditionally used parametrised model of data Y if the parameters of model are known while the latter allows to express further information about the unknown parameters (it is dissimilarity from other techniques). If other considered elements besides the parameters Θ belong to ignorance part $\tilde{\mathcal{F}}_\xi$ they have to appear in these models besides the parameters.

Lymphoscintigraphy quantification is based on modelling of scintigraphic response on the limbs. Parametrised model describes this response by few parameters while the prior distribution of parameters narrows the group of considered models according the expert knowledge.

Loss function

The *loss* function $\mathcal{Z}(\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$ quantifies the consequences of $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi)$, i.e. the degree of the achievement of the designer's aim. It should express these wishes as precisely as possible. The choice of the *loss* function should be also harmonised with the model adopted. If it is violated then design provides non-acceptable decision rule.

The reconstructed time activity curve and other quantifiers of interest are presented for simplicity by means of point estimate — expected value, that is a consequence of used popular *quadratic* loss, see Remark(s) 3.2.1. Other decision tasks employs the loss functions constructed with respect to models employed.

3.2.5 Specific Cases of Static Decision Task

The specific well-known instances of employment of general static decision task, Section 3.2.2, used in this work are listed bellow.

Point estimation

Point estimation can be cast in the considered framework as follows:

- $\mathcal{Q} = (Y, \hat{\Theta}, \Theta) \equiv$(data at disposal, point estimate, unknown parameter)
- Admissible rules are of the form $\mathcal{R}: \; Y^* \to \hat{\Theta}^*, \Theta^* \subset \hat{\Theta}^*$
- Loss function $\mathcal{Z}$ measures a distance of $\hat{\Theta}$ and Θ. This distance may generally depend on data Y, too.

The ignorance $\mathcal{F}_{\hat{\Theta}}$ coincides with the unknown parameter Θ. Thus, the outer model of the system needed for decision making is $f\left(\Theta|Y, \hat{\Theta}\right)$. The adopted natural conditions of decision making (3.7) imply that the decision $\hat{\Theta}$ is superfluous in the conditioning. Thus, the needed model coincides with the generalized Bayesian estimate $f(\Theta|Y)$. For given Y, the optimal point estimate is

$$\hat{\Theta} \in \arg\min_{\hat{\Theta}\in\hat{\Theta}^*} \int_{\Theta^*} \mathcal{Z}\left(Y, \hat{\Theta}, \Theta\right) f(\Theta|Y)\, d\Theta.$$

Remark(s) 3.2.1

(i) The Bayes point estimate with respect to the specific popular quadratic *form of the loss function* $\mathcal{Z}\left(Y, \hat{\Theta}, \Theta\right) = (\Theta - \hat{\Theta})^T \boldsymbol{M} (\Theta - \hat{\Theta})$ *is the mean of* $f(\Theta|Y)$, $\hat{\Theta} = E[\Theta|Y]$, *if* $\boldsymbol{M}^{-1}$ *exists.*

(ii) The point estimate can be done for any part of multidimensional parameter Θ or its some function. Task of parameter point estimation can be similarly adapted for point prediction of potentially observable data that are part of ignorance set.

Testing of hypothesis

Hypothesis testing, a selection of the best variant among several alternatives, can be put within our framework as follows.

- $\mathcal{Q} = (Y, \hat{h}, \{H_h\}_{h\in h^*})$ ≡(data at disposal, estimate $\hat{h} \in h^* \equiv \{1, \ldots, \mathring{h} < \infty\}$ of h, {hypothesis list})
- Admissible rules are of the form $\mathcal{R}: \ Y^* \to h^*$
- Loss function $\mathcal{Z}$ is $(\mathring{h}, \mathring{h})$-table with non-negative entries $\mathcal{Z}\left(Y, \hat{h}, h\right)$, usually with zero diagonal as no penalty is paid when $\hat{h} = h$.

Outer model needed is $f\left(h|\, Y, \hat{h}\right)$. The adopted natural conditions of decision making (3.7) imply that the decision $\hat{h}$ is superfluous in conditioning. Thus, we need $f\left(h|\, Y\right)$. For a given Y, the optimal decision in hypothesis testing is

$$\hat{h} \in \arg\min_{\hat{h}\in\hat{h}^*} \sum_{h\in h^*} \mathcal{Z}\left(Y, \hat{h}, h\right) f\left(h|\, Y\right).$$

Remark(s) 3.2.2

(i) Unlike in classical hypothesis testing [75], the testing is performed within a completely specified set of alternatives.

(ii) If a pair of hypotheses is compared, $\mathring{h} = 2$, the decision rule constructed according the described methodology coincides with the celebrated Neymann-Pearson lemma [15].

In this case, off diagonal elements of $\mathcal{Z}\left(\hat{h}, h\right)$ penalise the classical errors of the 1st and 2nd kind [75].

(iii) Selection of the best model from the set of considered alternative models can be treated in this framework.

Reporting beliefs as the decision problem

It can seem that problems of reporting inferences do not fall within the framework of decision tasks, however, converse can be simply demonstrated [6].

Let's consider the decision task which aim is to select from the list of various inference statements $\hat{f}\left(\Theta|\, Y\right)$ about unknown parameter Θ to be the best alternative for reporting beliefs about Θ. The reason why the inference statement is selected from $\hat{f}^*$ is that computation of actual belief $f\left(\Theta|\, Y\right)$ may be difficult. Then the tractable approximation $\hat{f}\left(\Theta|\, Y\right)$ in $\hat{f}^*$ is sought, which is in some sense "close" to $f\left(\Theta|\, Y\right)$. Then it can be cast in the considered framework as follows.

- $\mathcal{Q} = (Y, \hat{f}, \Theta)$ ≡(data at disposal, reported inference $\hat{f}$ about parameter Θ, unknown parameter)

- Admissible rules are of the form $\mathcal{R}: \; Y^* \rightarrow \hat{f}^*, \; \hat{f}(\Theta|\, Y) \geq 0, \; \int_{\Theta^*} \hat{f}(\Theta|\, Y)\, d\Theta = 1,$ $(\hat{f}(\Theta|\, Y))$ is *pdf*

- Loss function $\mathcal{Z}\left(Y, \hat{f}, \Theta\right)$ measures the quality of inference report $\hat{f}$ about Θ.

The outer model needed is $f\left(\Theta|\, Y, \hat{f}\right)$. The adopted natural conditions of decision making (3.7) imply that the chosen report $\hat{f}$ is superfluous in conditioning. Thus, the needed model coincides with the generalized Bayesian estimate $f(\Theta|\, Y)$. For given Y, the optimal inference reported is

$$\hat{f} \in \arg\min_{\hat{f} \in \hat{f}^*} \int_{\Theta^*} \mathcal{Z}\left(Y, \hat{f}, \Theta\right) f(\Theta|\, Y)\, d\Theta. \tag{3.10}$$

The considered proper loss function $\mathcal{Z}$ in (3.10) should respect here, that the best action is to state one's actual belief, i.e. $f(\Theta|\, Y)$. The utility functions $\mathcal{U}$ (i.e. the negative loss function $\mathcal{Z}$) for this problem is often called *score* functions . For the problem of reporting *pure inference* statements, it is natural to restrict to the class of *local* loss functions, $\mathcal{Z}\left(Y, \hat{f}, \Theta\right) \equiv \mathcal{Z}\left(\hat{f}(\Theta|\, Y)\right)$. The *local* loss functions assess the value of distribution, $\hat{f}$, according the probability it assigns to the "actual" outcome Θ. The only one existing smooth proper local loss function is of the form (for more details see [6]):

$$\mathcal{Z}_{\log}\left(\hat{f}(\Theta|\, Y)\right) = -A \log \hat{f}(\Theta|\, Y) - B(\Theta), \quad A > 0, \tag{3.11}$$

i.e. it holds for (3.11)

$$E\left[\mathcal{Z}_{\log}\left(\hat{f}(\Theta|\, Y)\right)|\, Y\right] \geq E\left[\mathcal{Z}_{\log}\left(f(\Theta|\, Y)\right)|\, Y\right].$$

Consequently, if preferences are described by a logarithmic loss function (3.11), the expected loss of reporting $\hat{f}(\Theta|\, Y)$ rather then $f(\Theta|\, Y)$ representing actual beliefs, is given by

$$\begin{aligned} \mathcal{D}\left(\hat{f}(\Theta|\, Y) \,||\, f(\Theta|\, Y)\right) &= E\left[\mathcal{Z}_{\log}\left(\hat{f}(\Theta|\, Y)\right)|\, Y\right] - E\left[\mathcal{Z}_{\log}\left(f(\Theta|\, Y)\right)|\, Y\right] \\ &= A \int_{\Theta^*} f(\Theta|\, Y) \log \frac{f(\Theta|\, Y)}{\hat{f}(\Theta|\, Y)} d\Theta, \quad A > 0, \end{aligned} \tag{3.12}$$

Moreover, $\mathcal{D}(\cdot||\, \cdot)$ is non-negative and is zero, if and only if, $\hat{f}(\Theta|\, Y) = f(\Theta|\, Y)$.

Remark(s) 3.2.3

(i) The measure (3.12) is natural general measure of discrepancy *between a distribution and its approximation (for simplicity, $A = 1$).*

(ii) The quantity $\mathcal{D}\left(\hat{f}||\, f\right)$, which arises here as a difference between two expected losses, was introduced by Kullback and Leibler [51] as an ad hoc *measure of directed divergence between two* pdf*s. This the reason of the name* Kullback-Leibler distance.

3.2.6 Advantages and Disadvantages of Bayesian Approach

Summary of the advantages and disadvantages closes this Chapter concerning the Bayesian decision making under uncertainty.

Choice of Bayesian methodology is determined especially by the following reasons:

- It combines in consistent way all accessible information sources, specifically theoretical description of the given system, measured data and expert knowledge and experience.
- It does not rely on an asymptotic behaviour of estimates.
- It yields the information about uncertainty of the estimate and employs it also for the decision tasks.
- It follows from previous items that this methodology is suitable to be used in cases of small and uncertain data.
- Learning and decision making phases are bound up in consistent way.

On the other hand, groups of analytical and numerical disadvantages occur with using Bayesian methodology. Among them belong:

- Impossibility to integrate *pdf*s analytically. This problem causes difficulties to normalize the posterior *pdf* (see e.g. (3.8)) and also prevents to integrate-out redundant quantities analytically [**?**, 86].
- Advantage of prior information is sometimes considered as the disadvantage, due to the problems how to treat it.
- "Curse of dimensionality" — problem of task dimension growing with increasing amount of data [77].
- Computational obstacles to evaluate non-normalized function (see Section A.1).

3.2.7 Particular Distributions Used

In this Section, particular univariate and multivariate distributions used in the thesis are listed. The books [43, 44, 45] provides a mass of detail on these and others distributions.

These distributions provide the building blocks for *statistical models* and are defined in terms of "parameters". Even in many cases these "parameters" coincide with considered parameters Θ appearing in parametric modelling, see Section 3.2.3, for the present, "parameters" should be simply regarded as "labels" of the various mathematical functions we shall be considering.

The Poisson distribution

A discrete random quantity x has a *Poisson* distribution with parameter λ ($\lambda > 0$) if its *pf* $\mathrm{Pn}(x|\,\lambda)$ is

$$\mathrm{Pn}(x|\,\lambda) = \exp(-\lambda)\frac{\lambda^x}{x!}, \quad x = 0,\, 1,\, 2, \ldots \tag{3.13}$$

The mean and variance are given by $E\,[x] = V\,[x] = \lambda$.

The continuous uniform distribution

A continuous random quantity x has a *continuous uniform* distribution with parameters (α, β) if its *pdf* $\mathrm{Un}(x|\,\alpha, \beta)$ is

$$\mathrm{Un}(x|\,\alpha, \beta) = (\beta - alpha)^{-1}, \quad \alpha < x < \beta \tag{3.14}$$

It has mean $E[x] = (\alpha + \beta)/2$ and variance $V[x] = (\beta - \alpha)^2/12$.

The gamma distribution

A continuous random quantity x has a *gamma* distribution with parameters (α, β) $(\alpha > 0,\ \beta > 0)$ if its *pdf* $\mathrm{Ga}(x|\,\alpha, \beta)$ is

$$\mathrm{Ga}(x|\,\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x > 0, \tag{3.15}$$

where $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)$ is Euler gamma function. Its moments are $E[x] = \alpha/\beta$ and $V[x] = \alpha/\beta^2$. If $\alpha = 1$, x is said to have an exponential distribution.

The Poisson-gamma distribution

A discrete random quantity x has a *Poisson-gamma* distribution with parameters (α, β, n) $(\alpha > 0,\ \beta > 0,\ n = 1, 2, \ldots)$ if its *pf* $\mathrm{Pg}(x|\,\alpha, \beta, n)$ is

$$\mathrm{Pg}(x|\,\alpha, \beta, n) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + x)\, n^x}{x!\,(\beta + n)^{\alpha+x}}, \quad x = 0,\ 1,\ 2, \ldots. \tag{3.16}$$

The distribution is generated by the mixture

$$\mathrm{Pg}(x|\,\alpha, \beta, n) = \int_0^\infty \mathrm{Pn}(x|\,n\lambda)\mathrm{Ga}(\lambda|\,\alpha, \beta) d\lambda.$$

The mean is $E[x] = n\alpha/\beta$, and the variance is $V[x] = n\alpha(\beta + n)/\beta^2$.

The normal distribution

A continuous random vector X of dimension k has a *multivariate normal Gaussian* distribution, with parameters $(M, \mathbf{\Lambda})$ $(M \in \Re^k,\ \mathbf{\Lambda}$ is $k \times k$ symmetric positive-definite matrix) if its *pdf* $\mathrm{N}(X|\,M, \mathbf{\Lambda})$ is

$$\mathrm{N}(X|\,M, \mathbf{\Lambda}) = \frac{|\mathbf{\Lambda}|^{1/2}}{(2\pi)^{k/2}} \exp\left(-\frac{1}{2}(X - M)^T \mathbf{\Lambda}(X - M)\right), \quad X \in \Re^k, \tag{3.17}$$

where $|\cdot|$ denotes determinant. The distribution is symmetrical about $X = M$. Its mean is $E[X] = M$ and the covariance matrix is $Cov[X] = \mathbf{\Lambda}^{-1}$, so that $\mathbf{\Lambda}$ here represents the *precision matrix* (inversion of covariance matrix) of the distribution.

If $k = 1$, so that M and $\mathbf{\Lambda}$ are scalars μ and λ, $\mathrm{N}(X|\,M, \mathbf{\Lambda})$ reduces to univariate normal *pdf* $\mathrm{N}(x|\,\mu, \lambda)$.

The Student distribution

A continuous random vector X of dimension k has a *multivariate Student* distribution, with parameters $(M, \mathbf{\Lambda}, \alpha)$ ($M \in \Re^k$, $\mathbf{\Lambda}$ is $k \times k$ symmetric positive-definite matrix, $\alpha > 0$) if its *pdf* $\mathrm{St}(X|\, M, \mathbf{\Lambda}, \alpha)$ is

$$\mathrm{St}(X|\, M, \mathbf{\Lambda}, \alpha) = \frac{\Gamma((\alpha+k)/2)|\mathbf{\Lambda}|^{1/2}}{\Gamma(\alpha/2)(\alpha\pi)^{k/2}} \left[1 + \frac{1}{\alpha}(X-M)^T \mathbf{\Lambda}(X-M)\right]^{-(\alpha+k)/2}, \quad X \in \Re^k. \quad (3.18)$$

If $k = 1$, so that M and $\mathbf{\Lambda}$ are scalars μ and λ, then it reduces to univariate Student (t) *pdf* $\mathrm{St}(x|\, \mu, \lambda, \alpha)$. The distribution is symmetrical about $X = M$. Its mean is $E[X] = M$ and the covariance matrix is $Cov[X] = \mathbf{\Lambda}^{-1}(\alpha/(\alpha-2))$.

The marginal *pdf* of vector X, constructed from joint distribution of (X, y) having the *pdf* of normal-gamma distribution, has Student distribution:

$$\mathrm{St}(X|\, M, \alpha\beta^{-1}\mathbf{\Lambda}, 2\alpha) = \int_0^\infty \mathrm{N}(X|\, M, \mathbf{\Lambda}y)\, \mathrm{Ga}(y|\, \alpha, \beta) dy.$$

Chapter 4

Modelling of Scintigraphy Response of Lymphatic System

The objective of this Chapter is to introduce the outer probabilistic model that describes scintigraphic visualization of the lymphatic system behaviour on the limbs of the patients involved in the study. Such model should serve for the description of responses of the lymphatic system on individual limbs during the scintigraphy only, giving up modelling of lymphatic system in general. The assumption is that the state of the lymphatic system influencing its behaviour manifests itself in the scintigraphic response, see Section 1.2.

The available scintigraphic data directly used in models are introduced in the Section 4.1. The Section 4.2 deals with the definition of the general outer probabilistic model of observed quantities describing the solved problem. Since the parametric modelling is employed, each from the pair, parametrised model and prior distribution about its parameters are naturally decomposed along the limbs. The simplification of limb structured global model to the set of parallel models on individual ROIs is defined at the end of the Section. Finally, the Section 4.3 is devoted to the choice of particular ROI model(s) employing the introduced general structuring and decomposition.

4.1 Data as a Link to Reality

The choice of theoretically available data (records of observable quantities) for the task of the Bayesian inference and decision making plays a key role on the result obtained. The basic informations about the scintigraphic data at disposal were given in Section 2.2.3 but till this occasion no formal notation has been given about the data employed in evaluation. Here is the right place to do that.

The raw 2-D data of scintigrams reduced to summarized scalar integral counts over ROIs and initial counts over the injection site are the raw data obtained for the analysis. For our purpose, two versions of used data will be considered. The former are the original raw data. The latter are normalized data, i.e. the relative activities on ROIs normalized with respect to that at initial injection site.

Raw Data Let t be the number of minutes since the administration time at $t = 0$. Then, for each limb l, set of sequences of non-negative integer counts ${}^{R}Y_{l,r}$ reflecting activities at the ROIs $r = 1, 2, 3$ (forearm, upper arm, axilla) together with the initial non-negative integer "injected"

activity count ${}^{R}y_{l,I}$ at the injection site represent all the raw data observed on the limb:

$$ {}^{R}Y_l = \left({}^{R}Y_{l,r=1}, {}^{R}Y_{l,r=2}, {}^{R}Y_{l,r=3}, {}^{R}y_{l,I}\right). $$

The superscript R denotes here the raw data while subscripts r and l refer to limb and ROI. The sequences of ROIs counts ${}^{R}Y_{l,r} = ({}^{R}y_{l,r;t_{l,1}}, {}^{R}y_{l,r;t_{l,2}}, \ldots, {}^{R}y_{l,r;t_{l,n_l}})$ are recorded at the same sampling times $\mathcal{T}_l = (t_{l,1}, t_{l,2}, \ldots, t_{l,n_l})$ for all ROIs. Their number n_l, is small, typically two or three.

Normalized data are derived from the raw data. They include only the sequences of relative activities ${}^{N}Y_{l,r}$ (superscript N denotes the normalized) on the individual ROIs defined as absolute sequences ${}^{R}Y_{l,r}$ normalized by initial count ${}^{R}y_{l,I}$ at the injection site:

$$ {}^{N}Y_l = \left({}^{N}Y_{l,r=1}, {}^{N}Y_{l,r=2}, {}^{N}Y_{l,r=3}\right), \quad \text{where } {}^{N}Y_{l,r} = {}^{R}Y_{l,r}/{}^{R}y_{l,I}. $$

Superscript N distinguishes here the normalized data. Such relative activities are real non-negative numbers. In this case, the information about the initial activity ${}^{R}y_{l,I}$ is superfluous and it is not included into observed data.

The reason why to normalize the data is to filter out the differences of injected amount among limbs and patients beforehand. We can focus then immediately on "standardized" responses on limbs that are appropriate for evaluation of the lymphatic system and avoid uninteresting amplification caused by injection.

For the sake of simplicity, the observable quantities on limbs in next general text are marked Y_l for both cases of considered data. The different marking ${}^{R}Y_l$ vs ${}^{N}Y_l$ for raw and normalized data is used again whenever we need to distinguish these versions.

4.2 Global Observation Probabilistic Model

Model, in general, is a suitable description of the real system. Mathematical model can be comprehended as a form of the relation between the quantities (both observable and non-observable on the system). Since all the quantities are understood as random in Bayesian context, the relation is defined in probabilistic terms too.

Observable quantities — their observations form data — have the major role as link between the reality and theoretical modelling. For that reason, the probability measure of observable quantities is the suitable model, which mathematically specifies subjectivist's degree of belief about them. Thinking in the whole context, the global observation model spans over all patients whose at least one limb is included in the scintigraphic study. Having the observable quantities Y_l on $\mathring{l}$ individual limbs defined in Section 4.1 the global joint model over limbs has the form:

$$ f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}}\right). \tag{4.1} $$

The specification of the predictive probability (4.1) is the task of the actual problem modelling.

The most general way how to determine it is to employ nonparametric models of empirical distribution functions and to assess their belief over function space [6]. But the assessing and representing such belief distribution over the set of all possible distribution functions is by no means straightforward, since such distribution functions are, effectively, an infinite-dimensional objects. In what follows, it is convenient to restrict attention to those cases where a corresponding representation holds in terms of density functions, labelled by a finite-dimensional parameter.

Conventionally, the representations in the finite-dimensional case are referred as parametric models, see Section 3.2.3. Let Θ be the finite-dimensional parameter, then using basic probability calculus the model (4.1) can be gained by marginalization of a model extended on Θ:

$$f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) = \int_{\Theta^*} f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}} \middle| \Theta\right) f\left(\Theta\right) d\Theta. \tag{4.2}$$

Consequently the model (4.1) is determined by the parametrised model $f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}} \middle| \Theta\right)$ and by the prior distribution $f\left(\Theta\right)$ on its parameter Θ according (4.2).

4.2.1 Limb Structured Parametrised Model

The use of specific parametrised models can often be given by a motivation or justification as the coherent representation of certain forms of belief. In this case the general parametrised model $f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}} \middle| \Theta\right)$ can be naturally structured along the limbs. Incorporating the individual limb related parameters $\Theta_l \subset \Theta$ together with supposed conditional independence (3.2) of observed quantities Y_l given these parameters, the global parametrised model can be expressed as a product of individual limb parametrised models

$$f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}} \middle| \Theta\right) = \prod_{l=1}^{\mathring{l}} f\left(Y_l \middle| \Theta\right) = \prod_{l=1}^{\mathring{l}} f\left(Y_l \middle| \Theta_l, l\right) = \prod_{l=1}^{\mathring{l}} f\left(Y_l \middle| \Theta_l\right). \tag{4.3}$$

The same structural forms are used for all limbs in the study, i.e. $f\left(Y_l \middle| \Theta_l, l\right) = f\left(Y_l \middle| \Theta_l\right)$, since it is supposed that the limb parametrised model is rich enough to cover almost all possible variants of scintigraphy responses of limb lymphatic system.

The definition of the limb related parameters Θ_l is general and does not try to describe any relations between them. However their specification is also one part of modelling. For our purpose we further distinguish two simple versions. Nevertheless, other versions are also possible.

Strictly limb local parameters It is the case where all the parameters are considered to be local for limbs. No part of the parameters describes common quality. In this case, the global parameter Θ is the Cartesian product of the local parameters Θ_l,

$$\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_{\mathring{l}}). \tag{4.4}$$

Limb local parameters with a common part ${}^{[c]}\Theta$ Sometimes it is useful to consider one part of parameters to be common for all limbs. It can represent e.g. the condition of standardized investigation, which is the same for all limbs. Then, employing a decomposition of the limb parameters Θ_l to the part common for all limbs ${}^{[c]}\Theta$ and the part strictly individual ${}^{[i]}\Theta_l$, i.e. $\Theta_l = ({}^{[c]}\Theta, {}^{[i]}\Theta_l)$, the global parameter consists of

$$\Theta = \left({}^{[c]}\Theta, {}^{[i]}\Theta_1, {}^{[i]}\Theta_2, \ldots, {}^{[i]}\Theta_{\mathring{l}}\right). \tag{4.5}$$

4.2.2 Limb Structured Prior Distribution

To make the probability model (4.2) complete, prior distribution $f\left(\Theta\right)$ about unknown parameters Θ has be determined. The decompositions (4.4) and (4.5) of the global parameter space simplify this task.

Strictly limb local parameters The specification is simpler for the parameter space of strictly limb local parameters (4.4). Supposed sufficient separation of peripheral lymphatic system of limbs enables us to define even the limb parameters to be mutually independent, see 3.2. Thus, the prior distribution about Θ is the product of the limb local priors of the same structural form,

$$f(\Theta) = \prod_{l=1}^{\mathring{l}} f(\Theta_l), \tag{4.6}$$

where the more detailed specification of the limb local prior $f(\Theta_l)$ depends on a selection of specific model and its parameters.

Limb local parameters with common part ${}^{[c]}\Theta$ For the second case (4.5) similar idea can be applied. The most general way is to define:

$$f(\Theta) = \prod_{l=1}^{\mathring{l}} f\left({}^{[i]}\Theta_l \middle|\, {}^{[c]}\Theta\right) f\left({}^{[c]}\Theta\right). \tag{4.7}$$

4.2.3 ROI Separated Models

The pair of limb structured parametrised model (4.3) and prior distribution (4.6) alternatively (4.7) defines the observation probabilistic model on the data of whole limbs. It means all ROIs are covered together in the limb model. It is the most general description. Though the considered utilization of whole modelling by doctors together with the obstacle of missing data on some ROIs for some limbs force us to simplify such a model and decompose it to local models on individual ROIs.

The advantage of such simplified description is that for single-ROI model only its data are necessary, not those of other ROIs. It reflects the opinion of doctors: If one ROI proves to be significant in evaluation of lymphedema, it is possible to reduce scintigraphic measurements only on this ROI. On the other hand, such a decomposition does not allow to express directly the spatial relations and interdependence among the ROIs.

The single ROI models are studied in this thesis only. The ROIs are evaluated in parallel without any mutual relations. The global model decomposition along limbs is assumed to stay valid. It only concerns individual ROIs. The alternatives of parametrised model on one ROI for two versions of data follow. The subscript r at parameters and data denoting the ROI can be omitted as model belongs to a single ROI.

Parametrised model on a single ROI and limb for raw data

$$f(Y_l|\,\Theta_l) = f\left({}^{R}Y_{l,r}, {}^{R}y_{l,I}|\,\Theta_l\right)$$

Parametrised model on a single ROI and limb for normalized data

$$f(Y_l|\,\Theta_l) = f\left({}^{N}Y_{l,r}|\,\Theta_l\right)$$

The structuring of prior distribution $f(\Theta)$ along limbs remains the same as in the Section 4.2.2, but it concerns a single ROI.

Remark(s) 4.2.1

Let's notice that even the parametrised models for raw data are constructed independently on individual ROIs, all they contain common injected activity count ${}^{R}y_{l,I}$.

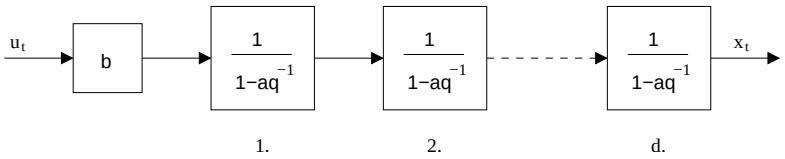


Figure 4.1: Cascade model whose impulse response models the lymphoscintigraphy time activity curves (q^{-1} is the backward shift operator, i.e. $q^{-1}x_t = x_{t-1}$).

4.3 ROI Model Choice

The discussion about probabilistic model has been held in the general way, up to now, concerning only its decomposition along limbs. This Section contains a description of the proposed particular model on the ROI and limb inserted into general decomposition.

The structure of the involved observation model is considered common for all ROIs. To distinguish ROIs, various settings of prior distributions are used. Both observation parametrised models for two versions of data include the common inner part describing the relative response of lymphatic system, see Section 4.3.1, while the outer parts, described in Section 4.3.2, differ according to the consideration of noisy samples for the versions of data. The Section 4.3.3 then summarizes all available information about established parameters together with the specification of suitable prior distributions.

4.3.1 Relative Lymphatic System Response

The model of the relative lymphatic system response is the main model, that attempts to describe a behaviour (i.e. response) of the lymphatic system on individual ROI during scintigraphy. It employs input-output discrete-time dynamic modelling verified from the system theory [16, 66]. To simplify notation, the subscripts labelling the limb and region are omitted as the same model for all ROIs and limbs is considered.

The image sampling is the natural reason to use discrete-time description. Let t be the number of minutes since the administration time at $t = 0$. Hence, real time is $\tau(t) = t\Delta$, $\Delta = 1$ (min). The injected tracer is a known input into the lymphatic limb system up to its absolute amount. The relative amount of the injected tracer is a unit impulse, i.e. $u_t = \delta_t$ where $\delta_0 = 1$ and $\delta_t = 0$, for $t > 0$. The discrete-time scalar *impulse response*, x_t, of the lymphatic system at the chosen ROI, known as a *time activity curve*, is the relative amount of the tracer in the ROI at time $\tau(t)$. Causality implies that $x_t = 0$ for $t < 0$.

The dynamic model relating the sequences u_t and x_t (see Figure 4.1) is chosen as a cascade of first-order linear models, with a common parameter a for each of the d sections, and with a single lumped gain parameter b. It is chosen as a compromise between the complex distributed nature of the lymphatic system and the need for a model with a few of unknown parameters. This cascading of simple sections describes the gradual penetration of tracer through the limb. To avoid misunderstanding it is necessary to stress, that the cascading of sections is in no dependence with physical ROI partition. It expresses only the abstract space partition.

Binomial expansion of the system denominator leads to the difference equation:

$$x_t = -\sum_{i=1}^{d} \binom{d}{i} (-a)^i x_{t-i} + bu_t.$$

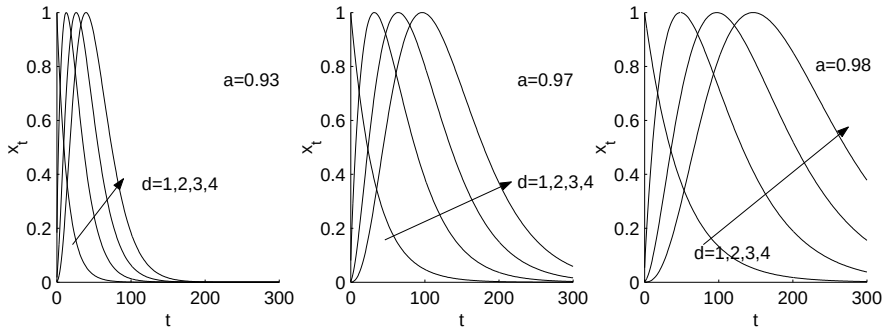


Figure 4.2: Typical time activity curves for modelling the time evolution of the lymphatic system at a particular ROI (b is normalized so that curve maximum is 1 for all cases).

For $u_t = \delta(t)$, x_t models the *impulse response* of the lymphatic system observed at times t at the particular ROI. Its closed form solution [65] is:

$$x_t = b\binom{t+d-1}{t}a^t, \quad t \geq 0. \tag{4.8}$$

A rich signal ensemble is generated by the proposed parametrised model. It successfully captures with only three free parameters the stable, slowly-decaying, non-oscillatory nature of the lymphatic system responses at a particular ROI. In particular, the order parameter, d, allows a rich set of candidate curves to be explored. Typical curves are illustrated in Figure 4.2, for b normalized so the maximum of all curves is 1.

The complete time activity curve is $X = (x_0, x_1, \ldots, x_t, \ldots)^T$, where ^{T} denotes transposition,

$$X = b\widetilde{A}_{\Xi}, \quad x_t = b\widetilde{A}_{\Xi,t+1}, \quad \widetilde{A}_{\Xi,t+1} = \tbinom{t+d-1}{t}a^t, \quad t = 0, 1, \ldots, \quad \Xi \equiv (a, d). \tag{4.9}$$

The points of the time activity curve in observation (sampling) times are gathered into the n-vector $X_o = bA_{\Xi}$ for the subset of times $\tau(t)$, $t \in \mathcal{T}$ ($A_{\Xi} = \widetilde{A}_{\Xi}(\mathcal{T})$).

4.3.2 Observation Parametrised Model

The inner hidden response of the lymphatic system, i.e. the amount of tracer on the ROI, is measured indirectly from the outside by means of scintigraphy counts. The noisy samples of the "absolute" time activity curve together with initial injection amount are observed in the raw image data. The individual counts on image pixels can be described by the the Poisson distribution [4], see Section 3.2.7. The raw source data available for the analysis are then obtained by summation of pixel noisy counts over ROIs.

The sequel considerations about the suitable observation model differ on the version of data and employ different degree of approximation.

Poisson observation model for raw data If the pixel counts have Poisson distribution then also the aggregated counts over pixels on the ROI can be well described by the Poisson distribution (the sum of independent quantities with Poisson distribution keeps at least the distribution type). Then the noisy sample of the initial injection amount can be described by Poisson distribution with the single parameter g_l representing its expectation. Parameter g_l

defines the degree of lymphatic system excitation, i.e. the actual amount of injected tracer. Similarly, the observations of individual points of absolute time activity curve at times $t_{l,i} \in \mathcal{T}_l$ can be described by means of Poisson distribution. Their expectations are given as the product of corresponding points $x_{t_{l,i}}$ of relative time activity curve and initial gain parameter g_l, i.e. $g_l x_{t_{l,i}}$. Long inter-sample intervals imply that observations are approximately conditionally independent. Thus, the parametrised model for the (n_l+1)-series of observations $Y_l = ({}^R Y_l, {}^R y_{l,I})$ is *pf*:

$$f\left(Y_l \middle| \Theta_l\right) = f\left({}^R Y_l, {}^R y_{l,I} \middle| \Theta_l\right) = \mathrm{Pn}({}^R y_{l,I} \middle| g_l) \prod_{i=1}^{n_l} \mathrm{Pn}({}^R y_{l;t_{l,i}} \middle| g_l b_l A_{\Xi_l,i}), \tag{4.10}$$

where $\mathrm{Pn}(y|\lambda)$ *pf* (3.13) of the Poisson distribution of a discrete quantity y determined by the common mean and variance parameter λ and $A_{\Xi_{l,i}}$ is i-th element of vector A_{Ξ_l}. This observation model is parametrised by the quadruple $\Theta_l = (a_l, d_l, b_l, g_l) = (\Xi_l, b_l, g_l)$, $(\Xi_l \equiv (a_l, d_l))$.

Briefly, the observation model (4.10) relates the observations Y_l to the parameters (Ξ_l, b_l) describing relative individual accumulation dynamics on the ROI (4.9) and to the input gain parameter g_l characterizing injected amount and influencing only the scale of the absolute accumulation dynamics.

The similar idea of the Poisson model for the normalized data has no sense as the normalized data are real values derived as the ratio of two non-negative integer counts.

Normal observation model for normalized data If we consider again that the absolute measured aggregated counts on ROI and injection site have Poisson distributions then the relative normalized counts are proportion of two quantities having Poisson distributions. It is not reasonable to employ it exactly here, therefore suitable approximation is in demand.

Since the Poisson distribution with the large value of its expectation parameter is close to Gaussian distribution, the aggregation and normalization of great number of counts permits the overall noise process to be approximated well by additive, zero-mean, normal noise $e_{l;t_{l,i}}$, i.e. ${}^N y_{l;t_{l,i}} = x_{l;t_{l,i}} + e_{l;t_{l,i}}$. The precision s_l of $e_{l;t_{l,i}}$ can be assumed (approximately) constant. Again, long inter-sample intervals imply that $e_{t_{l,i}}$, $t_{l,i} \in \mathcal{T}_l$, are approximately conditionally independent. Thus, the probability density function (*pdf*) for the n_l-vector of observations $Y_l = {}^N Y_l$ is:

$$f\left(Y_l \middle| \Theta_l\right) = f\left({}^N Y_l \middle| \Theta_l\right) = \mathrm{N}({}^N Y_l \middle| b_l A_{\Xi_l}, s_l \boldsymbol{I}_{n_l}) = \left(\frac{s_l}{2\pi}\right)^{\frac{n_l}{2}} \exp\left(-\frac{s_l}{2} || {}^N Y_l - b_l A_{\Xi_l} ||^2\right). \tag{4.11}$$

$\mathrm{N}(Y|M, \boldsymbol{\Lambda})$ stands for *pdf* (3.17) of normal distribution of continuous vector Y determined by its mean M and precision matrix $\boldsymbol{\Lambda}$. $\boldsymbol{I}_{n_l}$ denotes identity matrix of dimension n_l and $||\cdot||$ the Euclidean norm. The observation model is parametrised here by the quadruple $\Theta_l = (a_l, d_l, b_l, s_l) = (\Xi_l, b_l, s_l)$.

In summary, the observation model (4.11) relates the available observations Y_l to the parameters (Ξ_l, b_l) describing individual accumulation dynamics (4.9) and the common precision parameter s_l characterizing measurement process. It is necessary to note here, that this model based on the noise approximation is simplification as it permits existence of the negative activities, i.e. it assigns the nonzero probability also to negative values that in the reality can not exist. Though, high signal to noise ratio makes influence of this approximation negligible.

The Gaussian distribution models the absolute error in data. Therefore it is reasonable to employ the proposed approximative normal model only on the normalized data. It is unsuitable for the absolute raw data due to generally different injected activities for individual limbs resulting in different data levels.

Θ entry	Prior ranges
$d_l = \Xi_{l,2}$	$1 = \underline{d} \leq d_l \leq \overline{d}$
$a_l = \Xi_{l,1}$	$\underline{a}_{d_l} < a_l < \overline{a}_{d_l} (\Leftarrow \underline{t}_{\max} < t_{\max} < \overline{t}_{\max})$
b_l	$\underline{b}_{\Xi_l} \leq b_l \leq \overline{b}_{\Xi_l} (\Leftarrow \underline{r}_{\max} \leq r_{\max,l} \leq \overline{r}_{\max})$
s	$s > 0$ $(\underline{s} < s < \overline{s})$
g_l	$\underline{g} < g_l < \overline{g}$

Table 4.1: Prior ranges of unknown parameters.

4.3.3 Prior Distribution

The parameters Θ_l in the limb observation models are unknown. To make the probability model complete, it requires the prior distribution about these parameters $f(\Theta_l)$ to be elicited.

The precision parameter, s_l, in model for normalized data (4.11) reflects only the measurement process, neither the property of the patient, limb nor the particular ROI. Thus, it can be considered to be common parameter for all limbs, i.e. $s = s_1 = s_2 = \ldots = s_{\tilde{l}}$, and for its estimation, data from various limbs can be used. The parameter g_l in model for raw data (4.10) seems in the first view as common for all limbs too. It describes the injected amount of tracer, which should be same or similar in the standardized examination. However, in many cases it is not possible to inject the whole amount of the prepared radiotracer in syringe due to various difficulties Therefore this parameter is considered strictly limb local. The remaining three parameters, $(a_l, d_l, b_l) = (\Xi_l, b_l)$, common for both models are strictly local to the limb of the patient, describing local accumulation dynamics. This triple is of our main interest for lymphedema evaluation. They have to be estimated using two or three local patient-specific measurements. This is impossible without prior information, which is rich in this case. This is the key advantage of the Bayesian paradigm in the inference of diagnostically significant quantities from the sparse data. The prior information is expressed through intervals of a priori possible Θ_l, given formally in Table 4.1, commented and specified hereafter.

Model order d_l The parameter describes the penetration rate through the limb and modifies the shape of time activity curve. The upper bound $\overline{d}$ was chosen in accordance with the supposed curve shape comparing it with model responses for various orders d_l. $\overline{d} = 5$ was selected as the conservative upper bound. For higher orders of model the shape of the curve becomes less suitable. The model order $d_l = 1$ is also permitted (i.e. $\underline{d} = 1$) even the first order model is standardly used for the description of depot clearance at the place of injection and the ROI is shifted from the depot. Thus for distant ROIs from injection site the higher orders should be employed. However for the near ROIs it is still suitable. There is one special case where the first order model is just suitable, even for more distant ROIs. It is case where the response is very low (about under 1% relative response). Then the higher orders cannot grasp the quick relative increase of response with respect to its maximum but slow to absolute values.

Partial time constant a_l The inspected responses for this parameter in the range $0 < a_l < 1$ are stable and non-oscillatory. It catches the basic assumption, that the response during scintigraphy subsides in time, i.e. the tracer is washed out from limb. The oscillations can exist in general, but they are caused by changing mode of limb movement. If they are eliminated, the interval can been shrunk to reflect practically observed slow accumulation dynamics. There emerges query how to set relatively simply this range if also the model order d_l significantly influence the response. Therefore information about the supposed

time interval of the maximum of the response, $\underline{t}_{\max} < t_{\max} < \overline{t}_{\max}(\min)$, is employed for the setting of intervals of parameter a_l. The maximum of the response for model with parameters $\Xi_l = (a_l, d_l)$ is reached at time

$$t_{\max} = \arg\max_t x_t = \text{floor}\left(\frac{d_l a_l - 1}{1 - a_l}\right) + 1, \tag{4.12}$$

so the time interval can be simply transformed to the interval on a_l for individual orders d_l. The exception is the model of the first order, $d_l = 1$, where the maximum is $\forall\ a_l$ at time $t = 0$. Then the time interval is related to the points where the response descents to the half of its maximum. Different setting of time ranges is used for individual ROIs. Time ranges were chosen $25 < t_{\max} < 250$ for forearm ROI, $60 < t_{\max} < 300$ for upper arm ROI and $100 < t_{\max} < 360$ for axilla ROI.

Lumped gain b_l The response is non-negative and cannot exceed the applied input $u_0 = 1$. This condition can be transformed to the upper bound $\overline{b}_{\Xi_l}$ using $t_{\max}$ from (4.12):

$$\overline{b}_{\Xi_l} = \left[\binom{t_{\max} + d_l - 1}{t} a_l^{t_{\max}}\right]^{-1}. \tag{4.13}$$

The necessary physical maximum is correct but the practical experience have shown that the response maximum is smaller. Therefore the upper bound of the response maximum $\overline{r}_{\max}$, consequently $\overline{b}_{\Xi_l}$, can be cautiously decreased. The harmless upper bound for all the ROIs has been selected so the maximum of the response cannot exceed 20% of the applied input, i.e. $\overline{r}_{\max} = 20\%$. Similarly to the upper bound, the positive lower bound $\underline{r}_{\max} = 0.01\%$ (consequently $\underline{b}_{\Xi_l}$) is determined. It expresses the distinguished nonzero relative activity.

Noise precision s Precision has to be non-negative. The initial prior range of the absolute precision has been chosen according the normalized supposed highest level of noise in the original raw data described by Poisson model. The exact values of the range is $1e5 < s < 1e8$. This range can be gradually refined by processing of many sets of ROI data. The common precision is *a priori* independent of other parameters.

Injection gain g_l The injected amount of tracer has to be positive and of a sufficient value in order to make scintigraphic imaging practicable and to distinguish it from the background radiation. The upper bound $\overline{g}$ is given by the standardized preparation of solution into syringe. The injected amount into limb is whole amount of syringe in ideal conditions. If some difficulties appears it is smaller. The independence to other parameters follows from it. The common interval $\underline{g} < g_l < \overline{g}$ has been selected conservatively to the range $1e5 < g_l < 4e5$.

The mixed-type prior distribution $f(\Theta_l)$ can be written as a product of conditional distributions in accordance with the described relationships above.

Prior distribution for Poisson model on raw data

$$f(\Theta_l) = f(\Xi_l, b_l, g_l) = f(b_l|\, a_l, d_l)\, f(a_l|\, d_l)\, f(d_l)\, f(g_l) \tag{4.14}$$

Prior distribution for normal model on normalized data

$$f(\Theta_l) = f(\Xi_l, b_l, s) = f(b_l|\, a_l, d_l)\, f(a_l|\, d_l)\, f(d_l)\, f(s) \tag{4.15}$$

Individual factors in (4.14) and (4.15) are discussed bellow.

The ranges of all the parameters listed is useful information, but the strict adherence to such parameter intervals in prior distributions would make the technical difficulties for the model analysis. The reason is that the included finite integrals cannot be evaluated analytically.

In general, numerical methods are often necessary to analyse models with constrained parameters; see [28, 86] for employed discretization or for the use of MC sampling in this context. Fortunately, the prior ranges of all the parameters do not have the same significance. The parameters of accumulation dynamics (Ξ_l, b_l) are those significant and their ranges are well specified. The priors for the remaining parameters g_l and s can be replaced by the appropriate unconstrained distributions.

Since further detailed information is unavailable, the (conditional) distributions on the a priori expected ranges of parameters a_l, b_l, d_l in (4.14) and (4.15) have been chosen uniform, see (3.14), justified via the principle of insufficient reason [41]. To get a feasible solution, the continuous parameters a_l, b_l have been discretized. Hence, the *pdf*s $f(a_l|\, d_l)$ and $f(b_l|\, a_l, d_l)$ are replaced by *pf*s on a discrete grids.

Besides, the alternative prior distribution on the a priori expected ranges of parameters a_l, b_l, d_l have been chosen which is based on a little different idea. Let's notice that the parameter a_l enters observation model in a nonlinear way (4.8). Is then reasonable to use the uniform distribution on this parameter? This is commonly discussed question in applications of Bayesian theory how to choose the so called non-informative prior distribution [73]. Choosing the "non-informative" uniform distribution on the model parameter its uniform characteristic is not preserved, according to (3.2.1), for another parameter which is a non-linear function of the former one but they are both alternatives for the model definition.

Herein the idea is based on the assumption that the global characteristics of the accumulation dynamics are of our interest. Such proper global characteristic is the area under the time activity curve. Accordingly, the uniform non-informative distribution should be defined on it. Yet, the area under the time activity curve has to be positive. So the uniform distribution of its logarithm seems to be better. It has the advantage that with decreasing area the "sensitivity" of the distribution is increasing. The logarithm of the area under the time activity can be expressed as $\log(b_l) - d_l \log(1 - a_l)$. Consequently the uniform distribution is chosen for d_l but the uniform distribution together with discretization for the remaining parameters a_l, b_l is chosen on the scales $\log(b_l)$ and $\log(1 - a_l)$.

The remaining *pdf*s on continuous-valued g_l for raw data case and s for model on normalized data are expressed in conjugate form, see Section 3.2.3 (details can be found in [5]). This flexible choice simplifies evaluations. In both cases the conjugate form of prior *pdf* is gamma distribution (3.15),

$$f(g_l) \;=\; \mathrm{Ga}(g_l|\, \alpha_g, \beta_g) = \frac{\beta_g^{\alpha_g}}{\Gamma(\alpha_g)} {g_l}^{\alpha_g - 1} \exp(-\beta_g g_l), \quad g_l > 0 \tag{4.16}$$

$$f(s) \;=\; \mathrm{Ga}(s|\, \alpha_s, \beta_s), \quad s > 0 \tag{4.17}$$

determined by pairs of nonnegative hyper-parameters (α_g, β_g) and (α_s, β_s). It has the mean $E[\bullet] = \alpha_\bullet/\beta_\bullet$ and variance $V[\bullet] = \alpha_\bullet/\beta_\bullet^2$. Use of the Gaussian approximations for individual *pdf*s and the physical confidence intervals with half-width equal to one standard deviation and centered around the mean, give the choice

$$\alpha_\bullet = \left(\frac{\underline{\bullet}_l + \overline{\bullet}_l}{\overline{\bullet}_l - \underline{\bullet}_l}\right)^2, \quad \beta_\bullet = \frac{2(\underline{\bullet}_l + \overline{\bullet}_l)}{(\overline{\bullet}_l - \underline{\bullet}_l)^2}, \quad \bullet = g \text{ or } s. \tag{4.18}$$

A more resolved choice is unnecessary as the intervals serve only as a conservative guess for specification of the prior *pdf*s. The injection gain g_l is considered as an auxiliary parameter

to relate time activity curve to raw data, so this setting of relatively flat prior distribution is sufficient. For the parameter s, it initializes the estimation of common s, which can be improved with each patients' limb.

Besides, instead of usage of uniform distribution on b_l and $\log(b_l)$ for the model on the normalized data the conjugate alternative is considered too [5]. It will decrease the dimension of space where the numerical technique is necessary to apply. However, it forces us to modify the prior distribution (4.14) and to loose the exact boundaries on a priori possible values. The conjugate prior *pdf* for the parameter b_l in normal observation model on normalized data (4.11) is the normal *pdf* (eq:normal) in the form:

$$f\left(b_l|\,\Xi_l, s\right) = \mathrm{N}(b_l|\,\hat{b}_{\Xi_l}, \omega_{\Xi_l} s) = \left(\frac{\omega_{\Xi_l} s}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\omega_{\Xi_l} s}{2}||b_l - \hat{b}_{\Xi_l}||^2\right), \tag{4.19}$$

where hyper-parameters $\omega_{\Xi_l} > 0$ and $\hat{b}_{\Xi_l}$ determine this prior. That is, to make the original conditional distribution $f\left(b_l|\,a_l, d_l\right) = f\left(b_l|\,\Xi_l\right)$ conjugate, its condition has to be extended by precision parameter s,

$$f\left(b_l|\,\Xi_l\right) \overset{\text{conjugate form}}{\rightarrow} f\left(b_l|\,\Xi_l, s\right).$$

The expected range of b_l given in Table 4.1, together with neglecting the lower bound $\underline{b}_{\Xi_l}$, i.e. $\underline{b}_{\Xi_l} = 0$, gives the rough setting of hyper-parameters of distribution (4.19)

$$\hat{b}_{\Xi_l} = \frac{\overline{b}_{\Xi_l}}{2}, \quad \omega_{\Xi_l} = \frac{4}{\hat{s}\overline{b}^2_{l,\Xi_l}}, \tag{4.20}$$

where $\hat{s}$ is a conservative estimate of the measurement precision s, leading to a relatively flat prior *pdf*. The adopted choice corresponds to one standard deviation on both sides of the mean again. It is necessary to note here that this setting is good for the initial basic physical bound $\overline{b}_{\Xi_l}$. However this setting has poor consequences if the upper bound $\overline{b}_{\Xi_l}$ is decreased. It will be demonstrated in the Section 5.1 in Remark(s) 5.1.2 and preventive solution proposed too.

All the considered probabilistic models as the combinations of parametrised models and prior distributions are summarized in Table 4.2.

Remark(s) 4.3.1

Let's notice that the common bounds for the local parameters in the Table 4.1, e.g. those for model order $(\underline{d}, \overline{d})$ or for the maximum gain of the response $(\underline{r}_{\max}, \overline{r}_{\max})$, alternatively common hyper-parameters of prior distribution, e.g. (α_g, β_g) for distribution $f\left(g_l\right)$ (4.16), can be understood as the common collective hyper-parameters. They can be considered in general as further common unknown *parameter that can be "learned" from the set of data similarly to common parameter s. Though, only the parameter s has straight physical ground. On that account the full "correct" Bayesian learning is done only for this common parameter herein, while the other common parameters are fixed. If it would be necessary to tune-up their values, they could be roughly selected by means of maximum likelihood (or maximum a posteriori) point estimates (see [6], discussion on hierarchical models and empirical Bayes).*

4.3.4 Discussion

Limitations on the used model The selection of the suitable probabilistic model for the scintigraphy observation of the lymphatic system response was limited already from the beginning by the need to employ only the limited number of data. On that account only simple classes of

	Parametrised model $f\left(Y_l \middle\vert \Theta_l\right)$	Prior distribution $f\left(\Theta_l\right)$
	Models for raw data $({}^{R}Y_l, {}^{R}y_{l,I})$	
P_1:	$\mathrm{Pn}({}^{R}y_{l,I} \middle\vert g_l)\prod_{i=1}^{n_l}\mathrm{Pn}({}^{R}y_{l,t_{l,i}} \middle\vert g_l b_l A_{\Xi_{l,i}})$	$\mathrm{Un}(b_l \middle\vert a_l, d_l)\mathrm{Un}_{a_l}(a_l \middle\vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(g_l)$
P_2:	$\cdots$	$\mathrm{Un}_{\log(b_l)}(b_l \vert a_l, d_l)\mathrm{Un}_{\log(1-a_l)}(a_l \vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(g_l)$
	Models for normalized data ${}^{N}Y_l$	
N_1:	$\mathrm{N}({}^{N}Y_l \middle\vert b_l A_{\Xi_l}, s\boldsymbol{I}_{n_l})$	$\mathrm{Un}(b_l \middle\vert a_l, d_l)\mathrm{Un}_{a_l}(a_l \middle\vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(s)$
N_2:		$\mathrm{Un}_{\log(b_l)}(b_l \vert a_l, d_l)\mathrm{Un}_{\log(1-a_l)}(a_l \vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(s)$
N_3:	$\cdots$	$\mathrm{N}(b_l \middle\vert a_l, d_l, s)\mathrm{Un}_{a_l}(a_l \middle\vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(s)$
N_4:		$\mathrm{N}(b_l \middle\vert a_l, d_l, s)\mathrm{Un}_{\log(1-a_l)}(a_l \middle\vert d_l)\mathrm{Un}(d_l)\mathrm{Ga}(s)$

Table 4.2: List of considered ROI probabilistic models.

parametrised models with limited number of parameters were considered to be applicable for our purpose. Consequently we have faced a problem how to design a model, which would capture relatively rich signal ensemble of lymphatic system responses with their observation, having yet only a few parameters.

Inner cascade model The choice of proposed inner cascade model was inspired by the simplified modelling of the complex systems with the distributed parameters [9] and compartmental models often used in medicine [14, 52, 67]. In contrast to these compartmental models the individual sections (i.e. compartments) are not bound up with the physical spatial partition here. Therefore the model order d_l is not fixed and is taken as unknown too. Since the models describe the responses on individual ROIs independently we do not consider even any relations between compartments on these ROIs. It would come in useful within the global model on the whole limb.

Ways of probabilistic modelling The models for two versions of data can be understood as different ways of probabilistic modelling. While for the case of normalized data only the output is described in probabilistic way and the "input" is considered deterministic, also the measured injection amount "input" is described probabilistically in model for the raw data.

The proposed models describe the lymphatic system as deterministic one having noisy measurements. The version of model for normalized data (4.11) is often called output error (OE) model while the version for raw data (4.10) errors-in-variables model, see e.g. [81]. Experimental results, see Chapter 7, indicate that description of the real system is sufficient under the given condition of a few measurements. It is however only approximation. We should consider that also the inner process of accumulation of radiotracer is stochastic process in general. There is a lot of factors (primarily unobservable) that can influence the actual amount of radiotracer so it cannot be well described deterministically. The ARX model popular in many fields for its "nice" properties is unsuitable for description of both the measurement and process noises. Its applicability is limited by lacking of freedom in describing the properties of the disturbance term since it only considers uncorrelated process noise. The more general ARMAX model satisfies integration of both process and measurement noises. The additional flexibility of this model is achieved by describing the error term as moving-average of white noise. The advantage of this more general description is however counterbalanced by difficulty of its parameter estimation. In the case of complete measurement, the system should be described by the last named model. The reason for using the different one is again through over-parametrisation related to the amount of data. In order to minimize the influence of the stochastic nature of inner process the whole inspection procedure is standardized.

Possible improvements of introduced model Some improvements of the basic model are possible and during the extensive use they can be found useful. For instance, the model can be enriched by a fixed time delay bridging the guaranteed lower bound on real time delays. This bound is several minutes for axilla and the need to cover it by the current model may harm its quality. This time delay can be in general defined also as the unknown parameter but it will increase the number of unknown parameters, that complicates estimation from few data. In this occasion the lower bounds on real time delays are not determined yet. Therefore the fixed delays are not incorporated into the models. We hope that omitting the delay has negligible influence on the final decision making that cares about slow dynamics primarily.

Also some alternatives of observation part of the model can be found appropriate. For instance, the log-normal version of model could be used on the normalized data, using the log-normal distribution instead of normal distribution. It would solve the problem of normal model permitting the negative relative activities. This version of model seems promising, but it does not appear here. It is one of possible items of future work.

4.4 Summary

All this Chapter was devoted to the probabilistic modelling of the solved scintigraphy problem. Its main outcomes are:

- general structure of the global observation probabilistic model necessary for the task solved on the set of patients' scintigraphic data;
- particular models for the description of the scintigraphy response of the limb lymphatic system on individual ROIs. They consist of:
  - common inner cascade parametrised model describing the relative response of lymphatic system during scintigraphic inspection defined only by triple of parameters
  - two versions of observation parametrised models for two version of data considered, that are build up on the common inner part
  - relatively rich prior information about the ranges of the model parameters transformed to various tested prior distributions

  All the models are treated as candidates for practical application.

Chapter 5

Parametric and Predictive Inference

The Chapter 4 was devoted to the specification of outer observation probabilistic model by means of couple parametrised model and prior distribution about its parameters. This Chapter gets on and expresses the parametric and predictive inference for all the considered models that are the basic building elements for the addressed tasks. The Section 5.1 concerns the local inference on individual limbs while the Section 5.2 deals with the global inference on the group of limbs.

5.1 Limb Related Inference

The main aim of the common routine inspection is the evaluation of the state of individual limbs. It is done for each limb separately except for the special tasks where both limbs of patient are directly compared. Therefore the local limb inference plays here the key role. Besides, the local limb inference can be employed properly in the global inference defined on the group of patients.

Having parametrised model $f(Y_l|\,\Theta_l)$ together with the prior distribution $f(\Theta_l)$ and data Y_l for the ROI of one limb both limb related predictive and parametric inference are

$$\begin{aligned} f(\Theta_l|\,Y_l) &= \frac{f(Y_l|\,\Theta_l)\,f(\Theta_l)}{f(Y_l)} \propto f(Y_l|\,\Theta_l)\,f(\Theta_l) \\ f(Y_l) &= \int_{\Theta_l^*} f(Y_l|\,\Theta_l)\,f(\Theta_l)\,d\Theta_l. \end{aligned}$$

They are computed below for all the models summarized in the Table 4.2. Except these formulas other useful marginal and conditional posterior distributions of model parameters are expressed too, hereafter we will refer to them repeatedly. Employed notation from the global list at the beginning of thesis is repeated:

$\Gamma(\cdot)$	Euler gamma function
$\mathrm{Ga}(\cdot\,|\,\alpha,\beta)$	*pdf* of gamma distribution with parameters (α,β)
$\mathrm{N}(\cdot\,|\,\mu,\lambda)$	*pdf* of Gaussian normal distribution with parameters (μ,λ)
$\mathrm{St}(\cdot\,|\,\mu,\lambda,\alpha)$	*pdf* of Student (t) distribution with parameters (μ,λ,α)
$\|\|\cdot\|\|$	Euclidean norm
$\cdot!$	factorial

Further information about the probability distributions can be found in the Section 3.2.7.

Inference for Poisson models $\mathbf{P_1}$ and $\mathbf{P_2}$ on raw data $Y_l = ({}^R Y_l, {}^R y_{l,I})$

$$
\begin{aligned}
f(\Xi_l, b_l | Y_l) &\propto w_{\Xi_l,b_l} = \chi_2^{-\chi_1} \prod_{i=1}^{n_l} (b_l A_{\Xi_l,i})^{{}^R y_{l,t_{l,i}}} && (5.1)\\
f(g_l | \Xi_l, b_l, Y_l) &= \mathrm{Ga}(g_l | \chi_1, \chi_2) = \frac{\chi_2^{\chi_1} g_l^{\chi_1 - 1} \exp(-g_l \chi_2)}{\Gamma(\chi_1)} && (5.2)\\
f(\Xi_l, b_l, g_l | Y_l) &\propto w_{\Xi_l,b_l} f(g_l | \Xi_l, b_l, Y_l) && (5.3)\\
f(Y_l) &= \frac{\Gamma(\chi_1)\beta_g^{\alpha_g} f(\Xi_l, b_l)}{\Gamma(\alpha_g)\, {}^R y_{l,I}! \prod_{i=1}^{n_l} {}^R y_{l,t_{l,i}}!} \sum_{\Xi_l^*, b_l^*} w_{\Xi_l,b_l}, && (5.4)
\end{aligned}
$$

where

$$
\begin{aligned}
\chi_1 &= {}^R y_{l,I} + \sum_{i=1}^{n_l} {}^R y_{l,t_{l,i}} + \alpha_g \\
\chi_2 &= 1 + \beta_g + \sum_{i=1}^{n_l} b_l A_{\Xi_{l,i}}
\end{aligned}
$$

Inference for normal models $\mathbf{N_1}$ and $\mathbf{N_2}$ on normalized data $Y_l = {}^N Y_l$

$$
\begin{aligned}
f(\Xi_l, b_l | Y_l) &\propto w_{\Xi_l,b_l} = \chi_2^{-\chi_1} && (5.5)\\
f(s | \Xi_l, b_l, Y_l) &= \mathrm{Ga}(s | \chi_1, \chi_2) = \frac{\chi_2^{\chi_1} s^{\chi_1 - 1} \exp(-s\chi_2)}{\Gamma(\chi_1)} && (5.6)\\
f(\Xi_l, b_l, s | Y_l) &\propto w_{\Xi_l,b_l} f(s | \Xi_l, b_l, Y_l) && (5.7)\\
f(s | Y_l) &\propto \sum_{\Xi_l^*, b_l^*} w_{\Xi_l,b_l} \mathrm{Ga}(s | \chi_1, \chi_2) && (5.8)\\
f(Y_l) &= \frac{\Gamma(\chi_1)\beta_s^{\alpha_s} f(\Xi_l, b_l)}{\Gamma(\alpha_s)(2\pi)^{\frac{n_l}{2}}} \sum_{\Xi_l^*, b_l^*} w_{\Xi_l,b_l}, && (5.9)
\end{aligned}
$$

where

$$
\begin{aligned}
\chi_1 &= \frac{n_l + 2\alpha_s}{2} \\
\chi_2 &= \beta_s + \frac{1}{2} || {}^N Y_l - b_l A_{\Xi_l} ||^2
\end{aligned}
$$

Inference for normal models $\mathbf{N_3}$ and $\mathbf{N_4}$ on normalized data $Y_l = {}^N Y_l$

$$
\begin{aligned}
f(\Xi_l | Y_l) &\propto w_{\Xi_l} = \chi_2^{-1/2} \chi_4^{-\chi_1} && (5.10)\\
f(s | \Xi_l, Y_l) &= \mathrm{Ga}\left(s | \chi_1, \frac{\chi_4}{2}\right) = \frac{\chi_4^{\chi_1} s^{\chi_1 - 1} \exp(\frac{-s\chi_4}{2})}{2^{\chi_1}\Gamma(\chi_1)} && (5.11)\\
f(b_l | \Xi_l, s, Y_l) &= \mathrm{N}\left(b_l | \frac{\chi_3}{\chi_2}, s\omega_{\Xi_l}\chi_2\right) = \left(\frac{\omega_{\Xi_l} s\chi_2}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\omega_{\Xi_l} s\chi_2}{2} || b - \frac{\chi_3}{\chi_2} ||^2\right) && (5.12)\\
f(\Xi_l, b_l, s | Y_l) &\propto w_{\Xi_l} f(b_l | \Xi_l, s, Y_l) f(s | \Xi_l, Y_l) && (5.13)\\
f(b_l | \Xi_l, Y_l) &= \mathrm{St}\left(b_l | \frac{\chi_3}{\chi_2}, \frac{2\omega_{\Xi_l}\chi_1\chi_2}{\chi_4}, 2\chi_1\right) \propto \left[1 + \frac{\omega_{\Xi_l}\chi_2}{\chi_4}(b_l - \frac{\chi_3}{\chi_2})\right]^{-\frac{2\chi_1+1}{2}} && (5.14)
\end{aligned}
$$

$$f\left(s|\,Y_l\right) \propto \sum_{\Xi_l^*} w_{\Xi_l} \mathrm{Ga}\left(s|\,\chi_1, \frac{\chi_4}{2}\right) \tag{5.15}$$

$$f\left(Y_l\right) = \frac{\Gamma(\chi_1)(2\beta_s)^{\alpha_s} f\left(\Xi_l\right)}{\Gamma(\alpha_s)\pi^{\frac{n_l}{2}}} \sum_{\Xi_l^*} w'_{\Xi_l} \tag{5.16}$$

where

$$\begin{aligned}
\chi_1 &= \frac{n_l + 2\alpha_s}{2} \\
\chi_2 &= 1 + \omega_{\Xi_l}^{-1} ||A_{\Xi_l}||^2 \\
\chi_3 &= \hat{b}_{\Xi_l} + \omega_{\Xi_l}^{-1} {}^N Y_l^T A_{\Xi_l} \\
\chi_4 &= ||{}^N Y_l||^2 + 2\beta_s + \omega_{\Xi_l} \left(\hat{b}_{\Xi_l}^2 - \frac{\chi_3^2}{\chi_2}\right)
\end{aligned}$$

The common formulas can be always given for the pairs of considered models because they only differ in the discretization of the same parameter that has no influence on these formulas. It is only necessary to keep in mind this difference in the discretization. The non-normalized weights $w_\bullet$ force us to use the proportion sign $\propto$ in equations for the posterior probabilities. In order to be able to substitute it by equality these weights have to be replaced by their normalized versions, $\ddot{w}_\bullet = w_\bullet / \sum_{i=1}^{n_l} w_\bullet$.

Remark(s) 5.1.1

(i) The conditioning and marginalization allows to express the joint posterior distributions of parameters for individual models as the product of posterior pf *of all discrete/discretized parameters and posterior* pdf*s of the continuous parameters. The joint distribution (5.3) for models $P_\bullet$ is a product of (5.1) and (5.2), for example.*

(ii) Due to the discrete/discretized parameters used, the predictive and some marginal parametric inferences are computed by the summation over these parameters, see e.g. predictive inferences (5.4), (5.9) and (5.16) for all the considered models.

(iii) The equations (5.2) for models $P_\bullet$, (5.6) for N_1 and N_2 and (5.11), (5.12) for N_3 and N_4 demonstrate that the corresponding prior pdf*s (4.16), (4.17) and (4.19) are conjugate.*

Remark(s) 5.1.2

Here is the proper point to demonstrate the adverse influence od decrease of upper bound $\overline{b}_{\Xi_l}$ together with the setting (4.20) of the Gaussian prior distribution $f\left(b_l|\,\Xi_l, s\right)$ for models N_3, N_4, forewarn in the Section 4.3.3. If we employ (4.20) then for smaller upper bound $\overline{b}_{\Xi_l}$ the distribution $f\left(b_l|\,\Xi_l, s\right)$ narrows. Consequently, the data-based information is suppressed overly. It is clear when evaluating the expectation of the conditional posterior distributions $f\left(b_l|\,\Xi_l, Y_l\right)$, (5.14), of the parameter b_l

$$E\left[b_l|\,\Xi_l, Y\right] = \frac{\hat{b}_{\Xi_l} + \omega_{\Xi_l}^{-1} {}^N Y_l^T A_{\Xi_l}}{1 + \omega_{\Xi_l}^{-1} ||A_{\Xi_l}||^2}.$$

If the upper gain is decreased to $1/r$ of the original upper gain, i.e. ${}^n\overline{b}_{\Xi_l} = \overline{b}_{\Xi_l}/r$, then the new precision parameter ${}^n\omega_{\Xi_l}$ set according (4.20) is ${}^n\omega_{\Xi_l} = r^2\omega_{\Xi_l}$. Consequently, the influence of the data to the posterior mean is suppressed by the square of r. With increasing r the parameter

of prior expectation dominates in the posterior expectation. The proposed solution how to prevent it is to decrease only the mean $\hat{b}_{\Xi_l}$ of the normal distribution $f(b_l|\,\Xi_l, s)$ (4.19) with decreasing upper bound $\overline{b}_{\Xi_l}$ while the precision parameter ω_{Ξ_l} to set according the original physical bound. Then prior distribution will be sufficiently flat on the a priori possible interval. Though also the values outside the interval will have non-negligible probability. The information loss about the gain parameter does not seem critical for the estimation.

5.2 Global Inference

While the Section 5.1 deals with the local inference on individual limbs necessary for their evaluation this Section is devoted to the global inference defined on the group of patients' limbs. Such inference comes in useful for the special tasks solved over the whole "experimental" group of limbs but it is also necessary for all the task done on more than one limb.

This Section is divided into parts according to two considered structures of global parameter Θ described in the Section 4.2.1. The Section 5.2.1 deals with the simpler case of global parametric and predictive inference for the model with strictly limb local parameters. The Section 5.2.2 concerns model with limb local parameters containing a common part for all limbs. The general notation from the Section 4.2 is preserved here. The membership of the considered models from the Table 4.2 into these two cases is repeated only. The concrete forms can be obtained then by the substitution of terms from Section 5.1 to the general forms. The approximations of the global estimate of the common parameter s and of the global predictive and parametric inference for models $\mathrm{N}_\bullet$ whose exact computation is difficultly realizable are proposed at the end of the Section 5.2.2.

5.2.1 Inference for Models with Strictly Limb Specific Parameters

In the list of the considered models only the Poisson models $\mathrm{P}_\bullet$ for the raw data have strictly limb local parameters. The total separation (4.6) of limb parameters Θ_l in the global parameter $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_{\mathring{l}})$, see Section, makes both global parametric and predictive inference clear:

$$f\left(\Theta|\, Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) = \prod_{l=1}^{\mathring{l}} f\left(\Theta_l|\, Y_l\right) \tag{5.17}$$

$$f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) = \prod_{l=1}^{\mathring{l}} f\left(Y_l\right). \tag{5.18}$$

Terms (5.3) and (5.4) can be simply used for their evaluation.

5.2.2 Inference for Models with Limb Local Parameters Containing Common Part ${}^{[c]}\Theta$

The considered normal models $\mathrm{N}_\bullet$ for the normalized data contain the measurement noise precision parameter s which is defined to be common for all limbs, i.e. ${}^{[c]}\Theta = s$ in the global parameter $\Theta = ({}^{[c]}\Theta, {}^{[i]}\Theta_1, {}^{[i]}\Theta_2, \ldots, {}^{[i]}\Theta_{\mathring{l}})$. This common part together with the decomposition (4.7) causes the global parametric and predictive inference to be more complicated:

$$f\left(\Theta|\, Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) \propto \left[\prod_{l=l_p+1}^{\mathring{l}} f\left(Y_l|\, {}^{[c]}\Theta, {}^{[i]}\Theta_l\right) f\left({}^{[i]}\Theta_l|\, {}^{[c]}\Theta\right)\right] \times \ldots$$

$$\ldots \times f\left({}^{[c]}\Theta, {}^{[i]}\Theta_1, {}^{[i]}\Theta_2, \ldots, {}^{[i]}\Theta_{l_p} \middle|\, Y_1, Y_2, \ldots, Y_{l_p}\right) \tag{5.19}$$

$$f\left({}^{[c]}\Theta \middle|\, Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) \propto \left[\prod_{l=1}^{\mathring{l}} f\left(Y_l \middle|\, {}^{[c]}\Theta\right)\right] f\left({}^{[c]}\Theta\right) \propto$$

$$\propto \left[\prod_{l=l_p+1}^{\mathring{l}} f\left(Y_l \middle|\, {}^{[c]}\Theta\right)\right] f\left({}^{[c]}\Theta \middle|\, Y_1, Y_2, \ldots, Y_{l_p}\right) \tag{5.20}$$

$$f\left(Y_1, Y_2, \ldots, Y_{\mathring{l}}\right) = \int_{{}^{[c]}\Theta^*} \left[\prod_{l=1}^{\mathring{l}} f\left(Y_l \middle|\, {}^{[c]}\Theta\right)\right] f\left({}^{[c]}\Theta\right) d^{[c]}\Theta =$$

$$= \int_{{}^{[c]}\Theta^*} \left[\prod_{l=l_p+1}^{\mathring{l}} f\left(Y_l \middle|\, {}^{[c]}\Theta\right)\right] f\left({}^{[c]}\Theta \middle|\, Y_1, Y_2, \ldots, Y_{l_p}\right) d^{[c]}\Theta \times \ldots$$

$$\ldots \times f\left(Y_1, Y_2, \ldots, Y_{l_p}\right) \tag{5.21}$$

where

$$f\left(Y_l \middle|\, {}^{[c]}\Theta\right) = \int_{{}^{[i]}\Theta_l^*} f\left(Y_l \middle|\, {}^{[c]}\Theta, {}^{[i]}\Theta_l\right) f\left({}^{[i]}\Theta_l \middle|\, {}^{[c]}\Theta\right) d^{[i]}\Theta_l.$$

These general forms are theoretically valid however their precise computation causes technical difficulties. Consequently some approximation is necessary to be established. The description of the proposed solution for all the inferences (5.19), (5.20), (5.21) follows.

Estimate of the common parameter ${}^{[c]}\Theta = s$

The posterior estimate of the common parameter s is one of the important results. Its prior range and corresponding *pdf* described in the Section 4.3.3 has been chosen roughly considering it can be refined by processing of sets of ROI data. Inference (5.20) is the right instrument for this refinement.

The prior *pdf* for the noise precision parameter s has been chosen to be conjugate gamma distribution (4.17) in the model specification, Section 4.3.3. Recursive estimate of l-th limb needs as its "temporary prior" distribution estimate done on the first $(l-1)$ limbs, see (5.20). Thus, the estimate done on one limb for the proposed normal models $\mathrm{N}_\bullet$ results in a sum of $\mathring{n}$ weighted gamma distributions, see (5.8) for models N_1, N_2 and (5.15) for models N_3, N_4, where $\mathring{n}$ is the number of "grid points" of the discrete or discretized parameters. Proceeding with the computation on the second limb the estimate consists of $\mathring{n}^2$ weighted gamma distributions. In summary, the resulting estimate on $\mathring{l}$ limbs consists of $\mathring{n}^{\mathring{l}}$ weighted gamma distributions so their number increases in the exponential manner. Though it is not possible to realize its computation.

The proposed formal solution how to overcome this problem is to approximate the posterior sum of $\mathring{n}$ weighted gamma distributions in each step of recursion through the limbs l by a single gamma distribution:

$$f\left(s \middle|\, Y_1, Y_2, \ldots, Y_l\right) = \sum_{n=1}^{\mathring{n}} w_n \mathrm{Ga}(s|\, \alpha_n, \beta_n) \overset{\text{approx.}}{\rightarrow} \mathrm{Ga}(s|\, \hat{\alpha}, \hat{\beta}) = \hat{f}\left(s \middle|\, Y_1, Y_2, \ldots, Y_l\right), \tag{5.22}$$

and employ it as "temporary prior" distribution in next step of the recursion. Let's stress here, that $\sum_{n=1}^{\mathring{n}} w_n = 1$. The poor property of this recursive approximation is, that its result depends on the order in which the limbs are processed. The possibility how to suppress this influence

is to make repetitive evaluations for different random permutations of limbs. Thereafter these partial results can be summarized into a single gamma distribution in a similar way as for approximation in each step of recursion. It allows us to hold the number of numerical evaluations and computation time reasonable.

Two alternatives of approximation (5.22) are proposed. Other approaches to this approximation could be applied, e.g. employing mixtures or replacement of arithmetic mean by geometric mean [**?**, **?**], but the proposed solutions have very simple interpretation.

Approximation (5.22) via the equality of the first and second moment of *pdf*s This approximation is based on a preservation of moments of the original *pdf* in its approximation. Since the gamma distribution is defined by two parameters, also only two important moments — mean and variance — were selected to be preserved. It means:

$$E_f\left[s\right] = E_{\hat{f}}\left[s\right], \quad V_f\left[s\right] = V_{\hat{f}}\left[s\right]. \tag{5.23}$$

Means and variances (5.23) of the original and approximative *pdf*s expressed by means of their parameters (see Section 3.2.7 for information on moments of this distribution) are

$$\begin{aligned} E_f\left[s\right] &= \textstyle\sum_{n=1}^{\mathring{n}} w_n \frac{\alpha_n}{\beta_n}, & E_{\hat{f}}\left[s\right] &= \frac{\hat{\alpha}}{\hat{\beta}} \\ V_f\left[s\right] &= \textstyle\sum_{n=1}^{\mathring{n}} w_n \frac{\alpha_n^2+\alpha_n}{\beta_n^2} - E_f\left[s\right]^2, & V_{\hat{f}}\left[s\right] &= \frac{\hat{\alpha}}{\hat{\beta}^2}. \end{aligned}$$

The parameters $(\hat{\alpha}, \hat{\beta})$ of single gamma distribution approximation can be then determined as

$$\hat{\alpha} = \frac{E_f\left[s\right]^2}{V_f\left[s\right]}, \quad \hat{\beta} = \frac{E_f\left[s\right]}{V_f\left[s\right]}.$$

Minimal discrepancy approximation The second type of approximation is based on the minimization of the discrepancy between the original distribution and its approximation, see Remark(s) 3.2.3(i). Such approximation is searched which minimizes the expected loss of the local logarithmic loss function of reporting approximative *pdf* rather than the original *pdf*, see Section 3.2.5. The optimal single gamma approximation is then determined by

$$\min_{\hat{\alpha},\hat{\beta}} \mathcal{D}\left(f\left(s\right) || \, \hat{f}\left(s\right)\right) = \min_{\hat{\alpha},\hat{\beta}} \int_{s=0}^{\infty} \sum_{n=1}^{\mathring{n}} w_n \mathrm{Ga}(s|\, \alpha_n, \beta_n) \log \frac{\sum_{n=1}^{\mathring{n}} w_n \mathrm{Ga}(s|\, \alpha_n, \beta_n)}{\mathrm{Ga}(s|\, \hat{\alpha}, \hat{\beta})} ds. \tag{5.24}$$

The numerator in the logarithm in (5.24) does not depend on the approximative *pdf*, therefore it has no influence on the result of the minimization and the solution can be found as the maximum over $(\hat{\alpha}, \hat{\beta})$ from

$$\begin{aligned} \textstyle\int_{s=0}^{\infty} \sum_{n=1}^{\mathring{n}} w_n \mathrm{Ga}(s|\, \alpha_n, \beta_n) \log \mathrm{Ga}(s|\, \hat{\alpha}, \hat{\beta}) ds \;&=\; \hat{\alpha} \log \hat{\beta} - \log \Gamma(\hat{\alpha}) + \ldots \\ &\quad + \textstyle\sum_{n=1}^{\mathring{n}} w_n \Big[(\hat{\alpha}-1)(\psi(\alpha_n) - \log \beta_n) - \hat{\beta} \frac{\alpha_n}{\beta_n} \Big], \end{aligned}$$

where $\psi(\cdot) = \Gamma(\cdot)'/\Gamma(\cdot)$ denotes the *digamma function* (see e.g. [**?**]).

The extreme can be found by searching the zero gradient of minimized term in (5.24), i.e. differentiating the term with respect to $\hat{\alpha}$ and $\hat{\beta}$ and setting them 0:

$$\frac{\partial \mathcal{D}\left(f\left(s\right) || \, \hat{f}\left(s\right)\right)}{\partial \hat{\alpha}} = \log(\hat{\beta}) - \psi(\hat{\alpha}) + \sum_{n=1}^{\mathring{n}} w_n [\psi(\alpha_n) - \log(\beta_n)] = 0 \tag{5.25}$$

$$\frac{\partial \mathcal{D}\left(f\left(s\right) || \, \hat{f}\left(s\right)\right)}{\partial \hat{\beta}} = \frac{\hat{\alpha}}{\hat{\beta}} - \sum_{n=1}^{\mathring{n}} w_n \frac{\alpha_n}{\beta_n} = 0 \tag{5.26}$$

By a substitution of $\hat{\beta}$ from the equation (5.26) to the equation (5.25) and knowing α_n is the same for each n (exactly $\alpha_n = n_l/2+\alpha_s$ for all models $\mathrm{N}_\bullet$, see Section 5.1), the following equation of single unknown variable $\hat{\alpha}$ have to be solved:

$$\log(\hat{\alpha}) - \psi(\hat{\alpha}) = \log(\alpha_n) - \psi(\alpha_n) + \sum_{n=1}^{\mathring{n}} w_n \log(\beta_n) + \log(\sum_{n=1}^{\mathring{n}} w_n/\beta_n). \tag{5.27}$$

The solution of (5.27) can not be found analytically, so the numerical techniques are employed. The function $\log(\hat{\alpha}) - \psi(\hat{\alpha})$ is purely monotonous so the a single solution can easily be found.

Remark(s) 5.2.1

(i) Let's notice one interesting thing clearly visible in (5.26). The mean value of the original distribution, $E_f[s] = \sum_{n=1}^{\mathring{n}} w_n \frac{\alpha_n}{\beta_n}$, is also preserved in its approximation via the minimal discrepancy $E_{\hat{f}}[s] = \frac{\hat{\alpha}}{\hat{\beta}}$, alike in the first proposed approximation via moments.

(ii) The proposed approximations differ. They would coincide only when the normal minimal discrepancy approximation would be searched for. Recall that the best normal approximation via discrepancy to pdf *is the normal distribution having the same mean and variance [6].*

Global prediction inference and joint parametric inference of limb local parameters

The estimate (5.20) of the common parameter s is one of the global inferences, whose results are employed for local inference necessary for individual limbs evaluation. However, solving various global decision tasks over the group of limbs further global inferences are in demand. Specifically, the global prediction inference (5.21) and joint parametric inference of strictly limb local parameters are required for decision tasks in the Chapter 6.

Similarly to the estimate of common parameter s, the precise computation of these global inferences causes technical difficulties. In general, we could employ approximations proposed for estimation of common parameter s. However, the results would be dependent on the order of processing of limbs again. The permuting of limbs order to suppress its influence is not practically realizable on the larger group of limbs here. For that reason, the simplification is in demand.

The proposed simplification is based on the relaxation of the assumption about the parameter s to be common for all limbs. The noise parameter s is considered here to be local for individual limbs, i.e. common s breaks into local ones, $s \to (s_1, s_2, \ldots, s_{\mathring{l}})$. The estimate of the common s is then employed as the prior *pdf* for individual s_l. Thereafter the predictive inference transforms to the form (5.18) and joint parametric inference of strictly limb local parameters to the form structurally same form as (5.17) (only the strictly limb local parameters are included in Θ_l here), that can be easily evaluated.

The proposed model simplification is not the the consequence of reassessment of original model but only a solution how to overcome computational difficulties. It is clear that employing this simplification we introduce some error into results. Though, we hope this error has negligible influence on final decisions. Experiments, Chapter 7 are encouraging in this respect.

5.3 Summary

In this Chapter all the local and global predictive and parametric inferences for all the considered models from Table 4.2 are presented. They form basic results of Bayesian inference, however, their main aim is to serve for further purposes, the Chapter 6 is devoted to.

Among all the results, the global estimation of the common parameter s should be stressed. It is employed for refinement of roughly chosen prior distribution. The proposed approximations are the case specific solutions of ever present computational complexity.

Chapter 6

Inference of Significant Quantities and Decision Tasks

The parametric and predictive inference described in the Chapter 5 are the basic results originated from modelling, but they are not the final aim of the work. They are employed for the solution of partial tasks that have arisen to be important or useful for the specific case of lymphoscintigraphy described in this Chapter. The beliefs are to be used directly as the basic elements in the solved decision tasks or are to be reported with the intention to solve incompletely defined decision tasks in the future.

The Section 6.1 concerns the estimate of time activity curve which forms the basis of various scintigraphy quantification techniques while Section 6.3 is devoted to the inference of quantitative parameters possibly useful for the assessment of the lymphedema. The Section 6.2 describes how the quality of the proposed models can be verified comparing the original data with trends in reconstructed output. The Section 6.4 deals more deeply with the problem of selection of sampling times which is important to make the evaluation practically realizable and maximally effective. At the end of the Chapter the approach how to compare quantitatively the patient's limbs is proposed in the Section 6.5.

6.1 Time Activity Curve Estimate

Construction of this curve has been one of the the principal motivations for the employed modelling. Estimate of the whole curve is needed as it serves both to visual and quantitative evaluations, see Section 1.2. The practical lack of measurements does not allow us to use or even modify these techniques utilizing data at disposal alone. Therefore time activity curve as a means for possible further evaluation is reconstructed from few measurements, only.

It follows from (4.9) that X_l is a deterministic function of the unknown parameters, (Ξ_l, b_l). The distribution on the ensemble of possible X_l, $f(X_l|\, Y_l)$, is therefore a highly non-linear function of parameters (Ξ_l, b_l) described by the posterior *pdf* $f(\Xi_l, b_l|\, Y_l)$. From a clinical perspective, the point estimate of the time activity curve is representative enough. The expected value, $E[X_l|\, Y_l]$, — the general result of the decision task searching point estimate with respect to quadratic loss function, see Section 3.2.5, Remark(s) 3.2.1 and for more details [6, 47] — can be calculated directly for all the considered models.

The expected value $E[X_l|\, Y_l]$ for the models $\mathrm{P}_\bullet$ and N_1, N_2, see Table 4.2, can be computed

similarly from (4.9) and (5.1) or (5.5) according to the considered model:

$$E\left[X_l|\,Y_l\right] \propto \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l} X_l(\Xi_l,b_l) = \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l} b_l \widetilde{A}_{\Xi_l}. \tag{6.1}$$

The calculation for remaining models N_3, N_4 from the Table 4.2 is little modified due to a possibility of analytical integration of b_l. Then employing (4.9), (5.10) and (5.14) it follows:

$$E\left[X_l|\,Y_l\right] \propto \sum_{\Xi_l^*} w_{\Xi_l} \widetilde{A}_{\Xi_l} \int_{b_l^*} b_l f\left(b_l|\,\Xi_l,Y_l\right) db_l = \sum_{\Xi_l^*} w_{\Xi_l} \widetilde{A}_{\Xi_l} E\left[b_l|\,\Xi_l,Y_l\right], \tag{6.2}$$

where an expectation $E\left[b_l|\,\Xi_l,Y_l\right]$ of the Student distributions $f\left(b_l|\,\Xi_l,Y_l\right)$ (5.14) is χ_3/χ_2. Hence, the numerical evaluation of (6.1) and (6.2) reduces to a simple summation over the grid of *a priori*-allowed values of (Ξ_l,b_l) in the first case and of Ξ_l in the second case.

Remark(s) 6.1.1

(i) Instead of the expected value, other point estimates like quantiles and medians [6] could be computed, but it would not be so easy and would require more intensive computations. Therefore the former solution has been selected as a single curve estimate report.

(ii) The Bayesian theory works with uncertainty, on that account some more information about the probability distribution of estimated time activity curve $f\left(X_l|\,Y_l\right)$ should be given. Unfortunately right credible regions, e.g. highest probability density regions [6], need too much computation. Therefore the probability distribution is summarized here by its moments. The covariance matrix (with marginal variance on the main diagonal) and higher moments can be computed similarly as the expected value.

6.2 Output Reconstruction at Non-Sampled Times

In the previous Section the estimate of the time activity curve is given. It is tightly connected with the observed data. It is important quantifier for our purpose but it can not say anything about the quality of the model how it fits the employed data. The right choice for such task is a prediction (reconstruction) of the output variable at the non-sampled times, i.e. times where the observation was not done $\widetilde{\mathcal{T}}_l = (0,1,2,\ldots,t,\ldots) \setminus \mathcal{T}_l$. The quality of the model can be then verified by comparing how the original data fits into the trend of predicted output variable at non-sampled times. For that purpose predictive *pf* $f\left({}^{R}\widetilde{Y}_l|\,Y_l\right)$ for the raw data, and *pdf* $f\left({}^{N}\widetilde{Y}_l|\,Y_l\right)$ for the normalized data need to be evaluated, where ${}^{\bullet}\widetilde{Y}_l = ({}^{\bullet}y_{l;\tilde{t}_{l;1}}, {}^{\bullet}y_{l;\tilde{t}_{l;2}},\ldots),\ \tilde{t}_{l;\cdot} \in \widetilde{\mathcal{T}}_l$.

The joint distribution of the whole vector ${}^{\bullet}\widetilde{Y}_l$, $f\left({}^{\bullet}\widetilde{Y}_l|\,Y_l\right)$ is of the general interest, though for comparing the data with the trend of prediction the marginal distributions of ${}^{\bullet}y_{l;t}$ at individual times t are sufficient. For the Poisson models $\mathrm{P}_{\bullet}$ in the Table 4.2 the marginal *pf* $f\left({}^{R}y_{l;t}|\,Y_l\right)$ follows from (4.10) and (5.3):

$$\begin{aligned} f\left({}^{R}y_{l;t}|\,Y_l\right) &= \sum_{\Xi_l^*,b_l^*} \int_{g_l^*} f\left({}^{R}y_{l;t}|\,\Xi_l,b_l,g_l\right) f\left(\Xi_l,b_l,g_l|\,Y_l\right) dg_l \\ &\propto \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l} \int_{g_l^*} \mathrm{Pn}({}^{R}y_{l;t}|\,g_l b_l \widetilde{A}_{\Xi_l,t+1}) \mathrm{Ga}(g_l|\,\chi_1,\chi_2) dg_l \\ &\propto \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l} \mathrm{Pg}({}^{R}y_{l;t}|\,\chi_1,\chi_2,b_l\widetilde{A}_{\Xi_l,t+1}), \quad t \in \widetilde{\mathcal{T}}_l, \end{aligned} \tag{6.3}$$

i.e. it is the sum of weighted Poisson-Gamma (Pg) distributions (3.16). The various moments of the distribution (6.3) can be then computed similarly to time activity curve estimate in the Section 6.1. For instance, its expected value

$$E\left[{}^{R}y_{l;t}|\,Y_l\right] \propto \sum_{\Xi_l^*,b_l^*} w_{\Xi,b_l} b_l \widetilde{A}_{\Xi_l,t+1} E\left[g_l|\,\Xi_l,b_l,Y_l\right]$$

is the weighted sum of relative time activity curve multiplied by the posterior mean value of the absolute gain parameter g_l.

The marginal distribution $f\left({}^{N}y_{l;t}|\,Y_l\right)$ for the first pair of models on normalized data N_1, N_2, see Table 4.2 follows from (4.11) and (5.7):

$$\begin{aligned} f\left({}^{N}y_{l;t}|\,Y_l\right) &= \sum_{\Xi_l^*,b_l^*}\int_{s^*} f\left({}^{N}y_{l;t}|\,\Xi_l,b_l,s\right) f\left(\Xi_l,b_l,s|\,Y_l\right) ds \\ &\propto \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l}\int_{s^*} \mathrm{N}({}^{N}y_{l;t}|\,b_l\widetilde{A}_{\Xi_l,t+1},s)\mathrm{Ga}(s|\,\chi_1,\chi_2)ds \\ &\propto \sum_{\Xi_l^*,b_l^*} w_{\Xi_l,b_l}\mathrm{St}\left({}^{N}y_{l;t}|\,b_l\widetilde{A}_{\Xi_l,t+1},\frac{\chi_1}{\chi_2},2\chi_1\right),\quad t\in\widetilde{\mathcal{T}}_l \end{aligned} \tag{6.4}$$

and for the second pair of normal models N_3, N_4 in the Table 4.2, employing (4.11) and (5.13) the similar form is obtained:

$$\begin{aligned} f\left({}^{N}y_{l;t}|\,Y_l\right) &= \sum_{\Xi_l^*}\int_{s^*,b_l^*} f\left({}^{N}y_{l;t}|\,\Xi_l,b_l,s\right) f\left(\Xi_l,b_l,s|\,Y_l\right) dsdb_l \\ &\propto \sum_{\vartheta_l^*} w_{\Xi_l}\int_{s^*,b_l^*} \mathrm{N}({}^{N}y_{l;t}|\,b_l\widetilde{A}_{\Xi_l,t+1},s)\mathrm{N}\left(b_l|\,\frac{\chi_3}{\chi_2},s\omega_{\vartheta_l}\chi_2\right)\mathrm{Ga}\left(s|\,\chi_1,\frac{\chi_4}{2}\right)dsdb_l \\ &\propto \sum_{\Xi_l^*} w_{\Xi_l}\mathrm{St}\left({}^{N}y_{l;t}|\,\widetilde{A}_{\Xi_l,t+1}\frac{\chi_3}{\chi_2},\frac{2\chi_1}{\chi_4}[1+(\omega_{\Xi_l}\chi_2)^{-1}\widetilde{A}^2_{\Xi_l,t+1}]^{-1},2\chi_1\right), t\in\widetilde{\mathcal{T}}_l. \end{aligned} \tag{6.5}$$

The marginal distribution in both cases (6.4), (6.5) is formed by the weighted sum of Student (St) distributions(3.16).

Remark(s) 6.2.1

(i) In comparison to the estimated time activity curves in the previous Section, the mean values of the marginal predictions are the same but the variances are greater. It is implied by the incorporating the noise of measurement into the prediction.

Let's notice e.g. that for the models N_3 and N_4 the conditional posterior variance of one point of the time activity curve x_t given Ξ_l is $V\left[x_t|\,\Xi_l,Y_l\right] = \widetilde{A}^2_{\Xi_l,t+1}\chi_4/(2\omega_{\Xi_l}(\chi_1-1)\chi_2)$. By contrast the variance of the prediction of ${}^{N}y_{l;t}$ (6.5) given Ξ_l:

$$V\left[{}^{N}y_{l;t}|\,\Xi_l,Y_l\right] = V\left[x_t|\,\Xi_l,Y_l\right] + \frac{\chi_4}{2(\chi_1-1)} = V\left[x_t|\,\Xi_l,Y_l\right] + E\left[s^{-1}|\,\Xi_l,Y_l\right]$$

is greater by the additive term constant for all times t, that is not anything else than the mean value of the reciprocal value of the parameter s, $E\left[s^{-1}|\,\Xi_l,Y_l\right]$, i.e. the posterior expectation of the variance of measurement noise given Ξ_l.

(ii) The approach proposed for the evaluation of the model quality can also be used for the verification of the chosen optimal sampling times. Exactly, the prediction employing only data in optimal sampling times can be compared with the rest of (not used) data from the extended set, how they "fit together".

6.3 Quantitative Parameters in General

It is expected that the parameter estimates can be used in clinical staging of lymphedema. Point or interval estimates of the elementary time constant, a_l, and the cascade length, d_l, are intuitively good indicators of accumulation dynamics. They are described by the analytically obtained marginal posterior distribution $f(\Xi_l|\, Y_l)$. Its discrete nature makes evaluation of moments of interest straightforward. Besides the model parameters (Ξ_l, b_l) alone describing the individual accumulation dynamics their various functions can be tested on the clinical significance. Some of promising quantifiers motivated by tradition in nuclear medicine are listed bellow.

Maximum of time activity curve and time of its reaching have been the first quantifiers proposed by the physicians. If they should be found from the measured data alone, it would require sequential scintigraphic imaging till the occasion when it is evident the response is falling down. Employing proposed modelling and a few data, their estimates can be found according (4.12) and (4.13) and corresponding posterior distributions of parameters. Unfortunately these quantifiers don't express the global characteristics so the next seems more suitable.

Residence time The residence time, ζ_l, is defined as the accumulated activity within a ROI, divided by the administered activity. It is widely accepted in nuclear medicine as a quantitative global characteristic of accumulation dynamics [37, 82]. With the adopted scaling, the residence time in minutes is found as the area under the time activity curve.

The area under the sampled time activity curve is $\zeta_l = \sum_{t=0}^{\infty} x_t \Delta$, with $\Delta = 1$ (min). This sum converges for the considered stable elementary models with $0 \leq a < 1$ to the value

$$\zeta_l(a_l, d_l, b_l) = \frac{b_l \Delta}{(1-a_l)^{d_l}}. \tag{6.6}$$

Again the expected residence times, $E[\zeta_l|\, Y_l]$, and variation, $V[\zeta_l|\, Y_l]$, (respectively standard deviation) are the most instructive characteristics of the *pdf* $f(\zeta_l|\, Y_l)$.

6.4 Selection of Appropriate Sampling Times

The selection of appropriate sampling times of limb imaging is the critical task of the quantitative scintigraphy evaluation. Routine application of the quantitative scintigraphy can be achieved only if just the measurements serving to morphologic (qualitative) inspections will be exploited. Exactly, it means that the number of images (except the initial one) employed for the quantitative evaluation can not exceed the number 3, yet the physicians' desire is to get along with 2 measurements.

It is obvious that the selection of sampling times has an influence on results of processing. The limited number of sampling times makes this influence even more significant. The basic selection can be done empirically (e.g. the senseless combinations of neighbouring times can be excluded and the rough selection can be done) but the further fine-tuning for the maximal efficiency of the method is necessary. Thus it is desirable to advance beyond the current empirical choice

of standardized sampling times. A collection of patient-specific data, measured more frequently than is the clinical norm, has provided the aggregated experimental data needed for Bayesian inference of an optimal sampling grid.

The basic principle of the task of selection of appropriate sampling times can be expressed as looking for such times for which the "desirable results are the most informative". Its transformation to the scheme of Bayesian decision structure strongly depends on the designer demands. This aim belongs to the global tasks for which data of all the patients from experimental group is employed. There will be described two alternative versions of the mathematical formulation of this task. The difference between them is in the demanded result. The first version aims to find optimal sampling times for reporting the inference (beliefs) about demanded quantities, Section 6.4.1, while the second version is done with respect to report only point estimate of these quantities, see Section 6.4.2. Let's remark that the discussion bellow is related to the proposed models from the Chapter 4, but its generalization to arbitrary models and parametrization is possible.

6.4.1 Optimal Sampling Times for Reporting Beliefs

The triple $(a_l, d_l, b_l) \equiv (\Xi_l, b_l) \equiv \Phi_l$ are those model parameters of interest that describe the limb local relative lymphatic system response (accumulation dynamics). Thus, reporting their beliefs is one of the main tasks. Therefore one possible approach to selection of sampling times should be related with respect to it.

Let's suppose there is available an extended observation set Y_l for each patients' limb within the considered over-sampled experimental set. We can define various sub-selections of sampling times, $\mathcal{S}_l \subset \mathcal{T}_l$, that divide this extended observation set into the selected observations ${}^{[\mathcal{S}]}Y_l$ and remainder ${}^{[\bar{\mathcal{S}}]}Y_l$:

$$Y_l = ({}^{[\mathcal{S}]}Y_l, {}^{[\bar{\mathcal{S}}]}Y_l).$$

Let's also consider now for the simplicity that the sampling times for all the limbs are the same, i.e. $\mathcal{T}_l = \mathcal{T}$. Then the task of selection of sampling times is transformed to the aim to choose sub-selection $\mathcal{S}$ of fixed size from the the extended set of sampling times $\mathcal{T}$.

The proposed approach of selection of sampling times in order to report the beliefs about the parameters of interest Φ_l is based on the following idea: "We want to choose such sub-selection $\mathcal{S}$ for which the estimate of the parameters of interest given the observations in this sub-selection ${}^{[\mathcal{S}]}Y_l$ is close to the estimate done on all the data Y_l." It means to select such sampling times for which the distance between the posterior distributions $f\left(\Phi_l |\, {}^{[\mathcal{S}]}Y_l\right)$ and $f\left(\Phi_l |\, Y_l\right)$ is minimal. The result of the selection of sampling times is considered to serve as the suggestion for the sampling times for all patients. Thus it is the global task solved over all the limbs in the experimental set. Then it will be searched the minimal distance between the joint distributions $f\left(\Phi |\, {}^{[\mathcal{S}]}Y\right)$ and $f\left(\Phi |\, Y\right)$, where $\Phi = (\Phi_1, \Phi_2, \ldots, \Phi_{\mathring{l}})$, $Y = (Y_1, Y_2, \ldots, Y_{\mathring{l}})$ and ${}^{[\mathcal{S}]}Y$ similarly.

Employing the framework of the general static decision task described in the Section 3.2.2 its specific instances for the solved task are specified:

- Decomposition of the behaviour $\mathcal{Q} = (\mathcal{P}_\xi,\, \xi,\, \mathcal{F}_\xi) \equiv (Y, \mathcal{S}, \Phi)$
- Admissible rules are of the form $\mathcal{R} : Y^* \to \mathcal{S}^*$
- Loss function $\mathcal{Z}\left(Y, \mathcal{S}, \Phi\right)$ measuring distance between $f\left(\Phi |\, {}^{[\mathcal{S}]}Y\right)$ and $f\left(\Phi |\, Y\right)$

These elements determine the outer model needed for the decision making, $f\left(\Phi |\, Y, \mathcal{S}\right)$. The adopted natural conditions of decision making (3.7) imply that the sub-selection $\mathcal{S}$ is superfluous

in the conditioning. For the given Y the optimal sub-selection is

$$\hat{\mathcal{S}} \in \arg\min_{\mathcal{S}} \int_{\Phi^*} \mathcal{Z}\left(Y, \mathcal{S}, \Phi\right) f\left(\Phi|\, Y\right) d\Phi. \tag{6.7}$$

It remains only to determine loss function that measures the distance of $f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right)$ and $f\left(\Phi|\, Y\right)$. The proper loss function for reporting beliefs in pure inference problems, consequently for distance between two *pdf*s, is the local logarithmic loss function, see Section 3.2.5. Employing it, the general optimization (6.7) can be expressed as:

$$\hat{\mathcal{S}} \in \arg\min_{\mathcal{S}} \int_{\Phi^*} f\left(\Phi|\, Y\right) \log \frac{f\left(\Phi|\, Y\right)}{f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right)} d\Phi = \arg\min_{\mathcal{S}} \mathcal{D}\left(f\left(\Phi|\, Y\right) || \, f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right)\right). \tag{6.8}$$

It is the minimization of the expected local logarithmic loss (3.12) in reporting *pdf* $f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right)$ rather than the *pdf* $f\left(\Phi|\, Y\right)$ (or minimization of the Kullback-Leibler distance between these distributions). Since the numerator in logarithm in (6.8) is not influenced by the sub-selection $\mathcal{S}$, the optimization can be simplified to

$$\hat{\mathcal{S}} \in \arg\max_{\mathcal{S}} \int_{\Phi^*} f\left(\Phi|\, Y\right) \log f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right) d\Phi. \tag{6.9}$$

Both the formulas (6.8) and (6.9) can be decomposed along the limbs if all the parameters Θ_l are strictly limb local. It is only the case of the models $\mathrm{P}_\bullet$ listed in the Table 4.2 at the page 40. In the case where some part of parametric space is common for more limbs the computations are more difficult. However the simplification described in Section 5.2.2, at the page 49, allows us to do it also for the considered models $\mathrm{N}_\bullet$ from the Table 4.2. Then the partial integrations can be done on individual limbs. Only the final optimization is done globally in the following manner:

$$\hat{\mathcal{S}} \in \arg\max_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \int_{\Phi_l^*} f\left(\Phi_l|\, Y_l\right) \log f\left(\Phi_l|\,{}^{[\mathcal{S}]}Y_l\right) d\Phi_l = \arg\min_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \mathcal{D}\left(f\left(\Phi_l|\, Y_l\right) || \, f\left(\Phi_l|\,{}^{[\mathcal{S}]}Y_l\right)\right). \tag{6.10}$$

Remark(s) 6.4.1

(i) In order to understand more deeply the optimized formula in (6.10) it is interesting to look what is hidden behind the term $\mathcal{D}\left(f\left(\Phi_l|\, Y_l\right) || \, f\left(\Phi_l|\,{}^{[\mathcal{S}]}Y_l\right)\right)$*. It can be simply rearranged with the help of basic probabilistic calculus into:*

$$\mathcal{D}\left(f\left(\Phi_l|\, Y_l\right) || \, f\left(\Phi_l|\,{}^{[\mathcal{S}]}Y_l\right)\right) = \int_{\Phi_l^*} f\left(\Phi_l|\, Y_l\right) \log f\left({}^{[\bar{\mathcal{S}}]}Y_l|\, \Phi_l, {}^{[\mathcal{S}]}Y_l\right) d\Phi_l - \log f\left({}^{[\bar{\mathcal{S}}]}Y_l|\,{}^{[\mathcal{S}]}Y_l\right). \tag{6.11}$$

It contains the predictive term $f\left({}^{[\bar{\mathcal{S}}]}Y_l|\,{}^{[\mathcal{S}]}Y_l\right)$ *which has been considered intuitively from the beginning as a good indicator for selecting the sampling times, however, the other term influences the result of optimization.*

(ii) It is also possible to propose modified approach, how to select sampling times in order to report the beliefs about the parameters of interest Φ*, employing little different loss function. The modification is that we want to minimize the distance between the posterior distribution* $f\left(\Phi|\,{}^{[\mathcal{S}]}Y\right)$ *given the observations in the sub-selection* ${}^{[\mathcal{S}]}Y$ *and similar posterior distribution* $f\left(\Phi|\, Y, \widetilde{Y}\right)$ *given not all the available data but all (even) hypothetically available data*

$(Y, \widetilde{Y})$. $\widetilde{Y}$ *denotes hypothetically available data in non-sampled times. Employing again the framework of the general static decision task described in the Section 3.2.2, only the decomposition of the behaviour* $\mathcal{Q} = (\mathcal{P}_\xi, \xi, \mathcal{F}_\xi) \equiv (Y, \mathcal{S}, (\Phi, \widetilde{Y}))$ *together with loss function* $\mathcal{Z}$ *differ from the previous task. The needed outer model with consideration of natural conditions of decision making is then* $f\left(\Phi, \widetilde{Y} | \, Y\right)$ *and optimal sub-selection is*

$$\hat{\mathcal{S}} \in \arg\min_{\mathcal{S}} \int_{\widetilde{Y}^*} \int_{\Phi^*} f\left(\Phi, \widetilde{Y} | \, Y\right) \log \frac{f\left(\Phi | \, Y, \widetilde{Y}\right)}{f\left(\Phi | \, {}^{[\mathcal{S}]}Y\right)} d\Phi d\widetilde{Y}.$$

The numerator in the logarithm is again not influenced by the sub-selection $\mathcal{S}$ *and in the remainder,* $\widetilde{Y}$ *can be integrated out so the optimization simplifies into*

$$\hat{\mathcal{S}} \in \arg\max_{\mathcal{S}} \int_{\Phi^*} f\left(\Phi | \, Y\right) \log f\left(\Phi | \, {}^{[\mathcal{S}]}Y\right) d\Phi,$$

that is identical with (6.9). It means the results of both proposed approaches are the same.

(iii) All the considerations here were related to the purpose of reporting beliefs about model parameters of interest describing relative lymphatic system response. It is due to the assumption these parameters are somehow important for the disease assessment, but the specific quantifier is not known yet. If some quantifier K_l *would be chosen as an appropriate for the disease assessment and its belief is of interest then, instead of (6.10), it would be better to do optimization with respect to this quantifier*

$$\hat{\mathcal{S}} \in \arg\min_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \mathcal{D}\left(f\left(K_l | \, Y_l\right) || \, f\left(K_l | \, {}^{[\mathcal{S}]}Y_l\right)\right).$$

The formulas for the computation of maximized limb term $\mathcal{W}_l = \int_{\Phi_l^*} f\left(\Phi_l | \, Y_l\right) \log f\left(\Phi_l | \, {}^{[\mathcal{S}]}Y_l\right) d\Phi_l$ in (6.10) for individual models from the Table 4.2 are listed here, for completeness.

Maximized limb term $\mathcal{W}_l$ for models $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{N_1}$, $\mathbf{N_2}$

$$\mathcal{W}_l = \left(\sum_{\Xi_l^*, b_l^*} w_{\Xi_l, b_l}\right)^{-1} \sum_{\Xi_l^*, b_l^*} w_{\Xi_l, b_l} \log^{[\mathcal{S}]} w_{\Xi_l, b_l} - \log \sum_{\Xi_l^*, b_l^*} {}^{[\mathcal{S}]}w_{\Xi_l, b_l}$$

Maximized limb term $\mathcal{W}_l$ for models $\mathbf{N_3}$, $\mathbf{N_4}$

$$\begin{aligned} \mathcal{W}_l \;=\; & -\log \sum_{\Xi_l^*} {}^{[\mathcal{S}]}w_{\Xi_l} + \left(\sum_{\Xi_l^*} w_{\Xi_l}\right)^{-1} \sum_{\Xi_l^*} w_{\Xi_l} \left[\log \frac{\Gamma({}^{[\mathcal{S}]}\chi_1 + 1/2)}{\Gamma({}^{[\mathcal{S}]}\chi_1)\Gamma(1/2)} + \log\left(\omega_{\Xi_l}^{1/2} \, {}^{[\mathcal{S}]}\chi_4^{-({}^{[\mathcal{S}]}\chi_1 + 1/2)}\right)\right. \\ & \left. -({}^{[\mathcal{S}]}\chi_1 + 1/2) \int_{b_l^*} \log\left(1 + \frac{\omega_{\Xi_l}\chi_2}{{}^{[\mathcal{S}]}\chi_4}\left(b_l - \frac{{}^{[\mathcal{S}]}\chi_3}{{}^{[\mathcal{S}]}\chi_2}\right)^2\right) \mathrm{St}\left(b_l | \frac{\chi_3}{\chi_2}, \frac{2\omega_{\Xi_l}\chi_1\chi_2}{\chi_4}, 2\chi_1\right) db_l\right] \end{aligned}$$

The integral in this term can not be evaluated analytically, therefore the numerical evaluation is necessary. In this case e.g. the Laplace approximation can be employed [6, 42].

6.4.2 Optimal Sampling Times for Reporting Point Estimates

The Section 6.4.1 was devoted to the selection of sampling times in order to report beliefs about quantities of interest. If, on the other hand, only point estimates of the quantities are sufficient for further processing then the approach of selection of sampling times can be related to these point estimates.

Let's consider the point estimate of a selected quantity K_l is wanted. Thereafter if the popular quadratic loss function is used, the optimal point estimate of K_l on the base of the selected observations ${}^{[\mathcal{S}]}Y_l$ is the conditional expectation ${}^{[\mathcal{S}]}\hat{K}_l = E\left[K_l|\,{}^{[\mathcal{S}]}Y_l\right]$, see Section 3.2.5. Then the proposed approach of selection of sampling times in order to report these point estimates can be formulated as follows: "We want to choose such sub-selection $\mathcal{S}$ for which the point estimate of the desired quantity K_l given the observations in this sub-selection ${}^{[\mathcal{S}]}Y_l$, ${}^{[\mathcal{S}]}\hat{K}_l = E\left[K_l|\,{}^{[\mathcal{S}]}Y_l\right]$, minimizes the expectation of the loss function $\mathcal{Z}$, that quantifies the quality of the point estimate." Again it is the global task solved over all limbs in the experimental set, then the task is solved with respect to global parameter $K = \left(K_1, K_2, \ldots, K_{\mathring{l}}\right)$ and its point estimate ${}^{[\mathcal{S}]}\hat{K}$.

Employing the framework of the general static decision task described in the Section 3.2.2 its specific instances for the solved task are similar to those in the previous Section:

- Decomposition of the behaviour $\mathcal{Q} = (\mathcal{P}_\xi,\, \xi,\, \mathcal{F}_\xi) \equiv (Y, \mathcal{S}, K)$
- Admissible rules $\mathcal{R} : Y^* \to \mathcal{S}^*$
- Loss function $\mathcal{Z}\,(Y, \mathcal{S}, K)$ quantifying the quality of the estimate ${}^{[\mathcal{S}]}\hat{K} = E\left[K|\,{}^{[\mathcal{S}]}Y\right]$

Outer model needed for the decision making is $f\,(K|\,Y, \mathcal{S})$. Again the adopted natural conditions of decision making (3.7) imply that the sub-selection $\mathcal{S}$ is superfluous in the conditioning. For the given Y the optimal sub-selection is

$$\hat{\mathcal{S}} \in \arg\min_{\mathcal{S}} \int_{\Phi^*} \mathcal{Z}\,(Y, \mathcal{S}, K)\, f\,(K|\,Y)\, dK. \tag{6.12}$$

The possible loss function for quantification of the quality of the estimate ${}^{[\mathcal{S}]}\hat{K}$, can be defined as

$$\mathcal{Z} = \sum_{l=1}^{\mathring{l}} \left(K_l - {}^{[\mathcal{S}]}\,\hat{K}_l\right)^T \boldsymbol{M} \left(K_l - {}^{[\mathcal{S}]}\,\hat{K}_l\right), \tag{6.13}$$

where $\boldsymbol{M}$ is a symmetric positive definitive matrix. Employing it together with the decomposition along limbs (5.17), or (5.19) (together with the model simplification proposed in Section 5.2.2, at the page 49, for the models $\mathrm{N}_\bullet$), the optimization (6.12) can be expressed as

$$\begin{aligned}
\hat{\mathcal{S}} \;&\in\; \arg\min_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \int_{K_l^*} \left(K_l - {}^{[\mathcal{S}]}\,\hat{K}_l\right)^T \boldsymbol{M} \left(K_l - {}^{[\mathcal{S}]}\,\hat{K}_l\right) f\,(K_l|\,Y_l)\, dK_l = \\
&=\; \arg\min_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \left[\mathrm{tr}(\boldsymbol{M} V\,[K_l, |\,Y_l]) + \left({}^{[\mathcal{S}]}\hat{K}_l - \hat{K}_l\right)^T \boldsymbol{M} \left({}^{[\mathcal{S}]}\hat{K}_l - \hat{K}_l\right)\right],
\end{aligned} \tag{6.14}$$

where $\hat{K}_l = E\,[K_l|\,Y_l]$ and $\mathrm{tr}(\cdot)$ denotes matrix trace. The first term of (6.14) does not depend on the sub-selection $\mathcal{S}$, therefore only the second term has to be minimized. If the evaluation is done for the scalar quantifier K_l only on one limb, i.e. $\mathring{l} = 1$ it is easily checked that the preferred sampling times are those for which the estimate of K_l is the nearest to estimate found from all observations. For more limbs the sampling times with minimal sum of square of difference between the point estimates done on selected and all data are preferred.

Remark(s) 6.4.2

(i) Different point estimate of a quantity K_l and different loss function $\mathcal{Z}$ (6.13) could be used in general. The quadratic forms were selected here only due to the analytical solution of optimization and accordingly possibility to interpret the results. However, the selection of the suitable loss function is not straightforward, see bellow.

(ii) This approach to the optimization has one general drawback that is clearly visible on the example with quadratic loss function. Both the local loss function that determines the point estimates ${}^{[\mathcal{S}]}\hat{K}_l$ and the loss function $\mathcal{Z}(Y, \mathcal{S}, K)$ in (6.12) of the global optimization of sampling times should be selected with respect to the posterior distributions $f\left({}^{[\mathcal{S}]}K_l|\, Y_l\right)$ and $f(K_l|\, Y_l)$. For instance, it is evident that the optimization (6.14) with constant matrix $\boldsymbol{M}$ for all limbs is not suitable for cases where the level of uncertainty of estimates is different for individual limbs. The quadratic loss functions are appropriate for normal models in general. In this specific case the loss function with constant $\boldsymbol{M}$ for all limbs suits only for normal models $f\left({}^{[\mathcal{S}]}K_l|\, {}^{[\mathcal{S}]}Y_l\right) = N({}^{[\mathcal{S}]}K_l|\, {}^{[\mathcal{S}]}\mu_l, s)$ with common precision parameter s. For other cases it can be unsuitable.

The optimization for reporting beliefs based on the Kullback-Leibler distance described in the Section 6.4.1 is in this respect better, as it takes into consideration the whole probability distribution. Employing the approach of optimization for reporting beliefs about the parameters of interest K_l, where $f\left({}^{[\mathcal{S}]}K_l|\, {}^{[\mathcal{S}]}Y_l\right) = N({}^{[\mathcal{S}]}K_l|\, {}^{[\mathcal{S}]}\mu_l, {}^{[\mathcal{S}]}s_l)$ $f(K_l|\, Y_l) = N(K_l|\, \mu_l, s_l)$, the result of optimization (6.10) is:

$$\hat{\mathcal{S}} \in \arg\max_{\mathcal{S}} \sum_{l=1}^{\mathring{l}} \left[\frac{1}{2} \log\left(\frac{{}^{[\mathcal{S}]}s_l}{2\pi} \right) - \frac{{}^{[\mathcal{S}]}s_l}{2} \left(({}^{[\mathcal{S}]}\mu_l - \mu_l)^2 + s_l^{-1} \right) \right].$$

It coincides with (6.14) only for the common ${}^{[\mathcal{S}]}s_l$ for all limbs and selections. If the distribution is normal then the Kullback-Leibler distance is practically the quadratic loss function, if not then the general Kullback-Leibler distance leads to the loss function appropriate for this distribution. Therefore the optimization for reporting beliefs in the Section 6.4 is superior to the approach described in this Section and only the former will be used onward.

(iii) It was considered commonly for both optimization versions at the very beginning that the sampling times for all the limbs are the same for the simplicity, i.e. $\mathcal{T}_l = \mathcal{T}$. However it is not possible to ensure it in reality. Therefore the task of selection of sampling times has to be generalized e.g. to intervals. It means to group similar sampling times into intervals and perform the optimization with respect to these intervals. It can be done since the theoretical conclusions for sampling times and intervals are the same.

6.5 Comparison of Accumulation on Both Limbs

Comparison between the responses of a particular patient's upper limbs is a very useful diagnostic aid, since, often, it is known that one limb is healthy, and it can act as a control for evaluation of the other limb. Technically, the need for a quantitative comparison of just two or three scintigraphic images is implied, a task which is hopeless without a Bayesian treatment. Over here the lexical terms that the responses on both limbs are the same/different is translated into the mathematical terms and the task of making statement transformed to the decision-making task.

This problem is formalized here in the general framework of hypothesis testing, see Section 3.2.5 [6, 47]. A selection of better variant between alternative hypotheses "response is same/different" is done. The individual hypotheses can be defined in the following manner:

H_0**: The same responses** The correspondence of responses on both limbs can be expressed as that the accumulation characteristics on both limbs are the same, i.e. the parameters $(\Xi_l, b_l) \equiv \Phi_l$ determining these characteristics are the same and common for both limbs:

$$\Phi_1 = \Phi_2 = \Phi.$$

H_1**: Different responses** The difference of responses implies the different accumulation characteristics, i.e. the parameters Φ_1 and Φ_2 are generally different and local for each limb.

Putting this testing of hypothesis into the general framework of static decision task, Section 3.2.2, we come to its specific instances for this task

- Decomposition of the behaviour $\mathcal{Q} = (\mathcal{P}_\xi,\, \xi,\, \mathcal{F}_\xi) \equiv ((Y_1, Y_2), \hat{h}, (\Phi_1, \Phi_2, \{H_h\}_{h\in\{0,1\}}))$
- Admissible rules are of the form $\mathcal{R} : (Y_1, Y_2)^* \to h^*$
- Loss function $\mathcal{Z}\left((Y_1, Y_2), \hat{h}, (\Phi_1, \Phi_2, h)\right) \geq 0$ is a $(2, 2)$ table typically with zero diagonal entries as no penalty is paid when $\hat{h} = h$. Positive entries reflect medical and economic consequences of a bad selection. Mostly, the simplified loss function $\mathcal{Z}\left(\hat{h}, h\right)$ is used.

The outer model needed is then $f\left(h, \Phi_1, \Phi_2 |\, Y_1, Y_2, \hat{h}\right)$. The natural conditions of decision making (3.7) imply the decision $\hat{h}$ is superfluous in conditioning. The Bayesian decision for the general loss function $\mathcal{Z}$ is the result of optimization

$$\hat{h} \in \arg\min_{\hat{h}\in\{0,1\}} \sum_{h\in\{0,1\}} \mathcal{Z}\left((Y_1, Y_2), \hat{h}, (\Phi_1, \Phi_2, h)\right) f\left(h, \Phi_1, \Phi_2 |\, Y_1, Y_2\right), \tag{6.15}$$

while for the simplified loss function $\mathcal{Z}\left(\hat{h}, h\right)$ it is

$$\hat{h} \in \arg\min_{\hat{h}\in\{0,1\}} \sum_{h\in\{0,1\}} \mathcal{Z}\left(\hat{h}, h\right) f\left(h |\, Y_1, Y_2\right). \tag{6.16}$$

The needed conditional *pf* $f\left(h |\, Y_1, Y_2\right)$ in (6.16) and mixed-type *p(d)f* $f\left(h, \Phi_1, \Phi_2 |\, Y_1, Y_2\right)$ in (6.15) can be obtained using the chain and Bayes rules:

$$\begin{aligned} f\left(h, \Phi_1, \Phi_2 |\, Y_1, Y_2\right) &= f\left(\Phi_1, \Phi_2 |\, h, Y_1, Y_2\right) f\left(h |\, Y_1, Y_2\right) \\ f\left(h |\, Y_1, Y_2\right) &\propto f\left(Y_1, Y_2 |\, h\right) f\left(h\right). \end{aligned} \tag{6.17}$$

The simplified loss function is employed only here, so the stress is on the computation of $f\left(Y_1, Y_2 |\, h\right)$. Though, for the general case $f\left(\Phi_1, \Phi_2 |\, h, Y_1, Y_2\right)$ can be obtained too.

The null hypothesis $h = 0$ defines the local parameters to be same for both limbs. Therefore the predictive *pdf* $f\left(Y_1, Y_2 |\, h = 0\right)$ cannot be be evaluated using formulas (5.18) and (5.21) in the Section 5.2, where the limb local parameters are considered generally different. In the case of models $\mathrm{P}_\bullet$, see Table 4.2, the model and null hypothesis assumptions imply that only parameter

g_l is limb local, the others are common for both limbs. Then employing parametrised model (4.10) and modified prior distribution, the predictive *pdf* $f\left(Y_1, Y_2 \middle| h=0\right)$ can be expressed as

$$f\left(Y_1, Y_2 \middle| h=0\right) = \sum_{\Xi^*, b^*} \left[\prod_{l=1}^{2} \int_{g_l^*} f\left(Y_l \middle| \Xi, b, g_l\right) f\left(g_l\right) dg_l \right] f\left(\Xi, b\right).$$

The same predictive *pdf* for normal models $\mathrm{N}_\bullet$ in the Table 4.2 is even simpler. The parameter s is considered to be common for all limbs from the beginning. Therefore all the parameters are common for both limbs in null hypothesis. Thus it follows for models N_1, N_2:

$$f\left(Y_1, Y_2 \middle| h=0\right) = \sum_{\Xi^*, b^*} \left[\prod_{l=1}^{2} f\left(Y_l \middle| \Xi, b, s\right) \right] f\left(s\right) ds\, f\left(\Xi, b\right).$$

Similar formula can be obtained also for the models N_3, N_4. The prediction $f\left(Y_1, Y_2 \middle| h=1\right)$ for the alternative hypothesis $h=1$ is computed according to (5.18) and (5.21) where $\mathring{l}=2$. Its evaluation is obvious for the models $\mathrm{P}_\bullet$. For the models $\mathrm{N}_\bullet$ application of the approximation of the *pdf* of common parameter s or model simplification described in the Section 5.2.2 are required.

Computation of the posterior distribution $f\left(h \middle| Y_1, Y_2\right)$, see (6.17), requires the prior probability of the hypotheses $f\left(h\right)$. Since in the moment of making the decision about the correspondence of limbs the information from clinicians is not available and the evaluation is done due to the doubt about the limb differences, the prior distribution on the space of the hypothesis is selected uniform. To eliminate the prior distribution $f\left(h\right)$ in the temporary computation, only the posterior to prior odds ratio called the Bayes factor [6] is computed:

$$\mathcal{B}_{01}(Y_1, Y_2) = \frac{f\left(Y_1, Y_2 \middle| h=0\right)}{f\left(Y_1, Y_2 \middle| h=1\right)} = \frac{f\left(h=0 \middle| Y_1, Y_2\right)}{f\left(h=1 \middle| Y_1, Y_2\right)} \Big/ \frac{f\left(h=0\right)}{f\left(h=1\right)}. \tag{6.18}$$

It provides a measure of whether the data (Y_1, Y_2) have increased or decreased the odds on $h=0$ relative to $h=1$. The posterior hypothesis probabilities are then given by

$$f\left(h=0 \middle| Y_1, Y_2\right) = \frac{\mathcal{B}_{01}(Y_1, Y_2) f_{01}}{1 + \mathcal{B}_{01}(Y_1, Y_2) f_{01}}, \quad f\left(h=1 \middle| Y_1, Y_2\right) = \frac{1}{1 + \mathcal{B}_{01}(Y_1, Y_2) f_{01}},$$

where $f_{01} = f\left(h=0\right) / f\left(h=1\right)$.

The posterior probability $f\left(h=0 \middle| Y_1, Y_2\right)$ is a temporary final result, if the simplified loss function (table) $\mathcal{Z}\left(\hat{h}, h\right)$ is not determined yet, that allows to make the final decision later. If the loss $\mathcal{Z}$ is specified then the hypothesis of limb equality $\hat{h}=0$ is accepted if:

$$f\left(h=0 \middle| Y_1, Y_2\right) \geq \mathcal{P} = \frac{\mathcal{Z}\left(\hat{h}=0, h=1\right)}{\mathcal{Z}\left(\hat{h}=0, h=1\right) + \mathcal{Z}\left(\hat{h}=1, h=0\right)}. \tag{6.19}$$

That is to say, the loss function $\mathcal{Z}\left(\hat{h}, h\right)$ determines the probability bound (threshold) $\mathcal{P}$ for acceptance of the null hypothesis. This threshold can be therefore understood as tuning knob of this decision task.

In order to make this section complete the exact formulas (6.18) for individual models are listed. It is supposed hereafter that the sampling times for both limbs are the same, i.e. $\mathcal{T}_1 = \mathcal{T}_2$.

Bayes factor for models $\mathbf{P_1}$, $\mathbf{P_2}$

$$\mathcal{B}_{01}(Y_1,Y_2) = \frac{\sum_{\Xi^*,b^*\{\Xi=\Xi_1=\Xi_2,b=b_1=b_2\}} w_{\Xi_1,b_1} w_{\Xi_2,b_2}}{\sum_{\Xi_1^*,b_1^*} w_{\Xi_1,b_1} \sum_{\Xi_2^*,b_2^*} w_{\Xi_2,b_2} f\left(\Xi_2,b_2\right)}$$

Bayes factor for models $\mathbf{N_1}$, $\mathbf{N_2}$

$$\mathcal{B}_{01}(Y_1,Y_2) = \frac{\Gamma(n+\alpha_s)\Gamma(\breve{\alpha}_s) \sum_{\Xi^*,b^*} \left[\beta_s + \frac{1}{2}\sum_{l=1}^{2} ||^N Y_l - bA_\Xi||^2\right]^{-(n+\alpha_s)}}{\Gamma(\chi_{1,l=1})\Gamma(\breve{\chi}_{1,l=2})\breve{\beta}_s^{\breve{\alpha}_s} \sum_{\Xi_1^*,b_1^*} w_{\Xi_1,b_1} \sum_{\Xi_2^*,b_2^*} \breve{w}_{\Xi_2,b_2} f\left(\Xi_2,b_2\right)} \tag{6.20}$$

Bayes factor for models $\mathbf{N_3}$, $\mathbf{N_4}$

$$\mathcal{B}_{01}(Y_1,Y_2) = \frac{\Gamma(n+\alpha_s)\Gamma(\breve{\alpha}_s) \sum_{\Xi^*,b^*} \frac{\left[\sum_{l=1}^{2} ||^N Y_l||^2 + 2\beta_s + \omega_\Xi \left(\hat{b}_\Xi^2 - \frac{\left(\hat{b}_{\Xi+\omega_\Xi^{-1}} \left(\sum_{l=1}^{2} {}^N Y_l\right)^T A_\Xi\right)^2}{1+2\omega_\Xi^{-1}||A_\Xi||^2}\right)\right]^{-(n+\alpha_s)}}{\left(1+2\omega_\Xi^{-1}||A_\Xi||^2\right)^{1/2}}}{\Gamma(\chi_{1,l=1})\Gamma(\breve{\chi}_{1,l=2})(2\breve{\beta}_s)^{\breve{\alpha}_s} \sum_{\Xi_1^*} w_{\Xi_1} \sum_{\Xi_2^*} \breve{w}_{\Xi_2} f\left(\Xi_2\right)} \tag{6.21}$$

The notation $\breve{\bullet}$ in (6.20) and (6.21) allows us to express the formulas for normal models $\mathrm{N}_\bullet$ for both versions of computation. The first version assumes parameter s to be common for both limbs and employs approximation (5.22). The second version employs model simplification with local limb parameter s_l. The former supposes that terms $\breve{\alpha}_s, \breve{\beta}_s, \breve{w}.$ are substituted or computed with the parameters α, β of approximating single Gamma distribution $\hat{f}\left(s|\, Y_1\right)$ while the latter uses repeatedly the hyper-parameters α_s, β_s of the prior distribution of limb specific parameter s_l. Similarly to the estimate of common parameter s, the results for the first version employing approximation depends on the order of limb processing.

6.6 Summary

This Chapter is an important part of this work, where the proposed modelling, Chapter 4, and Bayesian inference, Chapter 5, are employed for the solution of majority of tasks useful for the scintigraphic evaluation of limbs listed in the objectives of the thesis, Section 1.3.

The main outcomes of the Chapter are:

- estimate of the "whole" time activity curve from sparse measurements. This curve can be used for visual evaluation alone. Besides it is the only chance how the quantification techniques proposed by another authors that need the whole or some points of time activity curve can be used having only sparse data. The mean and marginal standard deviation of time activity curve are basic outcomes that express also uncertainty of the estimate.

- quantitative test, that compares two limbs on the base of just two or three scintigraphic images. It is based on the Bayesian hypothesis testing of two alternative hypotheses, whose outcome is binary decision if the accumulation characteristics on both limbs are the same or different.

- Bayesian treatment of a selection of appropriate sampling times for scintigraphic evaluation. According my best knowledge, the proposed approach in the Bayesian framework has not been described in any publication, yet. Though the idea behind is straightforward. Two

versions of sampling time optimization according the demanded results of scintigraphy quantification were proposed, however only the superior version was decided to use. They were constructed for the specific scintigraphic problem and considered model, however the results can be generalized to any problem and model.

Besides some promising quantifiers for the assessment of staging of lymphedema are proposed, that can be computed by means of estimated model. This Chapter and also the whole work in this occasion does not solve the problem of searching a suitable quantitative parameter, that are correlated with the disease staging. It belongs still to open problems. Finally, the sense of output reconstruction at non-sampled times, Section 6.2, is for the verification of proposed models and also for verification of the optimized sampling times.

The Chapter 7 demonstrates all the theoretical results on practical experiments.

Chapter 7

Experiments

This Chapter is devoted to the experiments on real data. It comprises the experimental results whose main aim is to verify applicability of the proposed models and algorithms on the quantitative scintigraphy problem. The general aim of the Chapter consists of several important primary goals:

(i) Demonstrate that practically all the algorithms work.

(ii) Show that the results are reliable.

(iii) Demonstrate that results are sufficiently insensitive to the optional parameters — tuning knobs.

(iv) Show that results are meaningful and give to the physicians expected benefit.

(v) Summarize open problems, which appear as doubts in experiments dedicated to the all previous goals.

All the presented experiments hereafter are meant to demonstrate some of these goals. Some of them can cover even more goals. The goal (i) is in the background of all experiments, for instance.

At the beginning, the Section 7.1 contains information about the set of real data used in all the experiments. The Section 7.2 deals with necessary and optional initial tuning of the algorithm, that have preceded to all the subsequent tasks. The extensive set of local limb and patient inferences for all the patients is demonstrated over here by the illustrative examples in the Section 7.3. The "best" model from the considered set of models is being selected in the Section 7.4. The Section 7.5 is devoted to the experimental verification of the robustness of the estimator to the optional parameters that determine the prior information. The results of the proposed quantitative comparative test is compared with the decisions of physicians in the Section 7.6 and the sampling times are optimized in the Section 7.7. At the end, the potential model employment for the disease staging assessment is outlined in the Section 7.8.

The basic information about the implementation of the algorithms employed for generating the results can be found in the Appendix A. There are also concentrated all the tuning knobs of algorithms with their default values implicitly used in the experiments if not specified otherwise.

7.1 Data Source

Testing of algorithms on data is important to consider their stability and reliability. However, testing on "artificial" simulated data can hide some pitfalls, because data values prepared artificially are "pretty" and need not reflect situations that can be met with real data. Therefore the real scintigraphic data are used.

The algorithms were tested on an experimental group of well checked data of 32 limbs. For each limb in the group the data on all ROIs are available. The number of data for each ROI and limb is between 3 and 4 (besides the initial injected activity). All the subjects whose limbs are used in the experiments are females after the breast cancer treatment with suspicion to lymphedema or its early stages. Limbs with the advanced stages of lymphedema were not included standardly (apart from one exception) in the experimental set intentionally, as the corresponding data sets were incomplete.

There is also available, besides the scintigraphic data, information about the assessment of the stage of lymphedema for each ROI and limb and about the correspondence between limbs made by the clinician and nuclear medicine expert (more details can be found in the Section 2.2.3). It allows us to compare some results with the conclusions of independent experts.

7.2 Tuning phase

Before the proposed models and resulting algorithms can be employed for the purpose of quantitative evaluation of scintigraphy it is necessary to tune up their parts, mainly those that have been specified only roughly during the definition of models. The whole experimental group of data is used to "learn" these parts. It concerns first and foremost the prior distribution of the noise precision parameter s of the normal models $\mathrm{N}_\bullet$, see Table 4.2. Though the rest of prior information could be possibly refined employing indicated but not fully formalized common collective hyper-parameters in the Remark(s) 4.3.1. It will be demonstrated here on the range of model orders d_l.

7.2.1 Estimate of the Common Noise Parameter s

Let's recall that the precision parameter s for normal models $\mathrm{N}_\bullet$ is considered to be common for all limbs characterizing the measurement process. It is the reciprocal value of the variance of measurement noise. Its prior distribution was initially determined in the Section 4.3.3 roughly according to the supposed range of a priori possible values. It was considered that the range and distribution of this parameter common for all limbs can be refined by processing of sets of ROI data. The right instrument how to refine it was described in the Section 5.2.2 at the page 47 where also two possible approximations for realizable computation were proposed. It was the initial task preceding all further processing.

Results The initial prior range of the parameter s was chosen $1e5 < s < 1e8$ and corresponding gamma distribution determined by means of (4.18) in the Section 4.3.3. This distribution is flat enough as its hyper-parameter $\alpha_s \approx 1$. Using this initial prior distribution the noise parameter s was estimated on the experimental set of data employing both proposed approximations. In order to suppress the influence of the order of processed limbs on the results the solution mentioned in Section 5.2.2 was applied. The estimation was done on 30 permutations of limbs and summarized into one gamma distribution. The results for individual ROIs and considered normal models $\mathrm{N}_\bullet$ are summarized in Table 7.1.

	Estimate $f(s\|Y) = Ga(s\|\alpha_s, \beta_s)$							
	Approx. via discrepancy				Approx. via moments			
Model	α_s	β_s	$E[s]$	$\sigma(s)$	α_s	β_s	$E[s]$	$\sigma(s)$
N_1	43.03	8.35e-5	5.15e5	7.86e4	43.07	8.35e-5	5.16e5	7.86e4
N_2	43.31	8.35e-5	5.18e5	7.88e4	43.40	8.36e-5	5.19e5	7.88e4
N_3	43.13	8.34e-5	5.17e5	7.87e4	43.16	8.34e-5	5.17e5	7.87e4
N_4	43.69	8.35e-5	5.23e5	7.91e4	43.73	8.35e-5	5.23e5	7.91e4

(a) ROI=axilla

	Estimate $f(s\|Y) = Ga(s\|\alpha_s, \beta_s)$							
	Approx. via discrepancy				Approx. via moments			
Model	α_s	β_s	$E[s]$	$\sigma(s)$	α_s	β_s	$E[s]$	$\sigma(s)$
N_1	42.39	1.14e-5	3.71e6	5.71e5	42.4	1.14e-5	3.72e6	5.71e5
N_2	42.54	1.15e-5	3.69e6	5.67e5	42.47	1.15e-5	3.69e6	5.67e5
N_3	42.87	1.12e-5	3.82e6	5.84e5	42.79	1.12e-5	3.82e6	5.85e5
N_4	42.99	1.13e-5	3.78e6	5.77e5	42.90	1.13e-5	3.78e6	5.77e5

(b) ROI=upper arm

	Estimate $f(s\|Y) = Ga(s\|\alpha_s, \beta_s)$							
	Approx. via discrepancy				Approx. via moments			
Model	α_s	β_s	$E[s]$	$\sigma(s)$	α_s	β_s	$E[s]$	$\sigma(s)$
N_1	42.46	2.58e-5	1.65e6	2.52e5	42.28	2.58e-5	1.63e6	2.52e5
N_2	41.76	2.61e-5	1.6e6	2.47e5	41.71	2.59e-5	1.64e6	2.49e5
N_3	43.47	2.60e-5	1.67e6	2.53e5	43.52	2.60e-5	1.67e6	2.53e5
N_4	42.63	2.64e-5	1.62e6	2.47e5	42.63	2.64e-5	1.62e6	2.47e5

(c) ROI=forearm

Table 7.1: Estimates of the precision parameter s from the data of experimental set. Tables for individual ROIs contain estimates for all models $N_\bullet$ and two proposed approximations. Both parameters (α_s, β_s) of Gamma distribution are given together with mean $E[s]$ and standard deviation $\sigma(s)$ of the distribution.

Discussion There are almost no difference between the pairs of results in Table 7.1 for both approximations described in the Section 5.2.2 and the proposed normal models from the Table 4.2. Employing both approximations similar results are obtained, hence there is no reason here to prefer some of them. The differences among estimates of s, i.e. noise levels, on the ROIs, can be explained by the used simplification in normal models. In the original Poisson model the level of noise depends on the level of the data. Used simplification for models $\mathrm{N}_\bullet$ assumes the constant noise level for all data. The consequence is that the estimated *pdf* of s reflects the highest level of the noise in the original raw data on ROIs. This idea was already used for the initial rough setting of the noise. The estimated noise precision corresponds to the 30% noise to signal ratio in the worst case of low data and about 1% in the best case. Such boundary ratios are reasonable for the estimation and the estimated ranges agree with the assumption about the noise. Thus the proposed normal models from this point of view seem meaningful. It can be also understood as a special experiment for the check of meaningfulness of the results, the goal (iv) from the list of primary goals at page 65.

The different results of this task for individual ROIs were then employed as the prior *pdf* about the parameter s for further processing.

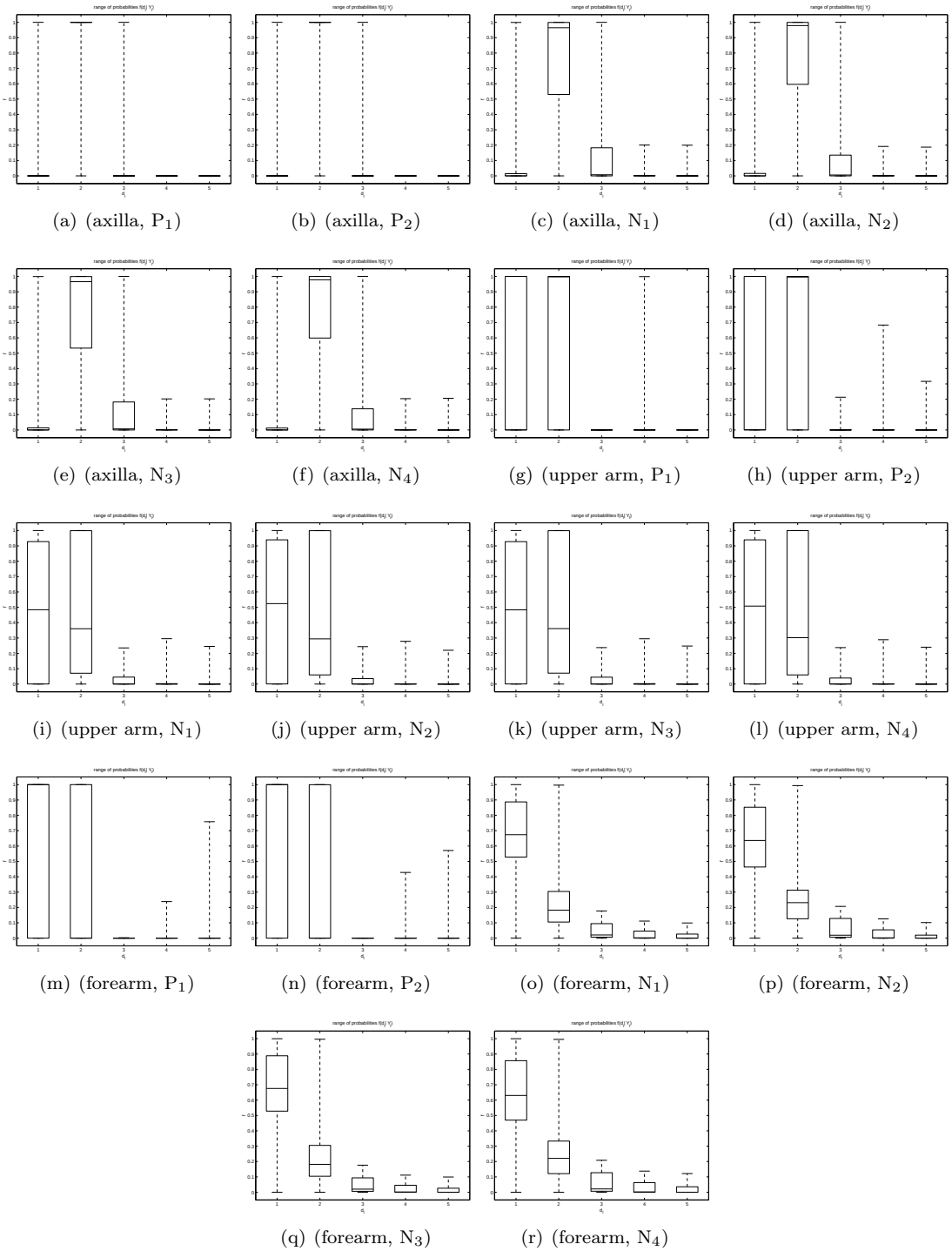
7.2.2 Range of the Model Orders d_l

The prior information about the model order d_l was determined in the Section 4.3.3 in the form of range of a priori possible model orders. It was chosen according the supposed time activity curve shape and comparing it with model responses for various orders d_l. The consequence of this consideration is the conservative range $1 = \underline{d} \leq d_l \leq \overline{d} = 5$. The objective of this Section is to demonstrate that this range is sufficiently large. Besides it will be shown how this range can be narrowed down employing non-formalized refinement through the collective bounds, only indicated in the Remark(s) 4.3.1.

Results The estimates of the model order d_l on individual limbs are good indicators for the testing of sufficiency of the overall model order range. The "significance" of individual model orders d_l on the individual limb and ROI is clearly visible from this result. All the posterior distributions $f\left(d_l \middle| Y_l\right)$ computed for each limb and ROI in the experimental group are summarized in box graphs along the order d_l for each ROI and considered model in the Figure 7.1. The lower and upper line of the boxes mark the 25% and 75% percentiles of the "samples" $f\left(d_l \middle| Y_l\right)$ while the middle line of box is median. The whiskers above and bellow box mark whole range of the "samples" $f\left(d_l \middle| Y_l\right)$.

In general, the fixed range $(\underline{d}, \overline{d})$ of model orders d_l can be considered to be unknown common collective hyper-parameter for all limbs. Then this range can be narrowed down employing the Bayesian learning on the experimental set of data. The posterior distribution $f\left(\underline{d}, \overline{d} \middle| Y\right)$ is the correct quantity from the Bayesian viewpoint necessary for the refinement, however under specific conditions, see e.g. [35], it coincides with the maximum likelihood estimate $\max_{(\underline{d},\overline{d})} f\left(Y \middle| \underline{d}, \overline{d}\right)$. In our case it is not reasonable to narrow the range $(\underline{d}, \overline{d})$ since the maximum likelihood estimate for the majority of models and ROIs coincides with the original conservative range. The only exceptions are the estimate $(1, 3)$ for the Poisson models $\mathrm{P}_\bullet$ on the ROI of axilla and $(1, 4)$ for model P_1 on the upper arm ROI.

Discussion The summarized results of individual posterior distributions $f\left(d_l \middle| Y_l\right)$ expressed by box graphs in the Figure 7.1 are very similar in the groups of normal models $\mathrm{N}_\bullet$ and Poisson models $\mathrm{P}_\bullet$ while there is visible difference in the results between those two groups, especially on the ROI of axilla. Regardless it is possible to make conclusion common for all models.



(a) (axilla, P_1) (b) (axilla, P_2) (c) (axilla, N_1) (d) (axilla, N_2)

(e) (axilla, N_3) (f) (axilla, N_4) (g) (upper arm, P_1) (h) (upper arm, P_2)

(i) (upper arm, N_1) (j) (upper arm, N_2) (k) (upper arm, N_3) (l) (upper arm, N_4)

(m) (forearm, P_1) (n) (forearm, P_2) (o) (forearm, N_1) (p) (forearm, N_2)

(q) (forearm, N_3) (r) (forearm, N_4)

Figure 7.1: Box plots of model order estimates $f\left(d_l \middle| Y_l\right)$ for each ROI and considered model. The lower and upper line of the box are the 25% and 75% "percentiles" of the "samples" $f\left(d_l \middle| Y_l\right)$, line in the middle is median and "whiskers" above and bellow show the whole range. Each graph refers to one ROI and one considered model (ROI, model).

Let's notice that the orders $d_l = 1, 2, 3$ have the major role and the influence of higher orders decreases. These results indicate that the selected range of model orders is sufficiently great. The model order $d_l = 1$ is "used" in the estimate on the distant ROI of axilla only in the sporadic cases while its significance increases for the nearer ROIs from the injection site. Finally also notice that for each considered model moving through the ROIs with greater distance from the injection the major role moves to higher orders, even no such information were included in the prior information and no relations among ROIs were defined. It indicates the meaningful results (goal (iv) from the list of goals at page 65) and verifies the usage of "compartmental" model to be reasonable even if the compartments are not bound up with regions.

The results of maximum likelihood estimate of the common collective range of model orders do not allow us to narrow down the initially selected range. Mostly the estimates coincide with the initial range chosen according to the supposed time activity curve shape. The reason is that even the range of model orders could be narrowed down on the great part of limbs there exist "exceptions" where it is not possible. Therefore the initial setting of the range of model orders remained for a further processing.

The similar approach of the refinement could be applied also on further collective hyper-parameters, e.g. bounds of the maximum gain of the response $(\underline{r}_{\max}, \overline{r}_{\max})$ and bounds of its time $(\underline{t}_{\max}, \overline{t}_{\max})$ from the Table 4.1, however it was not employed here. The demonstrated refinement of the range of model order serves only as the illustration of further possibility how to employ Bayesian inference. The bounds are standardly considered fixed herein. When we had done some refinement (especially of the maximum gain of the response) the rough setting came out from the empirical knowledge of physicians.

7.3 Illustrative Examples

Most of the practical experiments have comprised computation of results of various tasks solved on individual limbs or their pairs. As it is not possible to show all extensive results, they are illustrated here by selected examples on the specific limbs and ROIs. Their aim is to demonstrate that the results are meaningful (goal (iv) from the list of goals at page 65) and to compare the results for individual models. In order to be objective, both poor and good results are shown.

Results Two examples of the results on the selected patients and ROIs are given here. In the first example the data comes from the ROI of axilla. There is evident difference between both limbs of patient. The results belong to better ones. The second example concerns the ROI of upper arm with similar data on both limbs whose results present the worse from the group of results. The presented outcomes are time activity curve estimate, output reconstruction, estimate of two promising quantitative parameters and comparison of limbs for all the considered models from the Table 4.2. Their discussion is the part of the next paragraph.

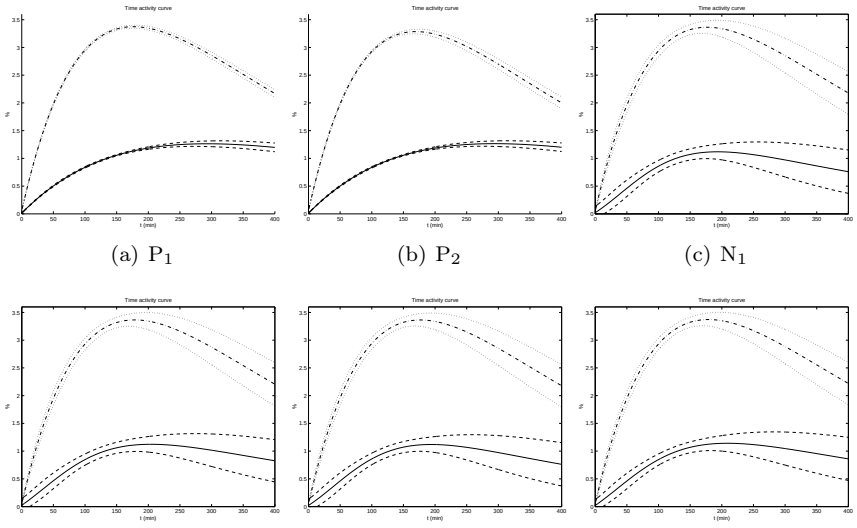
Example 1: Illustrative results on one axilla ROI

Time Activity Curve (TAC) — see Figure 7.2.

Output Reconstruction at Non-sampled Times — see Figure 7.3.

Residence time and time activity curve (TAC) maximum — the resulting estimates from both limbs are summarized in the Table 7.2.

Comparison of accumulation on both limbs — the posterior probability $f\,(h = 0|\, Y_1, Y_2)$, (6.19), of zero hypothesis in the case of chosen uniform distribution on alternative hypothesis is almost zero (1e-29).



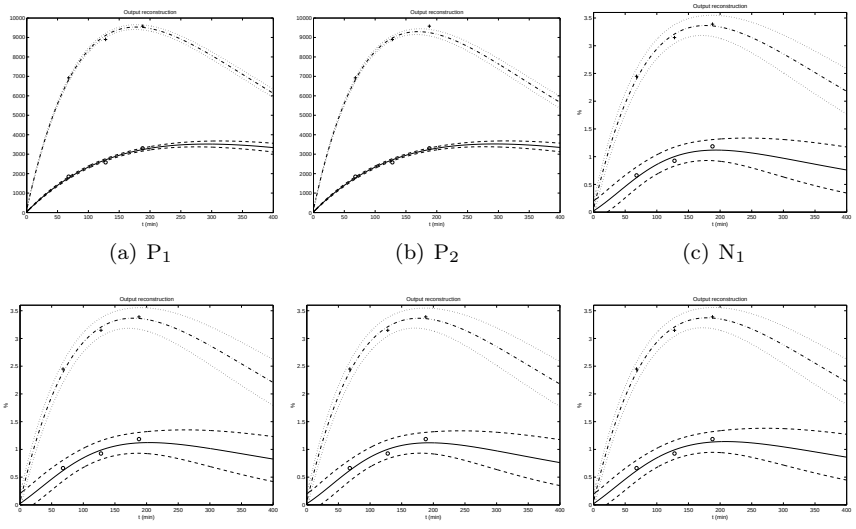
(a) P_1 (b) P_2 (c) N_1

(d) N_2 (e) N_3 (f) N_4

Figure 7.2: (Example 1) Time activity curve estimates on both limbs for all the considered models. The solid line represents the expected time activity curve $E[X_l|Y_l]$ on the left limb. Dashed lines represent its uncertainty (mean value±marginal standard deviation). The dash-dotted line and two dotted lines are used for the same results on the right limb. Y-axes give % values of relative normalized estimate.

	Residence time				TAC maximum			
	L		R		L		R	
Model	E	σ	E	σ	E	σ	E	σ
P_1	10.17	1.134	16.314	0.444	1.27e-2	5e-4	3.37e-2	3e-4
P_2	10.21	1.130	15.251	0.676	1.27e-2	5e-4	3.29e-2	5e-4
N_1	6.35	3.53	16.55	2.48	1.18e-2	1.6e-3	3.39e-2	1.2e-3
N_2	6.93	3.73	16.74	2.57	1.19e-2	1.6e-3	3.39e-2	1.3e-3
N_3	6.36	3.51	16.55	2.46	1.18e-2	1.6e-3	3.39e-2	1.2e-3
N_4	7.16	3.71	16.82	2.54	1.21e-2	1.6e-3	3.39e-2	1.3e-3

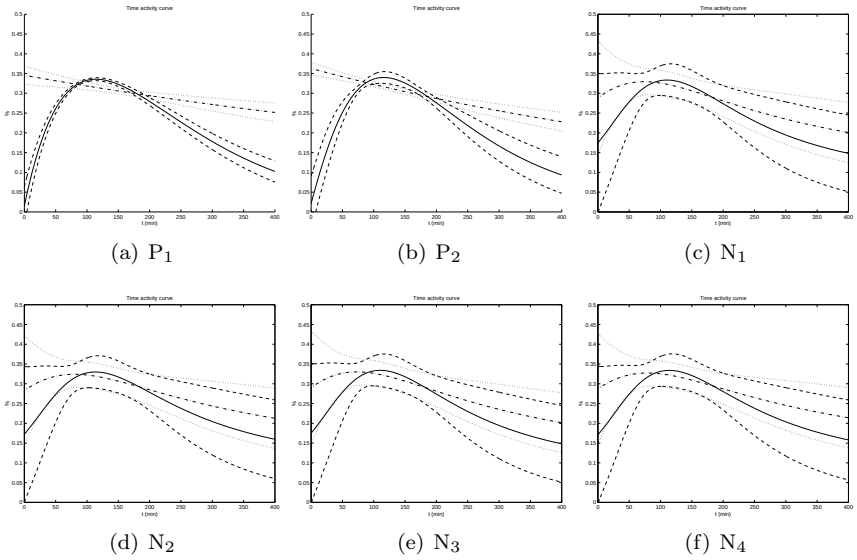
Table 7.2: (Example 1) Residence time and maximum of time activity curve (TAC) estimates for all the considered models. The mean value E and standard deviation σ for both left (L) and right (R) limbs are presented.



(a) P_1 (b) P_2 (c) N_1

(d) N_2 (e) N_3 (f) N_4

Figure 7.3: (Example 1) Output reconstruction at non-sampled times for all the considered models. The circles and plus signs denote the measured data on the left/right limb. The solid line represents the mean of the output reconstruction on the left limb. Dashed lines represent its uncertainty (mean value±marginal standard deviation). The dash-dotted line and two dotted lines are used for the same results on the right limb. Y-axes give the absolute values on the raw data scale for the models $P_\bullet$ whereas % values on the normalized data scale for the models $N_\bullet$.



(a) P_1 (b) P_2 (c) N_1

(d) N_2 (e) N_3 (f) N_4

Figure 7.4: (Example 2) Time activity curve estimates on both limbs for all the considered models. The graphical conventions are those of Figure 7.2.

Example 2: Illustrative estimate on the upper arm ROI

Time Activity Curve (TAC) — see Figure 7.4.

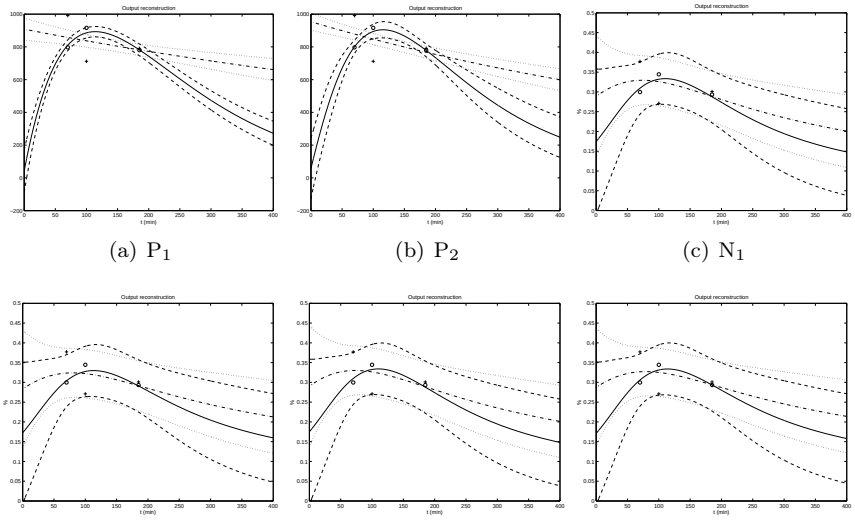
Output Reconstruction at Non-sampled Times — see Figure 7.5.

Residence time and time activity curve (TAC) maximum — the resulting estimates from both limbs are summarized in the Table 7.3.

Comparison of accumulation on both limbs — the posterior probability $f\,(h = 0|\,Y_1, Y_2)$, (6.19), of zero hypothesis in the case of chosen uniform distribution on alternative hypothesis is summarized in the Table 7.4 for all the considered models.

All the possibilities of computation of this quantity for the normal models according the proposed simplification or computation approximation involving the common parameter s is given.

Discussion The "confidence" intervals of the time activity curve estimates (mean±marginal standard deviation) in the Figures 7.2, 7.4 appear to be the same (or very similar) in the groups of normal models $N_\bullet$ and Poisson models $P_\bullet$. But again similarly to the results in the Section 7.2.2 the results between those two group of models differ. The obvious difference in the Example 1 is only greater standard deviation for all the normal models comparing to Poisson models. There are also variances among individual models too but their are minimal at a glance. In the Example 2 even the shapes of estimated time activity curves differ. Poisson models can seem more suitable from this viewpoint as they give the narrower estimate. However the output reconstruction in non-sampled times shown in the Figures 7.3, 7.5 casts another sights on the results. The tight estimates for Poisson models cause that the original employed data lie outside the "confidence interval" for output reconstruction similar to those for time activity curve estimate. It is especially noticeable in the Example 2. The normal models are more "tolerant", they cover majority of data



(a) P_1 (b) P_2 (c) N_1

(d) N_2 (e) N_3 (f) N_4

Figure 7.5: (Example 2) Output reconstruction at non-sampled times for all the considered models. The circles and plus signs denote the measured data on the left/right limb. The graphical conventions are those of Figure 7.3.

	Residence time				TAC maximum			
	L		R		L		R	
Model	E	σ	E	σ	E	σ	E	σ
P_1	1.161	0.726	4.998	1.68	3.4e-3	1.37e-5	3.5e-3	2.29e-4
P_2	1.236	1.172	3.466	1.39	3.4e-3	1.38e-4	3.6e-3	1.59e-4
N_1	2.386	1.9	3.238	1.867	3.6e-3	3.91e-4	3.6e-3	3.76e-4
N_2	2.802	2.285	3.866	2.215	3.5e-3	3.95e-4	3.5e-3	3.7e-4
N_3	2.379	1.897	3.235	1.853	3.6e-3	3.93e-4	3.6e-3	3.74e-4
N_4	2.752	2.267	3.849	2.203	3.6e-3	3.96e-4	3.5e-3	3.7e-4

Table 7.3: (Example 2) Residence time and maximum of time activity curve (TAC) estimates for all the considered models. The mean value E and standard deviation σ for both left (L) and right (R) limbs are presented.

Model	$f(h=0\mid Y_1,Y_2)$		
	Limb local parameters	Approx. via discrepancy	Approx. via moments
P_1	0.939	—	—
P_2	0.866	—	—
N_1	0.9902	0.9903	0.9903
N_2	0.9705	0.9706	0.9706
N_3	0.9996	0.9996	0.9996
N_4	0.9996	0.9996	0.9996

Table 7.4: (Example 2) The posterior probability of the hypothesis, that the accumulation characteristics on both limbs are the same. The results for all the considered models are listed. For the normal models $N_\bullet$, the values in the first column are obtained employing the simplification that the parameter s is limb local while next two columns correspond to the proposed approximation of computation of posterior distribution of common s.

by the estimated confidence intervals.

The estimates of residence time and maximum of time activity curve are given here as examples of proposed quantitative parameters. They result from the estimated time activity curve, therefore the similar statements could be done also here. Besides further interesting things are revealed here. The estimated mean value of the maximum of time activity curve is very similar for all the considered models even the whole time activity curves differ (especially in the Example 2). It can be understood that the maximum is (in time) local quantifier of the time activity curve, whose corresponding part is estimated similarly for all the models. In contrast the residence time is global quantifier where the small difference in the estimate of model parameters and consequently of time activity curve manifests itself distinctly to it. This is the reason why the estimates of residence time so differ for individual models. Its second consequence is that the relative uncertainty of estimate of global quantifiers is greater than for the local quantifiers. Finally let's notice that the estimates of residence time for pairs of odd and even normal models are similar while they differ each other.

The result of comparison of accumulation on both limbs in the Example 1 is obvious immediately from the visual evaluation of the estimated time activity curves in the Figure 7.2. The difference in estimated time activity curves is confirmed by the "zero" posterior probability of hypothesis that the accumulation characteristics on both limbs are the same $f(h=0\mid Y_1,Y_2)$. The results for all the considered models and variants of its computation for the normal models is ≈1e-29. The posterior probabilities from the Example 2 are given in the Table 7.4. The first column corresponds to the case, where all the parameters are limb local. It is the case of Poisson models $P_\bullet$. For the normal models $N_\bullet$ the values are obtained by the proposed simplification that the parameter s is limb local too. The next two columns correspond to the case where parameter s is still considered to be common for all (here for both) limbs and two proposed approximations of its posterior distribution are used. Employing this approximation the results depend in general on the order of processing of limbs. However the results for both sequences left/right and right/left limbs are the "same" over here and almost the same as the result employing the simplification about common s. The differences in the results for individual models manifest also distinctly in the results of limbs comparison while its manner of computation for the individual models has small influence on them. The greater values for the models N_3, N_4 comparing to models N_1, N_2 can be explained by the introduced greater uncertainty through the flat prior distribution $f(b_l\mid \Xi_l,s)=\mathrm{N}(b_l\mid \hat{b}_{\Xi_l},\omega_{\Xi_l}s)$ for models N_3, N_4. The partial prediction $f(Y_1\mid h=0)$ on

the first limb is greater for the models N_3, N_4 but it is cancelled by the same term for the alternative hypothesis. The significant role has the posterior parametric distribution processing one limb $f(\Xi, b, s|\, Y_1, h = 0)$. It is tighter for the models N_1, N_2 because of more precise introduced prior distribution. Since the data on both limbs are not exactly same, the prediction for the shifted data on the second limb, $f(Y_2|\, Y_1, h = 0) = \int_{\Xi^*, b^*, s^*} f(Y_2|\, \Xi, b, s)\, f(\Xi, b, s|\, Y_1, h = 0)\, d\Xi\, db\, ds$, is smaller employing tighter parameter estimate from the first limb for the models N_1, N_2.

The illustrative examples have demonstrated that some results for all the considered models are similar but in many tasks the results differ more or less. Therefore it is necessary to select the "best" model from the proposed list of models and evaluate all the partial results. Then the results for this model have to be considered as the consequence of model selection. A few results in the given illustrative examples indicate that the normal models $N_\bullet$ are better for the description of the problem than the Poisson models $P_\bullet$. The proper selection of the model solves the next Section 7.4.

7.4 Choice among Alternative Models

Up to now the results for all the considered models have been presented. However, only a single model can be employed for the practical application. Therefore it is necessary to select the best one among alternative models. The Section 7.3 has shown that the results differ partially for individual models which asks for the choice.

In the previous Section there was an attempt to make a choice among the models, that was based on a simple comparison of the results for all the models. However its only outcome was the preference of normal models towards the Poisson models with no further specification. Fortunately the task of model selection is again special decision problem which can be solved using the Bayesian approach. The solution of the model choice from the range of possible models is described well in [6]. The task of hypothesis testing can be applied on, see Remark(s) iii at page 23. Let's recall that to be able to make decision from the Bayesian viewpoint it is necessary to have probabilistic model $f(m_i|\, Y)$ for all the considered models m_i and corresponding loss function. The needed model can be obtained simply using the Bayes rule (3.1):

$$f(m_i|\, Y) \propto f(Y|\, m_i)\, f(m_i)\,,$$

where $f(Y|\, m_i) = f\left(Y_1, Y_2, \ldots, Y_{\hat{l}}|\, m_i\right)$ is the prediction (5.18), (5.21) already evaluated for individual models in the Section 5.2.

Results In order to evaluate the posterior distribution $f(m_i|\, Y)$ essential for the Bayesian decision, it is necessary to determine prior distribution $f(m_i)$ on the space of models m_i besides the predictions $\log f(Y|\, m_i)$. Since there is no reason to a priori prefer some model(s) the prior distribution on the space of models has been chosen uniform. The obtained results are summarized for all ROIs by the pairs $\log f(Y|\, m_i)$ and $f(m_i|\, Y)$ in the Table 7.5. The posterior distributions $f(m_i|\, Y)$ on the space of models for all ROIs are peaked enough with a single maximum so the choice of the model is straightforward. It reduces to the maximum of $f(m_i|\, Y)$ no matter of the selected (reasonably smooth) loss function. The model N_2 is then selected as the best one for the ROIs of forearm and upper arm while model N_1 is preferred on the ROI of axilla. If only one common model would be necessary to select for all ROIs then the model N_2 is the result of global optimization over all ROIs. From the perspective of lymphedema evaluation, the ROI on axilla does not appear to be suitable due to the unilaterally removed axillary nodes (see Sections 1.2, 2.1.1). Therefore the decision has simplified even more as the optimal results on remaining ROIs are the same.

Model	ROI=axilla		ROI=upper arm		ROI=forearm	
m_i	$\log f(Y\|m_i)$	$f(m_i\|Y)$	$\log f(Y\|m_i)$	$f(m_i\|Y)$	$\log f(Y\|m_i)$	$f(m_i\|Y)$
P_1	-1.86e3	0	-1.33e3	0	-1.38e3	0
P_2	-1.92e3	0	-1.31e3	0	-1.38e3	0
N_1	-9.82e1	$\boxed{1}$	-1.4e1	0	-1.79e1	0
N_2	-1.23e2	0	0	$\boxed{1}$	-2e-1	$\boxed{1}$
N_3	-1.99e2	0	-1.14e2	0	-1.18e2	0
N_4	-1.97e2	0	-1.12e2	0	-1.14e2	0

Table 7.5: Values $\propto \log f(Y|m_i)$ and $f(m_i|Y)$ necessary for the selection of model from the list of considered models on individual ROIs. The $\log f(Y|m_i)$ is normalized according its maximum over all ROIs. It is given here for the illustration since this information is hidden in the peak posterior distributions $f(m_i|Y)$. Maximums of the posterior distributions $f(m_i|Y)$ on individual ROIs in the frames determine the optimal selection of model.

In order to verify the results, some auxiliary quantifiers are evaluated too and compared if they are not in a contradiction with the results of Bayesian decision. The sum of absolute deviation of the normalized data from the mean of the reconstructed output in the sampling times over all the limbs $q = \sum_{l=1}^{i} \sum_{1}^{n_l} |{}^{N}y_{l;t} - E\left[{}^{N}y_{l;t}|\,Y_l\right]|$ is the quantifier which describes the quality of the point estimate. The results for the Poisson models on the raw non-normalized data are obtained similarly, only they have to be normalized with respect to the injected amount in order to be comparable with the results for normal models. The reconstructed output is computed in the times where the data from these times are already employed for the reconstruction. It is not absolutely fair, however it serve only as the auxiliary indicator and is not used further. This quantifier is employed here for the comparison of individual models but also for the demonstration that the proposed models have far better quality compared to the the primitive constant model defined as the average of data on individual limbs. The values of the quantifier q for all the considered models and simple constant model are

Quantifier $q = \sum_{l=1}^{i} \sum_{1}^{n_l} \lvert {}^{N}y_{l;t} - E\left[{}^{N}y_{l;t}\lvert\, Y_l\right]\rvert$							
ROI$\backslash$Models	P_1	P_2	N_1	N_2	N_3	N_4	constant
axilla	9.72e-2	1e-1	8.9e-2	9.41e-2	8.89e-2	8.76e-2	6.3e-1
upper arm	3.6e-2	3.61e-2	3.77e-2	3.8e-2	3.78e-2	3.76e-2	1.18e-1
forearm	3.5e-2	3.5e-2	4.05e-2	4.25e-2	4.05e-2	4.07e-2	7.3e-2

Another useful indicator of the output reconstruction is the relative deviation of the data from the mean of reconstructed output. It can be e.g. expressed by the frequency with which the original data fall into the "confidence" intervals determined as the mean value± marginal standard deviation of the output reconstruction. It tries to depict also the quality of the uncertainty of the probabilistic estimate. While the relative frequencies range from 80% to 90% on individual ROIs similarly for all the normal models $N_\bullet$ the same frequencies for Poisson models $P_\bullet$ range from 30% to 50%. It corresponds to the results of output reconstruction for the illustrative examples in the previous Section. For all the considered models the empirical *p(d)f* of relative deviations of data from the mean of reconstructed output is symmetric round the zero mean.

Discussion The interpretation of the results of Bayesian model selection from the Table 7.5 is a hard task. In spite of it, we attempt to find and explain the reasons of these results, especially why the Poisson models $\mathrm{P}_\bullet$ come out from the comparison as the worst ones.

The explanation of the failure of Poisson models is based on the idea that the model of inner response of lymphatic system is not absolutely perfect description of the reality. Therefore the differences between the ideal and real responses exist. For the case of normal models $\mathrm{N}_\bullet$ these differences can be hidden (incorporated) into the noise of the measurement process, because the variance of the noise is determined by the separate parameter s. For the Poisson models $\mathrm{P}_\bullet$ the noise is determined only by the the signal level, that does not allow to incorporate the error of modelling. Its additional impulse gain parameter g_l is tightly related with the initial measurement so the error of modelling cannot be incorporated into the model here. It is possible perhaps to find further reasons but this one seems to be the most decisive one. Accordingly, the parameter s of the normal models does not describe only the noise of the measurement process but also inner model error.

The reason why the models N_1, N_2 are preferred against the remaining models N_3, N_4 is in the difference between those models, i.e. $f\left(b_l \middle|\, \Xi_l, s\right) = \mathrm{N}(b_l |\, \hat{b}_{\Xi_l}, \omega_{\Xi_l} s)$ for N_3, N_4. This prior distribution has to be selected flat enough, see the end of the Section 5.1, so it is also non-negligible outside the range $(\underline{b}_{\Xi_l}, \overline{b}_{\Xi_l})$. It incorporates more uncertainty to the estimate for the models N_3, N_4 then for the models N_1, N_2 where the prior distribution is zero outside the range $(\underline{b}_{\Xi_l}, \overline{b}_{\Xi_l})$. It results then in a smaller value of predictive *pdf* $f\left(Y \middle|\, m_i\right)$. Interpretation of the results for the pair of models N_1, N_2 is much more difficult. Comparing their prior distributions the second one prefers slower and smaller responses of lymphatic systems. The uniform distribution on $\log(1 - a_l)$ of model N_2 induces more flat distribution of term a_l^t which appears in observation model (4.11) comparing to uniform distribution on a_l. Then the prior distribution $f\left(b_l A_{\Xi_l}\right)$ of the mean value of the observation model $f\left({}^N Y_l \middle|\, \Theta_l\right) = \mathrm{N}({}^N Y_l |\, b_l A_{\Xi_l}, s \boldsymbol{I}_{n_l})$ is more flat. This seems to be practically better for the expressing of the adequate uncertainty about these parameters. The reason why the model N_2 was not selected on the ROI of axilla could be caused by the uniform distribution on $\log b_l$ together with its discretization. It handicaps the greater responses (common for axilla) by the assigned smaller prior probability and consequently greater discretization grid. Since the essential influence on the prediction $f\left(Y \middle|\, m_i\right)$ has an absolute difference between the data and modelling response, the greater discretization grid on b_l causes, on average, greater difference between them and consequently the smaller value of predictive *pdf* $f\left(Y \middle|\, m_i\right)$. It is not significant drawback of the model N_2.

The auxiliary quantifier q gives similar values for all the proposed models only the values for a primitive constant model are far greater. It indicates the proposed modelling is of better quality then simple averaging. Thus the proposed models are useful. It would be also possible to make selection among the proposed models according these values, however the differences (especially in the groups of normal and Poisson models) are very small with respect to the significance of this quantifier. The relative frequencies with which the original data fall into the "confidence" intervals of the output reconstruction confirm the idea why the Poisson models come out from the comparison as the worst described at the beginning of this discussion. According to it, further choice among models is not possible. The auxiliary quantifiers do not allow to make the detailed choice among models, but they are not also in a sharp contradiction with the results of Bayesian decision. Therefore the results of Bayesian decision cannot be disproved from this viewpoint.

The main outcome of this Section describing the choice of model from the set of proposed models listed in the Table 4.2 is the decision that the model N_2 is the best candidate for practical usage. Therefore from this point of text all the subsequent Sections in this Chapter include only the results employing the model N_2. If it would be necessary to speed up the computation due to the some technical reasons, then model N_4 is the best alternative.

7.5 Robustness of Estimator

The prior distribution on the space of parameters forms an important part of the whole probabilistic model but it is sometimes considered by non-Bayesians as the biggest obstacle to employ the Bayesian approach. Sometimes an effort appears to suppress its significance even by some Bayesians (non-informative priors, reference analysis, etc. [6, 73]) but it is reasonable only in the cases with a sufficient amount of data. On the other hand, a rich set of prior information is exploited in this work. It supplements sparse data available in order to make inferences important for scintigraphic evaluation. Without it the whole evaluation would be more difficult or even impossible in the critical situations. Therefore the prior information has the significant role on the final results definitely. The question arising here is a robustness of the estimator to the setting of the prior information. In spite of the importance of prior information the estimator should be sufficiently insensitive to its setting. It means the small changes in the setting of prior information should not cause (distinct) difference in the final results. The intent of this Section is to demonstrate that results are (sufficiently) insensitive to the optional knobs of algorithms that serve for setting of prior distribution (goal (iii) from the list of goals at page 65).

Results Two examples from the Section 7.3 are again used here for the demonstration of the sensitivity of the final results to the changes in the prior information. Results for four different settings of prior information are compared together. The used original setting is that defined as default in the Table A.1. The changes has been applied then on three important (group of) tuning knobs of algorithm — number of discretization samples `nA`, `nB` of parameters a_l and b_l, upper bound `gain_h` of relative response maximum $\overline{r}_{\max}$ and the supposed time interval `[t_max_l t_max_h]` of the response maximum. The setting of tuning knobs for all compared versions is

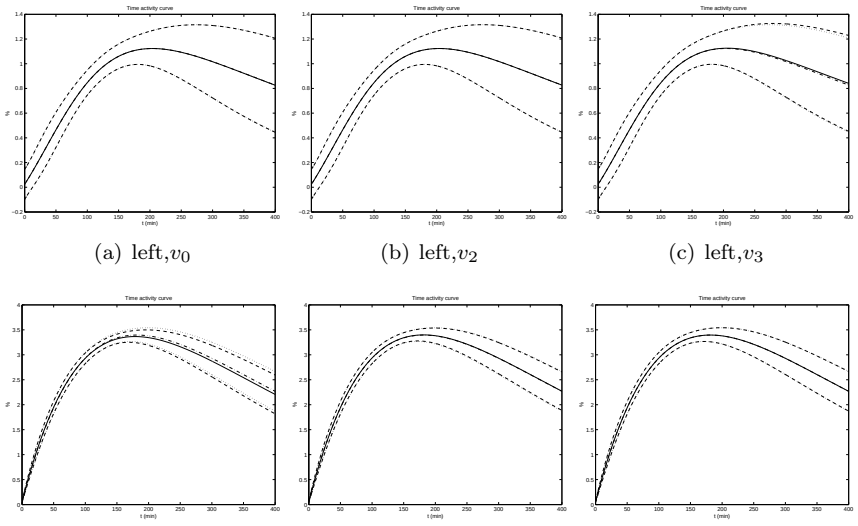
	Setting of tuning knobs of algorithm (differences)
v_0	see defaults in the Table A.1
v_1	`nA=105`, `nB=105`
v_2	`nA=105`, `nB=105`, `gain_h=1.3*gain_h` (v_0)
v_3	`nA=105`, `nB=105`, `t_max_l=t_max_l-5`, `t_max_h=t_max_h+10` (v_0)

The given results are similar to those in the Section 7.3. Estimates of time activity curve are compared in the Figures 7.6 and 7.7 for both Examples. Tables 7.6 and 7.7 collect the estimates of residence time and of response maximum, while the Table 7.8 contains the results of limb comparison.

Discussion The results presented on the examples have demonstrated that the proposed estimator is not sensitive to the small changes in the setting of important part of the prior distribution.

The difference of upper bound `gain_h` of relative response maximum in the version v_2 has almost no influence on the final results. The only case where it can cause troubles is too low selected upper bound which does not cover all responses. Since it is selected very conservatively and also the approach of refinement of the collective bounds, described in the Section 7.2.2 for the range of model orders, can be employed for the upper bound `gain_h`, this danger can be eliminated.

The setting of number of discretization samples `nA`, `nB` of parameters a_l and b_l and of the supposed time interval `[t_max_l t_max_h]` of the response maximum influences the results that is most visible on the estimates of time activity curves in the Figure 7.7 and estimates of the



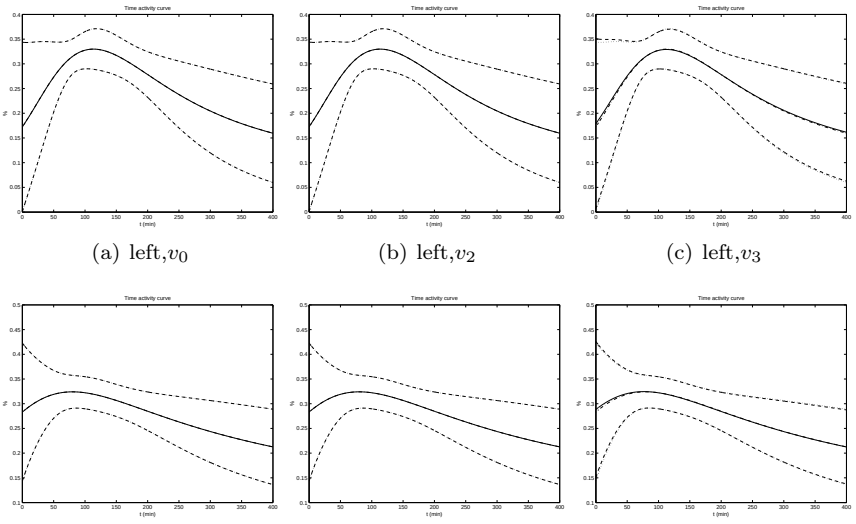
(a) left,v_0 (b) left,v_2 (c) left,v_3

(d) right,v_0 (e) right,v_2 (f) right,v_2

Figure 7.6: (Example 1) Comparison of time activity curve estimates for various settings of prior information. In all graphs the result for the version v_1 of setting is compared with the remaining versions on single limbs (left/right). The dash-dotted line represents the expected time activity curve $E[X_l|Y_l]$ for the version v_1. Dashed lines represent its uncertainty (mean value±marginal standard deviation). The solid line and two dashed lines are used for the results for the other versions $v_\bullet$. Y-axes give % values of relative normalized estimate.

Prior	Residence time				TAC maximum			
	L		R		L		R	
setting	E	σ	E	σ	E	σ	E	σ
v_0	6.94	3.74	16.74	2.57	1.19e-2	1.6e-3	3.39e-2	1.3e-3
v_1	6.95	3.74	17.15	2.59	1.19e-2	1.6e-3	3.42e-2	1.3e-3
v_2	6.95	3.75	17.16	2.51	1.19e-2	1.6e-3	3.42e-2	1.3e-3
v_3	7.12	3.91	17.15	2.59	1.19e-2	1.6e-3	3.42e-2	1.3e-3

Table 7.6: (Example 1) Residence time and maximum of time activity curve (TAC) estimates for all the versions $v_\bullet$ of prior information. The mean value E and standard deviation σ for both left (L) and right (R) limbs are presented.



(a) left,v_0 (b) left,v_2 (c) left,v_3

(d) right,v_0 (e) right,v_2 (f) right,v_2

Figure 7.7: (Example 2) Comparison of time activity curve estimates for various settings of prior information. In all graphs the result for the version v_1 of setting is compared with the remaining versions on single limbs (left/right). The graphical conventions are those of Figure 7.7

Prior	Residence time				TAC maximum			
	L		R		L		R	
setting	E	σ	E	σ	E	σ	E	σ
v_0	2.802	2.285	3.866	2.22	3.5e-3	3.95e-4	3.5e-3	3.7e-4
v_1	2.805	2.289	3.869	2.22	3.5e-3	3.95e-4	3.5e-3	3.7e-4
v_2	2.806	2.289	3.869	2.22	3.5e-3	3.95e-4	3.5e-3	3.7e-4
v_3	2.859	2.347	3.872	2.28	3.5e-3	3.97e-4	3.5e-3	3.8e-4

Table 7.7: (Example 2) Residence time and maximum of time activity curve (TAC) estimates for all the versions $v_\bullet$ of prior information. The mean value E and standard deviation σ for both left (L) and right (R) limbs are presented.

Prior	$f\,(h=0\,\vert\,Y_1,Y_2)$	
setting	Example 1	Example 2
v_0	8.6e-30	0.9706
v_1	8.3e-30	0.9705
v_2	8.5e-30	0.9706
v_3	9.1e-30	0.9716

Table 7.8: The posterior probabilities of the hypothesis, that the accumulation characteristics on both limbs are the same. The results for all the versions $v_\bullet$ of prior information and for both Examples are listed. The comparative test is computed employing the simplification that the noise precision parameter s is limb local.

residence time in Tables 7.6, 7.7. The remaining results are again almost the same. The idea for the global quantifier described in the Section 7.3 can be employed also here for the explanation why the difference is the most visible on the estimates of residence time.

The numbers of discretization samples `nA`, `nB` are not the pure hyper-parameters that determine the prior distribution. The discretization only serve for the replacement of the continuous uniform distributions in order to make the whole computation feasible. The discretization with greater number of grid points is better in general, but it is not possible to increase it to almost "infinity" as it influences the computational time. For instance, the computational time for the version v_0 of prior information is double of the computational time of version v_1 for all tasks. Therefore it is necessary to select such compromise when the estimate is not sensitive to the changes of sampling grid of discretization and the computational time is still reasonable. The proposed solution how to determine the number of grid points `nA`, `nB` employed here is to compare the various results for the used deterministic uniform discretization with the results employing the *sampling-importance-resampling* computational technique [76, 78, 80]. It works similarly but it employs the random discretization. The default values `nA`, `nB` in Table A.1 were selected then so as the results for both approaches were same.

The supposed time interval `[t_max_l t_max_h]` of the response maximum, i.e. $(\underline{t}_{\max}, \overline{t}_{\max})$ in Table 4.1, is the most important part of the prior information that reduces the space of suitable models to those with slow dynamics. Therefore it has influence on the results. Its aim is to eliminate those models that could be wrongly estimated employing only few measurements. Fortunately the sensitivity of the results to the small changes in its setting is also small.

There are two remaining tuning knobs of the algorithm that can be tested for the sensitivity of the algorithm. However it is not done here. The reason is that the distribution of the precision parameter s is learned by processing of sets of ROI data, see Section 7.2.1, and the range of model orders $(\underline{d}, \overline{d})$ is refined by processing of data in the Section 7.2.2. Over against, the former was determined only on the base of the empirical knowledge.

The sensitivity was demonstrated here on the examples with 3 measurements. It will be probably greater in the case of only 2 measurements, but decisive for our purpose is the behaviour of algorithm for 3 measurements, see Section 7.7.

7.6 Comparative Test of Limbs vs. Assessment of Physicians

Comparison between the responses of a particular patient's upper limbs is a very useful diagnostic aid. The proposed solution described in the Section 6.5 allows to do it even from the sparse data.

Without staging assessment, this is the most important task of the scintigraphic evaluation of the lymphedema. The outcome of this comparative test can be compared here with the independent decisions of clinicians and nuclear medicine experts at disposal. It allows us to demonstrate that the proposed quantification is working but, first and foremost that results are meaningful and can give the expected benefit to the physicians (the goals (i) and (iv) from the list of primary goals for this Chapter at page 65).

For this simple binary decision the quality of the diagnostic test can be evaluated by the 2x2 decision matrix (table)

Are the characteristics of both limbs on the ROI same?

Quantitative	Physicians' decision	
test	Yes	No
Yes	n_{TP}	n_{FP}
No	n_{FN}	n_{TN}

which summarizes comparison of the results of proposed comparative test with analogous decisions of physicians [18, 4]. n_{TP} and n_{TN} are number of cases (patients) where the both decisions of test and physicians are the same positive or negative, n_{FP} is number of cases where the quantitative test is positive and decision of physicians negative and n_{FN} is opposite to n_{FP}. This comparison is very popular and frequently used in medical papers [3, 19, 61]. The modification of this decision table from its analogy in [4] is in the part of clinicians' decisions. Here the decision of clinicians is again the consequence of evaluation of various observed factors that give some evidence for this decision, so this decision is not absolute too (it is not guaranteed that it is really correct). On the contrary it is considered in [4] that this part is really true. The most important measures for the evaluation of the test are the *sensitivity* and *specificity* :

$$sensitivity = \frac{n_{TP}}{n_{TP} + n_{FN}}, \quad specificity = \frac{n_{TN}}{n_{TN} + n_{FP}},$$

rates of correctly diagnosed same/different limbs of patients on individual ROI. Another important measure is the *accuracy* of the quantitative test

$$accuracy = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{TN} + n_{FP} + n_{FN}},$$

which is the rate of correctly diagnosed both the same and different limbs on individual ROI. Both high *sensitivity* and high *specificity* are indicators of efficient test.

None of these measures should be regarded as an absolute. It has been considered in the previous paragraph that the proposed test is fixed. But it is not absolutely true. The test would be fixed if also the loss function $\mathcal{Z}$ of decision task (6.16) would be determined, that is not done yet. It should be determined according the medical and economic consequences of bad selection. As it is not fixed on the base of this consideration, it represents the tuning knob of the whole decision task. Instead of the loss function, the quantity $\mathcal{P}$ determined by this loss function, see (6.19), presents the threshold of the whole comparative test that influences the final outcome of the test. It is useful then to make the comparison of results for various values of this threshold. A convenient way of representing these results is in terms of a *receiver operating characteristics* or ROC curve [4, **?**]. This curve is plot of sensitivity vs (1-specificity), with various points obtained simply by the varying threshold. It can give the helpful information about the proposed test.

Results The limbs should be theoretically compared on all the ROIs to reveal the differences on them. However it makes no sense to compare limbs on the ROI of axilla if the axillary lymphatic

Quantitative test (QD)	Clinical difference (CD) C^+	C^-
Q^+	9	0
Q^-	1	5

(a) QD vs CD

Quantitative test (QD)	Visual scint. difference (ViD) V^+	V^-
Q^+	7	1
Q^-	2	5

(b) QD vs ViD

Visual scint. difference (ViD)	Clinical difference (CD) C^+	C^-
V^+	7	1
V^-	3	4

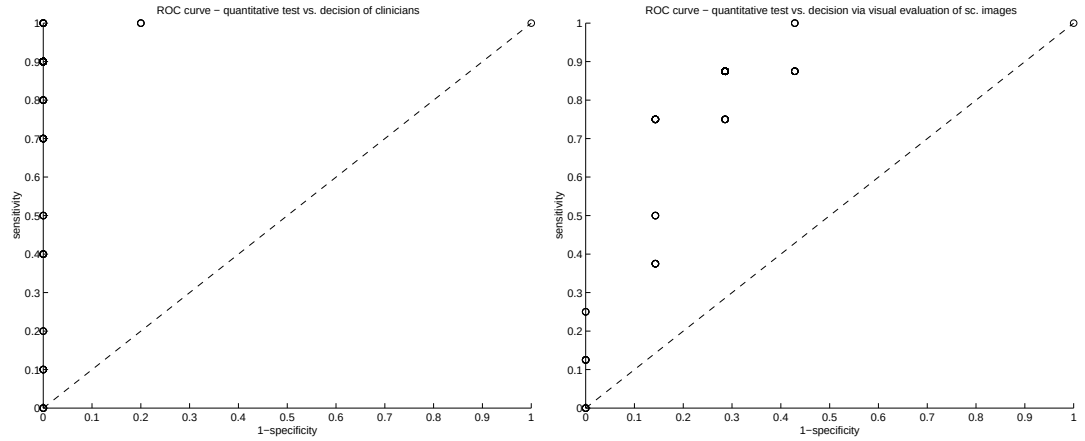
(c) ViD vs CD

Table 7.9: Comparison of decisions via the proposed Bayesian quantitative test (QD) (a) with clinical (CD) and (b) visual (ViD) decisions on forearm equality. Table (c) compares the decisions of clinicians (CD) with decisions of nuclear medicine expert about qualitative visual scintigraphic difference (ViD). C^+/C^- denotes the cases when the limbs are taken as the same/different from the CD viewpoint. V^+, V^- and Q^+, Q^- denote corresponding values for ViD and QD respectively. The number of cases belonging to the individual groups are quoted in the tables.

nodes are unilaterally removed during the breast cancer surgery. Then the local structure of lymphatic system is so different. There is reduced in some cases almost no accumulation in this ROI. Anyway only two remaining ROIs are then suitable for the comparison of limbs and lymphedema evaluation.

Table 7.9 illustrates the quality of the proposed Bayesian quantitative test on the forearms of patients in the experimental set for $\mathcal{P} = 0.8$ in (6.19). Limb equality was judged by a clinician who was treating the patients (CD decisions). An independent decision was made by an experienced nuclear medicine expert–who had access to the basic clinical patient data—via visual evaluation of the raw (non-reduced) scintigraphic images (ViD decisions). These medical decisions serve for practical evaluation of the Bayesian quantitative (QD) test (6.19). Comparing with the CD decisions, the QD decisions have sensitivity 90%, specificity 100% and accuracy 93%. Compared to the ViD decisions, the sensitivity is 78%, specificity 83% and accuracy 80%. The corresponding ROC curves for both comparisons of decisions QD vs CD and QD vs ViD on forearm equality are given in the Figure 7.8. Let's notice, there are listed single results for only one way of possible computation of quantitative test for normal model N_2 employing the simplification that the parameter s is limb local, while in the Table 7.4 the all ways are listed. The reason is that the results for all ways of computation are (almost) the same. Table 7.9(c) demonstrates the correspondence of visual scintigraphic decision ViD with the conclusion of clinicians CD. It can serve for the evaluation if the proposed quantitative test QD or the present visual scintigraphic decision ViD give the better results that coincide more with the clinical decisions CD. Comparing with the CD decisions, the ViD decisions have sensitivity 70%, specificity 80% and accuracy 73%.

The situation is more complicated on the ROI of upper arm. Some patients have the limbs with the lymphatic nodes that are located in the upper arm ROI. Unfortunately these nodes do not exist symmetrically on both limbs. It has led to the hypothesis that it makes no sense to compare the accumulation characteristics from the standpoint of the possible disease detection



(a) QD vs CD (b) QD vs ViD

Figure 7.8: ROC curves for the comparison of decisions via the proposed Bayesian quantitative test (QD) (a) with clinical (CD) and (b) visual (ViD) decisions on forearm equality. ∘ denote the points of ROC curve while the diagonal dashed line is boundary ROC curve of absolutely useless test.

Quantitative	Visual scint. difference (ViD)	
test (QD)	V^+	V^-
Q^+	7	1
Q^-	2	5

Table 7.10: Comparison of decisions via the Bayesian quantitative test (QD) with visual (ViD) decisions on *scintigraphic* equality on upper arm ROI. Q^+/Q^- denotes the cases when the aggregated scintigraphic responses are taken as the same/different by the proposed quantitative test (QD) V^+, V^- denote corresponding values for ViD. The number of cases belonging to the individual groups are quoted in the table.

as the accumulated activity in these nodes is not "eliminated" from the integral counts on the whole ROI. Consequently there is no reason to compare the results of quantitative test with a clinical decision. The only possible comparison is that of decisions via the proposed Bayesian quantitative test (QD) with visual (ViD) decisions on scintigraphic equality of upper arm ROI, see Table 7.10. Such results have to be interpreted in a different way then in the previous cases.

Discussion The results must be taken as preliminary, since the amount of available patient data is limited, and definite clinical conclusions are incomplete. It will not be possible to evaluate the quality of the Bayesian quantitative test properly until the enough data will be at disposal. The decision what is necessary amount of data to evaluate properly the test is another task suitable for Bayesian treatment.

Taking into account the discrepancies between experts, however, the high degree of correspon-

dence with the conclusions of both physicians on the forearm ROI is impressive. It is necessary to stress that the information in the aggregated data for the quantitative test is much smaller than in the raw (non-reduced) scintigraphic images employed by the nuclear medicine expert. Regardless the results are comparable even the Bayesian quantitative test is nearer to the clinical decisions, see Table 7.9. It underlines benefit of the Bayesian approach in lymphoscintigraphy. Also ROC curves bow away from the diagonal towards the upper left-hand corner which indicates the usefulness of the quantitative test. The ROC curve for the comparison of Bayesian decision QD with the clinical decision CD in the Figure 7.8(a) even reaches the absolute optimum in the left-hand corner. Though the reason is limited number of data. The ROC curve becomes more smooth for increasing number of data so it looses the tooth in the left-hand corner.

The complications with the nodes on the upper arm ROI extend the list of the open problems for the nearest future. The hypothesis that it has no sense to compare the accumulation characteristics from the standpoint of the possible disease detection and consequently the results of quantitative test with clinical decision seems reasonable but it has to be inspected further on. If it would be true then it is necessary to search the way how to make the results comparable. It means to eliminate the activity in nodes from integral counts on the ROI or to introduce further information about the existence of nodes into the process of staging assessment and limb comparison. Till this occasion we have received already aggregated data for individual ROIs, so there was no chance to cope with this problem. From the standpoint of only accumulation characteristics the results are correct. The comparison from the Table 7.10 can therefore serve for the decision if the both scintigraphic evaluations are the same.

The nature of this test is to evaluate if the accumulation characteristics on both limbs are the same or different. It is not designated for the decision if the both responses can be classified to the same stage of the disease. It belongs to the assessment of the disease staging. It is necessary to have this in mind.

7.7 Selection of Appropriate Sampling Times

The selection of appropriate sampling times is very important subtask, that should ensure that the whole proposed scintigraphy quantification will be routinely applicable. It forces us to select the sampling times very carefully for the maximal efficiency of the method. This Section presents the optimization results of the solution described in the Section 6.4 obtained on the experimental group of patients.

Results The number of data obtained on each ROI of patients' limbs in the experimental group differs from 3 to 4. Therefore the proposed optimization can be solved for the case of 2 and 3 sampling times.

The physicians' desire is to get along with 2 measurements. In this case the choice of the pair of sampling times on the model quality is very significant, as is evidenced by the series of results in Figure 7.9. There are shown the estimates of time activity curve together with output reconstruction for all the combinations of used 2 measurements. The situation is not so critical for 3 measurements, however the choice of sampling times has still significant influence on the model quality, see Figure 7.10. Optimization of the sampling time is therefore essential for both cases.

The sampling times differ from patient to patient in the experimental set. Hence, they were grouped into sets of similar values, and measurement intervals—rather than individual measurement times—were assessed as it was proposed already in the Remark(s) 6.4.2 (iii). The available intervals were

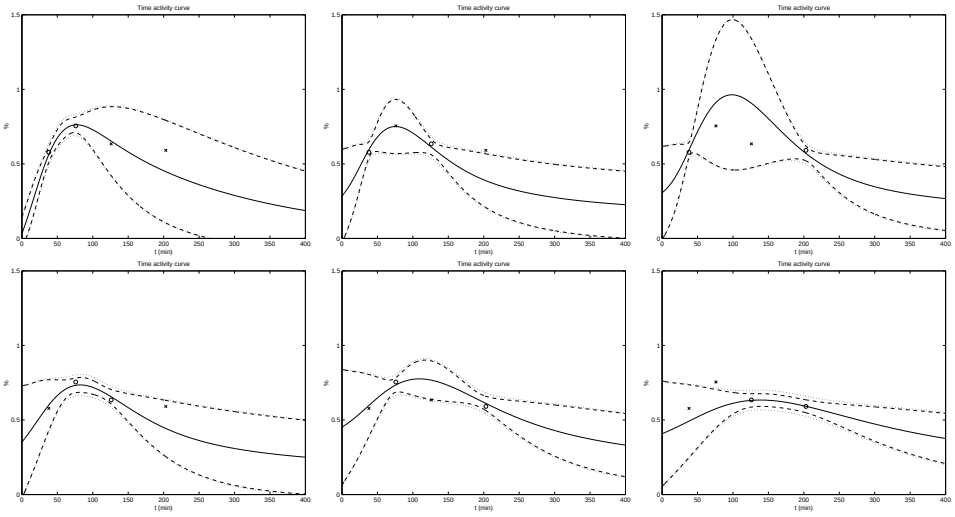


Figure 7.9: Example of time activity curve estimate from 2 measurements. ∘/× denotes data used/unused for estimation. All possible combinations are demonstrated. Solid line represents the expected time activity curves. Dashed lines represent its uncertainty (mean value± marginal standard deviation). Dotted lines represent uncertainty interval for response with measurement noise, defined in the same way. Y-axes give % values of relative normalized estimate.

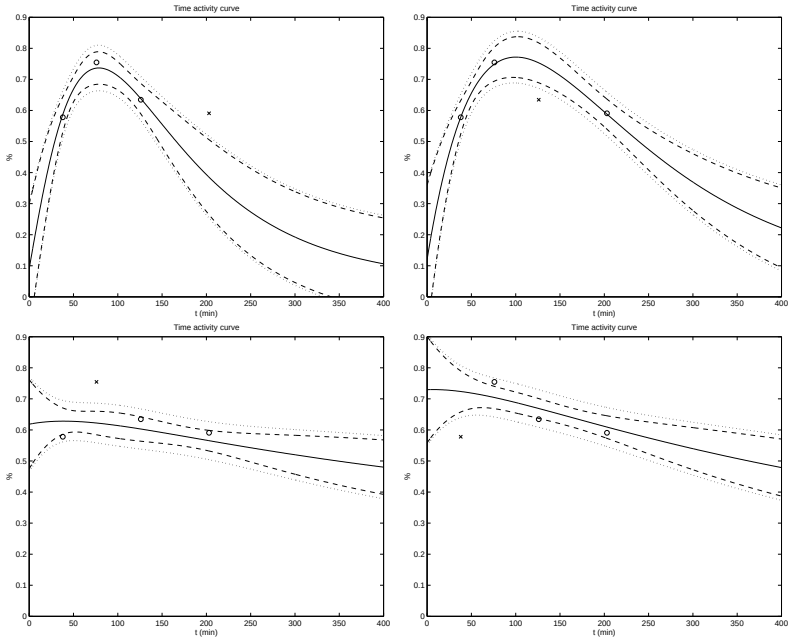


Figure 7.10: Example of time activity curve estimate from 3 measurements. ∘/× denotes data used/unused for estimation. All possible combinations are demonstrated. Graphical conventions are those of Figure 7.9.

Optimal combination of intervals $\hat{\mathcal{S}}$		
ROI	2 measurements	3 measurements
axilla	$\{\mathcal{I}_1, \mathcal{I}_3\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$
upper arm	$\{\mathcal{I}_1, \mathcal{I}_4\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$
forearm	$\{\mathcal{I}_1, \mathcal{I}_3\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$

(a) Optimization for reporting beliefs (6.10)

Optimal combination of intervals $\hat{\mathcal{S}}$		
ROI	2 measurements	3 measurements
axilla	$\{\mathcal{I}_1, \mathcal{I}_3\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$
upper arm	$\{\mathcal{I}_1, \mathcal{I}_4\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$
forearm	$\{\mathcal{I}_1, \mathcal{I}_4\}$	$\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_4\}$

(b) Former version of optimization via (7.1)

Table 7.11: Optimal choice of 2 and 3 measurement times for each ROI. Table (a) shows the results of optimization for reporting beliefs about triple (a_l, d_l, b_l), (6.10). Table (b) summarizes the results of optimization based on the predictions $f\left({}^{[\bar{\mathcal{S}}]}Y_l | {}^{[\mathcal{S}]}Y_l\right)$, (7.1).

Interval	Range of intervals (mins)
$\mathcal{I}_1$	$\langle 30, 50\rangle$
$\mathcal{I}_2$	$\langle 55, 95\rangle$
$\mathcal{I}_3$	$\langle 115, 150\rangle$
$\mathcal{I}_4$	$\langle 175, 220\rangle$

The results of optimization proposed for reporting beliefs about the triple (a_l, d_l, b_l) (parameters of interest that describe the limb local relative lymphatic system response), see Section 6.4.1, are summarized for individual ROIs in Table 7.11(a). For the comparison, also the optimization

$$\hat{\mathcal{S}} \in \arg\max_{\mathcal{S}} \sum_{l=1}^{\dot{l}} \log f\left({}^{[\bar{\mathcal{S}}]}Y_l | {}^{[\mathcal{S}]}Y_l\right) \tag{7.1}$$

that was considered intuitively from the beginning as a good indicator for the selection of sampling times, see Section 6.4.1 (6.11), is evaluated for individual ROIs in the Table 7.11(b).

In the case of global optimization done on all ROIs together the optimization results coincide for both the proposed version in Section 6.4.1 and the version (7.1). They are $\{\mathcal{I}_1, \mathcal{I}_4\}$ for two measurements and $\{\mathcal{I}_1, \mathcal{I}_2\, \mathcal{I}_4\}$ for three measurements.

Quality of the chosen model N_2 as well as of the performed optimization is confirmed by comparing the output reconstruction, Section 6.2, done from optimally placed measurements with position of real data unused for estimation. The results are summarized in Table 7.12 where the relative frequency of data in estimated "confidence" intervals determined as the mean value± marginal standard deviation of the output reconstruction are given. The average half-width of confidence intervals are listed also there.

Discussion Tables 7.11(a), 7.11(b) show the difference between the results of the proposed optimization for belief reporting (6.10) and the version (7.1) for 2 measurements on the ROI of forearm while on the other ROIs and for 3 measurements all the results coincide. The theoretical

ROI	2 measurements fr	2 measurements hw	3 measurements fr	3 measurements hw
axilla	76%	4.9e-3	64%	1.8e-3
upper arm	78%	2.5e-3	61%	0.7e-3
forearm	85%	2.6e-3	68%	1e-3

Table 7.12: Relative frequency fr of real data unused for estimation in "confidence" intervals estimated from optimally placed 2 and 3 measurements and the average half-width hw of confidence intervals.

difference between those results is evident from (6.11). Term $\log f\left({}^{[\bar{s}]}Y_l \middle| {}^{[s]}Y_l\right)$ used in the optimization (7.1) is only a part of minimized term of (6.10). It has caused little different results on individual limbs, however this difference has manifested in summary over all limbs just only on the forearm ROI. It seems reasonable that the second time from the pair of optimal times for upper arm ROI for the new version of optimization is delayed compared to forearm. It corresponds with fact that the upper arm is more distant from the injection site than forearm. Therefore it is surprising that for both versions of optimization the optimal pair of times on the axilla is not delayed even the second time foreruns times on other ROIs. The possible explanation is that the interval $\mathcal{I}_3$ is somehow important for the determination of the shape of time activity curve on the ROI of axilla. Since it is not known which ROI is the most important (if any) for evaluation of lymphedema the compromise in the choice of time intervals has to be made. The results of global optimization over all the ROIs is suitable for it. It is important and plausible that the optimization leads to the recommendations that conform with the best current practice.

The optimization for reporting beliefs (6.10) proposed in the Section 6.4.1 is preferred to the version (7.1) since the whole task of optimization is well defined in this case. Another reason is that it compares the combinations of sampling times on the common space of parameters, while in the optimization (7.1) the predictions for various times are compared. Besides in the latter case, the interpolations and extrapolations are mixed together.

Table 7.12 can serve for the demonstration that (ii) the results are (relatively) reliable. The relative frequencies of data unused for estimation in the chosen confidence intervals — mean value± marginal standard deviation — for three measurements are in the range 60–70%. It is near the equivalent confidence interval for normal Gaussian distribution ($f(\mu \pm \sigma) \approx 2/3$ for $f(\cdot) \equiv \mathrm{N}(\cdot|\,\mu, \sigma)$). The values are even greater for pair of measurements. The reason is that the estimate done only from 2 measurements is not learned enough, so it has not good quality and is still quite uncertain. Consequently the confidence intervals for the pair of measurements are much more wider than for 3 measurements, see Table 7.12 and Figures 7.9, 7.10, and the relative frequencies of data in intervals are greater. The relative frequencies of data in confidence intervals can be used as the indicator of the quality of the model together with optimization however they are not suitable for the comparison of the combinations of sampling times (the reason is the same as why not to use optimization (7.1) described in the previous paragraph).

It is necessary to repeat again that the optimization is searching among the existing sampling times of data in the experimental set. It does not solve the absolute optimization of sampling times also outside of the employed times. Therefore the result of optimization depends on the data at disposal also in this way. The optimization results are therefore the consequence of data available for the decision. For the different initial conditions the results need not be then the same.

All the gained experience leads to the following practical conclusions. Use of a pair of images together with the proposed modelling is at the border of feasibility and three images provide the best compromise between the accuracy of results and measurement demands. The Bayesian estimator is capable to give the results from 2 measurements where the other techniques already collapse. However the estimate is still quite uncertain. Permitting the first model order , $d_l = 1$, in the probabilistic model it is better to have 3 measurements to estimate properly the basic shape of time activity curve. Using only 2 measurements can cause the model of the first order is employed inappropriately.

The outcome of the optimization is the suggestion on the 3 measurements where the first should be taken after 30 minutes from the time of injection of tracer the second about after 90 minutes and the last after 3 hours.

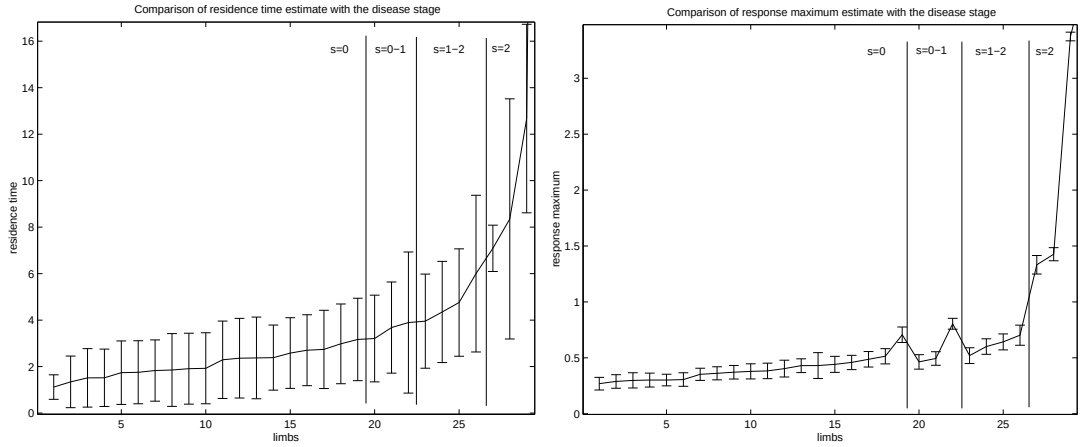
7.8 Potential Model Employment for the Disease Staging Assessment

The disease staging assessment is the only task that is still missing to achieve final goal of quantification. Staging estimate is the most detailed information that is desired to be obtained as the outcome of the whole processing. The proposed modelling that forms the main part of the thesis was considered from the very beginning as the aid of scintigraphy quantification so we hope it can be used also for the lymphedema staging assessment.

There were proposed some promising quantitative parameters in the Section 6.3 that should be tested with respect to the suitability of the disease assessment. It means to test various parameters for the correlation with the disease staging. It is also possible to search for further quantifiers outside the proposed set and combine them together. However all it faces the problem of limited amount of available patient data. Therefore no final conclusion can be done here. Instead of it, only results are given that illustrate the potential of two proposed quantifiers.

Results The open problem with existing lymphatic nodes on the ROI of upper arm that concern also the disease staging assessment was already described in the Section 7.6. Therefore the results for the forearm ROI are only demonstrated here. The Figure 7.11 tries to relate the (a) estimate of the residence time and (b) estimate of response (time activity curve) maximum with the lymphedema staging, see Section 2.1.3. The employed information about the staging comes from the decision of the clinicians. The limbs were divided into the groups according the stage and sorted in these groups according the expected value of the quantifier. Then the quantifier estimates (mean value±standard deviation) for individual limbs in the groups are plotted. The Figure 7.12 demonstrates the same relation for the couple of quantifiers (residence time, TAC maximum) in the scatter plot where the circles denote their point estimates on individual limbs while their radius depends on the lymphedema stage. It would be possible to give here results for other quantifiers, however the work does not solve their selection. Thus the demonstrated results serve only as a motivation for their search.

Discussion The Figures 7.11, 7.12 indicate that there exists some dependence between the disease staging and two proposed quantifiers. The residence time seems from these plots to be the better candidate as a suitable quantifier then TAC maximum. However it will be probably necessary to supplement it with another quantifier. The selection of the suitable quantifier(s) from the list of proposed quantifiers could be solved employing again within the framework of Bayesian decision theory. Since it is not possible to realize this search now, it belongs still to open problems and the most important practical result for the physicians is the comparative test of limbs.



(a) Estimate of residence time vs clinical staging (b) Estimate of TAC maximum vs clinical staging

Figure 7.11: Comparison of the estimate of (a) residence time and (b) TAC maximum with the clinicians' decisions about the lymphedema staging on the forearm. The limbs in the experimental set are divided into the groups according the stage marked by plumb lines together with legend in the plot. The sorted estimates (mean value±standard deviation) of quantifiers for individual limbs in the groups are then plotted.

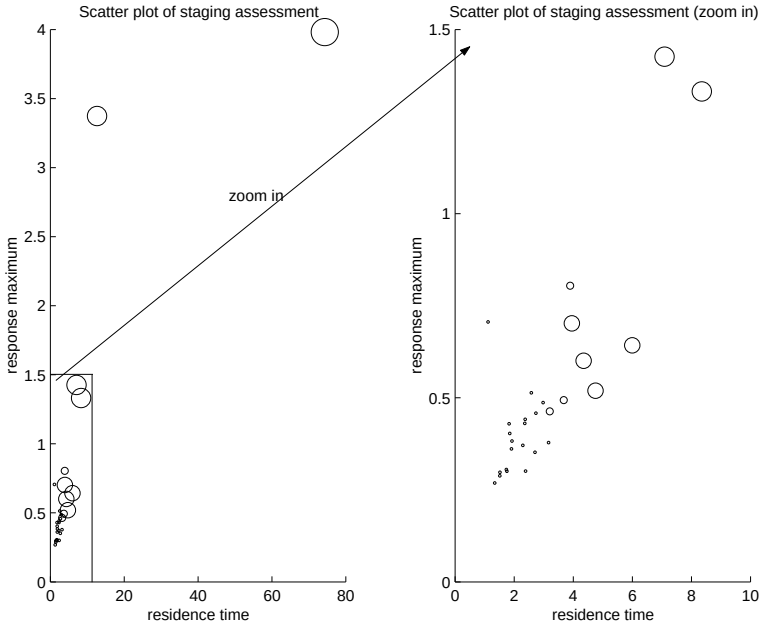


Figure 7.12: Scatter plot comparing the point estimate of the pair of quantifiers (residence time, TAC maximum) with the clinicians' decisions about the lymphedema staging on the forearm. ∘ denote the point estimates of pair of quantifiers on individual limbs while the staging is denoted by its diameter. The higher is the stage the greater is diameter.

7.9 Summary

The methodology described in the previous Chapters has been experimentally evaluated. The important outcomes from the whole set of experiments are emphasized in the following points:

- The normal model N_2 was selected as the best alternative from the set of considered models for the description of scintigraphic response of the lymphatic system in the Section 7.4. The experiments have confirmed that the idea of the compartmental model (even the compartments are not bound up with physical ROIs) are adequate for the modelling of the scintigraphic response on the ROIs even having only few data. The normal models are preferred against the Poisson models as they have more freedom to respect the error of "ideal" modelling of the inner deterministic response.

- The proposed estimator is robust. Small changes in the setting of the tuning knobs of algorithms expressing mainly the prior information cause minimal difference of results, see Section 7.5. Besides the both approximations of the computation of the distribution of the common parameter s for normal models proposed in Section 5.2.2 at page 47 give very similar acceptable results for individual normal models, see Section 7.2.1. Thus, none of the approximations is preferred here. The results of the quantitative comparisons of the limbs demonstrated in the Section 7.3 have shown that there is a negligible difference in the results among its various computation, employing various computation approximations and simplification about the common parameter s. Therefore the simplest version of computation based on the simplification that parameter s is limb local can be employed, despite of the fact the other versions are theoretically more immaculate.

- The common model for all three ROIs, that differ only in the setting of prior distribution, proved to be suitable for the description of the scintigraphy response on all ROIs. Since it is not known yet if some ROI is dominant for the disease assessment the estimates of time activity curves are done on each ROI. It has an added value as it provides the possibility to determine the position of lymph tapering or blocking. It underlines significance of the obtained estimates of time activity curves.

  From the perspective of disease assessment the ROI of axilla appears to be improper as the axillary lymphatic nodes are unilaterally removed during the breast cancer surgery. Anyway two remaining ROIs are suitable for the comparison of limbs and lymphedema evaluation. For the completeness, it is necessary to resolve the problem how to treat the data of specific limbs having also the lymphatic nodes on the upper arm ROI to make them comparable with other limbs.

- The expected sensitivity to measurement times was confirmed in the Section 7.7 and the optimization led to the recommendations that conform with the best current practice. Three images were selected for the employment together with the proposed model as the best compromise between the accuracy of results and measurement demands. In spite of the fact the Bayesian estimator can give results using only pair of images it is not enough to estimate properly the basic shape of time activity curve. The selected intervals employing the proposed optimization then resulted in the conclusion that is not in contradiction with the current practice.

- The results of the proposed quantitative test of the comparison of limbs on the limited amount of available patient data for the ROI of forearm have shown the high degree of correspondence with the conclusions of both physician and nuclear medicine expert. Despite the amount of information in the aggregated data for the quantitative test is much smaller

than in the raw (non-reduced) scintigraphic images employed by the nuclear medicine expert, the results of quantitative test and conclusions of physicians are comparable. Even the proposed quantitative test compared to visual qualitative scintigraphic evaluation gives better results that are closer to decisions of clinicians. Evaluating the quality of the quantitative test it has promising sensitivity and specificity. Also the rough drafts of ROC curves have the good shape.

The comparison of limbs on the ROI of upper arm faces the previously mentioned problem of the lymphatic nodes on this ROI. If the nodes exist symmetrically on both limbs the comparison is possible. The sense of alone comparison without incorporating the information about existence of nodes or elimination of the activity in them for the case of unilateral nodes has to be inspected further on.

- The quality of the proposed model together with Bayesian processing was verified from several viewpoints. All the previous items can be included among them. The fitting of used data on the output reconstruction and also the comparison of unused data with the output reconstruction from the data in optimal times belong to those more direct verification. However also the correspondence of comparative tests with the decisions of physicians (and another results) indicate the quality of model at least indirectly. At present, it remains to find the suitable quantifier for the lymphedema assessment on the base of the proposed model to achieve the final aim.

Chapter 8

Conclusions

The quantitative lymphoscintigraphy is the most promising technique at present, how to evaluate the state of the lymphatic system in the limbs and to diagnose even the latent stage of the lymphedema disease. The present work is the core of the project that aims to develop such a methodology of quantitative evaluation, missing up to now, that is reliable and routinely applicable. It means that such methodology should get along with the practically limited amount of measurements.

Thesis, that finalize my preliminary results [23, 24, 25, 26, 27], is focused on the design, implementation and verification of the methodology of scintigraphy quantification. The Bayesian processing has been selected as the viable option for this purpose. The solution is based on the simplified modelling of the accumulation dynamics of radiotracer on the pre-specified regions of interest of limb. It comes out from the presumption that the dynamic properties on the whole limbs are important for the lymphedema evaluation.

The results of the thesis indicate that the aims of the work were successfully achieved. The main result is the design and verification of a completely new methodology of quantitative scintigraphy. Its outcomes are in a good correspondence with the conclusions of physicians. It is necessary precondition to employ the proposed inspection for the recognition of the lymphedema already in its latent phase and evaluation of therapy effects.

8.1 Summary of Contributions

The whole thesis presents interdisciplinary work that tries to employ Bayesian theory for the real problem of scintigraphy. Consequently, main contributions of the work are determined by it. The contribution to the quantitative scintigraphic evaluation dominates but the application of the general Bayesian theory on the real problem is significant too.

In summary of the most important general contributions, the following points have to be listed:

- The real problem of quantitative scintigraphy is resolved in an integral and consistent way from the initial formulation, through the modelling, algorithmic and numerical solutions up to successful verification on the real data.
- The strength of the Bayesian theory is emphasized by the comprehensiveness of the solution and the rate of its systematic utilization. It serves also for the demonstration, that the Bayesian processing is suitable especially in inference problems based on few measurements only.

- Suitable class of models describing the scintigraphic response on the ROIs of limb for the limited amount of measurements is adopted and finally the best one selected. They distinguish between the local and common limb parameters. The prior distribution on parameters employing the expert knowledge serves for the reduction of the space of a priori suitable models.
- Non-standard Bayesian methodology of estimation respecting both the common and limb local parameters is elaborated.
- The approach how to solve particular non-standard (decision) problems within the Bayesian framework is proposed and verified.
- A practical diagnostic problem with a significant impact on the disease treatment is resolved. It opens the way for its further improvement.

In order to make the summary complete the accomplishment of partial objectives of the thesis is described in details.

Modelling of accumulation dynamics of tracer in the ROIs A significant part of the work is devoted to the modelling of the accumulation dynamics. It forms the core of the proposed quantification, so it has the decisive influence on the success of the proposed solution. In this respect:

- The chosen form of model describes well the response on individual ROIs separately. The cascade structure of the model whose impulse response expresses the scintigraphic response of lymphatic system is selected as the minimalist but sufficiently flexible model for all ROIs. It has a small number of unknown parameters. The prior distribution on the space of parameters necessary for the Bayesian paradigm allowes us to reduce the space of models to a priori suitable ones and to differentiate models for individual ROIs.
- Due to the computational difficulties related to the used common parameters for all limbs (noise precision parameter s for normal models $\mathrm{N}_\bullet$), the approximations of global parametric and simplification of predictive inference computation are proposed. The solution is designed directly to the model used, but the idea can be generalized.

  Two alternative approximative estimations of common precision s for normal models are proposed. It is the parameter, whose prior distribution is set only roughly. As it is defined as the common parameter for all limbs, processing of sets of ROI data could be used for its refinement. The obtained results of both approximations on the set of real data are found very similar, so they are accepted for further processing.
- Since the various alternatives of the basic model are proposed to be suitable for the description without the preference of anyone, the choice of the best one has to be solved. Fortunately, the well known general decision problem involving model choice in Bayesian theory offered natural solution.
- The adequacy of the adopted model is verified in experiments on the set of real data.

Reconstruction of time activity curves on individual ROIs The evaluation of this quantity is the direct consequence of the model employment. The expected value and standard deviation are computed for its comprehensible and simple description. Other characteristics could be evaluated, though it would be more complicated.

Selection of appropriate sampling times The solution of this task important for scintigraphy quantification is useful, in our opinion, for all the problems with the limited amount of data. To our best knowledge this problem has not been elaborated yet within the Bayesian framework.

- The proposed solution selects the best combination of sampling times from those included in the set of available data. It selects such a combination, for which the estimate of quantity of interest employing the data in the corresponding sampling times is nearest to the estimate from all data. Three parameters describing individual accumulation characteristic are chosen as the quantity of interest for the scintigraphy problem. It is demonstrated that the best variant is to compare the whole generalized Bayesian estimates, i.e. tho whole *pdf*s, rather then their partial characteristics. Such an optimization task is applicable for arbitrary probabilistic models. It does not suffer the problems, what characteristics should be compared and how to select the proper loss function with respect to the form of *pdf*s. The presented solution improves the optimization proposed formerly in [26].
- Outcomes of the optimization on the real scintigraphic data suggest optimal sampling times that conform with the best current practice.

Comparison of patient's limbs The judging whether the limbs differ or not is realized as the Bayesian test of the hypothesis whether the local limb parameters describing individual accumulation characteristic are the same for both limbs or they differ. Using this approach:

- The posterior probabilities of the hypotheses are its temporary outcome. The final binary decision then depends on the simply interpretable loss table, whose entries penalize both possible bad decisions.
- The results on the experimental set of data showes good correspondence with the conclusions of physicians. Both sensitivity and specificity are high, even the ROC curves have the right shape.

Computation of various quantitative parameters Several quantitative parameters that are promising with respect to the disease assessment are put forward and their estimates computed. The conclusive tests of their suitability, i.e. if they are correlated with the disease staging, could not be done yet due to the limited set of patient data. Instead of it, the potential of some quantifiers is illustrated at least. The results seem to justify the selected research direction, though search of suitable quantifier for the disease assessment belongs to open problems.

Methodology verification The proposed methodology is verified from the several perspectives. The quality of the proposed modelling and the robustness of the estimator together with the solutions of decision tasks are tested on the real scintigraphic data. The outcomes of the methodology are then compared with available conclusions of physicians.

The important outcome of the thesis is the implementation of the proposed methodology. It is realized as the package of macros for *MATLAB* environment. It is at the disposal for the physicians at the Clinics of Nuclear Medicine of the Faculty Hospital Motol for a further testing.

8.2 Open problems

The very initial attempt served for the indication of the direction, how to solve the problem of quantitative scintigraphy. At present, this work comprises pivotal part of the running project that copes with this problem. Results of the thesis indicate that the proposed methodology based on the employment of the Bayesian theory can give reliable outcomes. Though, several tasks have to be still solved to achieve the final goals of the whole project. They are beyond the scope of the present work, but they are addressed within the research project supported by IGA MZ ČR. Let us mention them:

- The whole methodology is to be verified on a sufficient amount of data to make the results relevant for physicians.
- The methodology is to be completed by finding suitable quantifiers for the lymphedema assessment. It means to test various proposed quantifiers, possibly together with available clinical information, if they are correlated with the disease staging and to select the best one or their combination.
- Complications with the common case of nodes on the ROI of upper arm have to be resolved. It has to be inspected carefully if their existence has significant influence on the quantification.
- Model describing the accumulation of radiotracer can be potentially further improved. Already in the body of the work, there were indicated some alternatives to the basic model that could improve the currently selected version of model. Instead of the local models for individual ROIs the common model for all ROIs with the spatial relations could be used.

Appendix A

Implementation Aspects

The proposed estimator and the solution of decision tasks were implemented as a package of macros for *MATLAB* environment [55]. It is numerically oriented software, which provides a lot of useful tools and enables a comfortable work with them. All algorithms code directly the formulas in Chapters 4, 5, 6. Already there, the formulas reflect the solution of found computational difficulties (e.g. discretization of continuous parameters, used approximations).

Though *MATLAB* is powerful tool it is necessary to take care of specific numerical problems related to *pdf*s. The one we met is described in the Section A.1. The algorithm tailor-made for formulas in this work that prevents this numerical problem is proposed there too.

All macros are constructed in such way they use the common initialization setup containing the tuning knobs that are mainly software counterparts of the prior information which allow to change and tune-up setting of algorithms. The list of all tuning knobs for all considered models together with their defaults are given in the Section A.2.

Finally, the code of two basic macros (functions) for evaluation of limbs for the chosen model N_2 from the considered set of models, see Section 7.4, is given. The first one computes estimate of time activity curve and estimate of residence time, the second implements quantitative comparison of limbs.

A.1 Numerical Realization — Treatment of Sharp Likelihoods

The problem of numerical overflow is common for work with the probability density functions (*pdf*s) mainly encountered in parameter estimation and in testing of hypothesis. In these cases, they have tendency to converge to very sharp functions close to Dirac delta function. Their values easily exceed the range of numbers that a software is able to operate with. The common solution how to overcome this problem is to work with logarithms of *pdf*s [47]. It means to work with logarithms of likelihood functions $\log(f(Y|\Theta))$ and before converting them into posterior *pdf*s subtract $\max_{\Theta\in\Theta^*}\log(f(Y|\Theta))$. It is possible because the likelihood is determined uniquely up to a Θ independent factor. It is advantageous as it prevents numerical overflows and the dynamical range of the posterior *pdf*s causes at most numerical underflows for values for which the posterior *pdf* is anyway negligible. Besides, it transforms the Bayes' rule using the multiplication to simpler version using addition of logarithms $\log(f(\Theta|Y)) \propto \log(f(Y|\Theta)) + \log(f(\Theta))$.

In our case all is complicated due to the included discrete and discretized parameters. If we want to compute e.g. some expectation or marginalization it is not advantageous to use logarithmic version of Bayes rule since the logarithms are applied on the sums. In addition to

it, the likelihoods of interest here are the terms $f(Y_l|\,\Xi_l, b_l) \propto w_{\Xi_l,b_l}$ for models $\mathrm{P}_\bullet$, N_1, N_2, see (5.1) and (5.5), and $f(Y_l|\,\Xi_l) \propto w_{\Xi_l}$ for models N_3, N_4, see (5.10), whose maximum can not be easily found. For that reason we have proposed algorithm how to realize the computation and to prevent overflows. Its core is demonstrated in Algorithm 1 on the computation of the posterior expectation $E[q|\,Y_l]$ of some quantifier $q = q(\Xi_l)$, which is a function of Ξ_l. The algorithm holds the information about the so far found maximum of log-likelihood in the variable wL_{max} and all the likelihoods relate to it. So if the maximum of log-likelihood wL_{max} does not exceed the range of numbers in the used software the overflow can not occur. Sum of weights of w_{Ξ_l} (that is proportional to predictive inference $f(Y_l)$) can be gained from the algorithm too, $\sum_{\Xi_l^*} w_{\Xi_l} = w_\Sigma \exp(wL_{max})$ (see the Algorithm 1).

Algorithm 1 Calculate $E[q|\,Y_l]$

Require: $\Xi_l^*,\ \log w_{\Xi_l},\ q(\Xi_l)$
Ensure: $E[q|\,Y_l] = E_q$
  $w_\Sigma \Leftarrow 0$
  $wL_{max} \Leftarrow []$
  $E_q \Leftarrow 0$
  for all $\Xi_l \in \Xi_l^*$ **do**
    $wL \Leftarrow \log w_{\Xi_l}$
    if $wL_{max} = []$ **then**
      $wL_{max} \Leftarrow wL$
      $w_\Sigma \Leftarrow 1$
      $E_q \Leftarrow q(\Xi_l)$
    else if $wL > wL_{max}$ **then**
      $w_T \Leftarrow \exp(\log w_\Sigma + wL_{max} - wL)$
      $w_\Sigma \Leftarrow 1 + w_T$
      $wL_{max} \Leftarrow wL$
      $E_q \Leftarrow (E_q \times w_T + q(\Xi_l))/w_\Sigma$
    else
      $w_T \Leftarrow \exp(wL - wL_{max})$
      $E_q \Leftarrow (E_q \times w_\Sigma + w_T \times q(\Xi_l))/(w_\Sigma + w_T)$
      $w_\Sigma \Leftarrow w_\Sigma + w_T$
    end if
  end for

A.2 Tuning Knobs of Algorithms

In order to make macros relatively simple and usable by non-experts they are constructed in such way that the number of their input parameters is limited to minimum. They only need corresponding data and times of measurements. In addition to it, there is possibility to influence the behaviour of macros without the need to modify them. All the macros employ the common initialization file, which contains software counterpart of the prior information setting described in Section 4.3.3. Consequently, this file contains a set of tuning knobs, that allow not only to make experiments with various settings of prior information but also to tune-up these settings and hide them for common use. All knobs used by individual models are listed in the Table A.1. It contains their description together with relation to terms in the Section 4.3.3 and Table 4.1. The default setting is given too. The initialization file contains also information about the ROI where the data comes from. It allows a specific setting for individual ROIs.

Models	Knobs	Description	Default setting
All	`[d_l d_h]`	common hyper-parameters $(\underline{d}, \overline{d})$ for all limbs that define the range of used model orders d_l; for their tuning see Section 7.2.2	`[1 5]`
	`gain_h`	upper bound of relative response maximum $\overline{r}_{\max}$; used for computation of $\overline{b}_{\Xi_l}$, (4.13)	`0.2`
	`[t_max_l t_max_h]`	supposed time interval $(\underline{t}_{\max}, \overline{t}_{\max})$ of the response maximum (4.12) used for setting of $(\underline{a}_{d_l}, \overline{a}_{d_l})$; their setting for $d_l = 1$ is modified, see Section 4.3.3; they are specific for individual ROIs, rows correspond to ROIs (axilla, upper arm, forearm)	$\begin{bmatrix} 100 & 360 \\ 60 & 280 \\ 30 & 210 \end{bmatrix}$
	`nA`	number of discretization samples of parameter a_l	`150`
P_1, P_2	`[alpha_g beta_g]`	hyper-parameters (α_g, β_g) of gamma distribution (4.16) for impulse gain g_l; they are determined according (4.18) with range setting 1e5 $< g_l <$ 4e5	`[2.78 1.11e-5]`
	`gain_l`	lower bound of recognized relative response maximum $\underline{r}_{\max}$; used for computation of $\underline{b}_{\Xi_l}$ similarly to `gain_h`	`1e-4`
	`nB`	number of discretization samples of parameter b_l	`150`
N_1, N_2	`[alpha_s,beta_s]`	hyper-parameters (α_s, β_s) of gamma distribution (4.17) for noise precision parameter s; for its refinement see Section 7.2.1; the default values given here are those before refinement obtained employing initially considered range 1e5 $< s <$ 1e8 and (4.18)	`[1.004 2e-8]`
	`gain_l`	same as for models P_1, P_2	
	`nB`	same as for models P_1, P_2	
N_3, N_4	`[alpha_s,beta_s]`	same as for models N_1, N_2	
	`gain_l`	same as for models N_1, N_2	
	`h_q`	the h_q-quantile for the computation of the conservative estimate of precision parameter $\hat{s}$ used for the setting of the Gaussian distribution $\mathrm{N}(b_l \mid \hat{b}_{\Xi_l}, \omega_{\Xi_l} s)$ according (4.20)	`0.95`
	`m_dev`	the hyper-parameter for setting of Gaussian distribution $\mathrm{N}(b_l \mid \hat{b}_{\Xi_l} \omega_{\Xi_l} s)$, (4.19), that determines the ratio of its standard deviation to the half-width of the range $(\underline{b}_{\Xi_l}, \overline{b}_{\Xi_l})$; (4.20) assumes `m_dev=1`, but it can cause problems when `gain_h` is decreased, see Remark(s) 5.1.2 where the proposed solution is given too.	`1`

Table A.1: List of tuning knobs of algorithms for individual models with their default settings

A.3 Code

It is not reasonable to write out here the code of all macros (functions) for all the considered models. Therefore, as an example, only two macros that are necessary for routine evaluation are given bellow. The first one computes estimates of time activity curve and residence time, see Sections 6.1 and 6.3, while the second one computes logarithm of the Bayes factor $\mathcal{B}_{01}(Y_1, Y_2)$, (6.18), necessary for comparison of limbs, see Section 6.5. The macros correspond to model N_2, that has been selected as the best one among all considered models, see Section 7.4. Given macro for comparison of the limbs employs (6.20), i.e. uses simplification that the noise parameter s is a local limb parameter.

Time activity curve and residence time estimates

```
function [E_tac,sdv_tac,E_rt,sdv_rt]=computeTAC(data,time,ROI);
% [E_tac,sdv_tac,E_rt,sdv_rt]=computeTAC(data,time,ROI);
% function computes mean (E) and standard deviation (sdv) of time activity
% curve (tac) and residence time (rt)
%
% Input:  data  column of data
%         time  column of samling times;
%               both data and time have to be of the same length
%         ROI   ROI, data comes from (1,2,3)
%
% Output: E_tac,sdv_tac     vectors of mean and marginal standard deviation of TAC
%         E_rt,sdv_rt         mean and standard deviation of residence time

% load the values of tuning knobs from the initialization file
% (contains alfa_s, beta_s, d_l, d_h, gain_l, gain_h, t_max_l, t_max_h, nA, nB;
% it contains also time t_h,  the upper bound of time, TAC is computed to
run initialization(ROI);

% macro based on the general Algorithm 1 in the Section A.1

% definition of necessary variables
n_data=length(data);
E_tac=zeros(1,t_h+1);
sdv_tac=zeros(1,t_h+1);
E_rt=zeros(1,1);
sdv_rt=zeros(1,1);
w_sum=0;
wL_max=[];

for d=d_l:d_h,          % cycle over model order d
    % computation of bounds (a_l,a_h) of parameter a for given order d
    % according the supposed range of times of TAC maximum
    a_l=find_A_with_maximum_time(d,t_max_l);
    a_h=find_A_with_maximum_time(d,t_max_h);
    a_lL=-log(1-a_l);
    a_hL=-log(1-a_h);
    % discretization step on log(1-a) approximating uniform distribution
    astepL=(a_hL-a_lL)/(nA-1);

    for a_i=1:nA,           % cycle over discretized parameter a
```

```
        % computation of bounds (b_l,b_h) of parameter b for given a,d
        % according the supposed range of TAC maximum determined by gain_l, gain_h
        a=1-exp((-1)*(a_lL+(a_i-1)*astepL));
        b_h=inverse_gain(d,a)*gain_h;
        b_l=b_h*gain_l;
        b_lL=log(b_l);
        b_hL=log(b_h);
        % discretization step on log(b) approximating uniform distribution
        bstepL=(b_hL-b_lL)/(nB-1);

        for b_i=1:nB,       % cycle over discretized parameter b
            b=exp(b_lL+(b_i-1)*bstepL);
            % TAC  (4.9) determined by parameters d,a,b
            pomtac=inner_response(d,a,b,t_h);
            % residence time (6.6) determined by parameters d,a,b
            pomrt=b*(1-a)^(-d);
            % loglikelihood log(f(data|d,a,b)) (wL corresponds to w_{d,a,b} in (5.5))
            wL=loglikelihood(data,time,d,a,b,alfa_s,beta_s);
            % summation with overflow treatment
            if isempty(wL_max),
                wL_max=wL;
                w_sum=1;
                E_tac=E_tac+pomtac;
                sdv_tac=sdv_tac+pomtac.^2;
                E_rt=E_rt+pomrt;
                sdv_rt=sdv_rt+pomrt^2;

            elseif wL>wL_max,
                w_T=exp(log(w_sum)+wL_max-wL);
                w_sum=1+w_T;
                E_tac=(E_tac*w_T+pomtac)/w_sum;
                sdv_tac=(sdv_tac*w_T+pomtac.^2)/w_sum;
                E_rt=(E_rt*w_T+pomrt)/w_sum;
                sdv_rt=(sdv_rt*w_T+pomrt^2)/w_sum;
                wL_max=wL;

            else
                w_T=exp(wL-wL_max);
                wnew=w_sum+w_T;
                E_tac(1,:)=(E_tac(1,:)*w_sum+w_T*pomtac)/wnew;
                sdv_tac(1,:)=(sdv_tac(1,:)*w_sum+w_T*pomtac.^2)/wnew;
                E_rt=(E_rt*w_sum+w_T*pomrt)/wnew;
                sdv_rt=(sdv_rt*w_sum+w_T*pomrt^2)/wnew;
                w_sum=wnew;
            end
        end
    end
end

sdv_tac=sqrt(sdv_tac-(E_tac).^2);
sdv_rt=sqrt(sdv_rt-E_rt^2);
```

Comparison of accumulation on both limbs

```
function rateL=computeComp(data,time,ROI);
% rateL=computeComp(data,time,ROI);
% function computes Bayes factor (6.20)
%   rateL=log(f(data|hyp.-limbs same)/f(data/hyp.-limbs generally different))
%
% Input:  data,time  two-column matrices that have the same number of rows;
%                    they contain measurements and their sampling times on both limbs;
%                    each column corresponds to one limb
%         ROI        ROI, data comes from (1,2,3)
%
% Output: rateL      see above

% load the values of tuning knobs from the initialization file
% (contains alfa_s, beta_s, d_l, d_h, gain_l, gain_h, t_max_l, t_max_h, nA, nB;
run initialization(ROI);

% macro based on the general Algorithm 1 in the Section A.1

% definition of necessary variables
n_data=size(data,1);
w_sum=0;
w_sum1=0;
w_sum2=0;
wL_max=[];
w1L_max=[];
w2L_max=[];
% data and sampling times on both limbs merged together
timeC=[time(:,1);time(:,2)];
dataC=[data(:,1);data(:,2)];
%--
for d=d_l:d_h,          % cycle over model order d
    % computation of bounds (a_l,a_h) of parameter a for given order d
    % according the supposed range of times of TAC maximum
    a_l=find_A_with_maximum_time(d,t_max_l);
    a_h=find_A_with_maximum_time(d,t_max_h);
    a_lL=-log(1-a_l);
    a_hL=-log(1-a_h);
    % discretization step on log(1-a) approximating uniform distribution
    astepL=(a_hL-a_lL)/(nA-1);

    for a_i=1:nA,           % cycle over discretized parameter a
        a=1-exp((-1)*(a_lL+(a_i-1)*astepL));
        b_h=inverse_gain(d,a)*gain_h;
        b_l=b_h*gain_l;
        b_lL=log(b_l);
        b_hL=log(b_h);
        % discretization step on log(b) approximating uniform distribution
        bstepL=(b_hL-b_lL)/(nB-1);

        for b_i=1:nB,    % cycle over discretized parameter b
            b=exp(b_lL+(b_i-1)*bstepL);
            % loglikelihood log(f(data_of_both_limbs|d,a,b))
```

```
            wL =loglikelihood(dataC,timeC,d,a,b,alfa_s,beta_s);
            % loglikelihood log(f(data_of_left_limb|d,a,b))
            wL1=loglikelihood(data(:,1),time(:,1),d,a,b,alfa_s,beta_s);
            % loglikelihood log(f(data_of_right_limb|d,a,b))
            wL2=loglikelihood(data(:,2),time(:,2),d,a,b,alfa_s,beta_s);
            % summation with overflow treatment
            if isempty(w1L_max),
                w1L_max=wL1;
                w_sum1=1;
            elseif wL1>w1L_max,
                w_T=exp(log(w_sum1)+w1L_max-wL1);
                w_sum1=1+w_T;
                w1L_max=wL1;
            else
                w_T=exp(wL1-w1L_max);
                w_sum1=w_sum1+w_T;
            end

            if isempty(w2L_max),
                w2L_max=wL2;
                w_sum2=1;
            elseif wL2>w2L_max,
                w_T=exp(log(w_sum2)+w2L_max-wL2);
                w_sum2=1+w_T;
                w2L_max=wL2;
            else
                w_T=exp(wL2-w2L_max);
                w_sum2=w_sum2+w_T;
            end

            if isempty(wL_max),
                wL_max=wL;
                w_sum=1;
            elseif wL>wL_max,
                w_T=exp(log(w_sum)+wL_max-wL);
                w_sum=1+w_T;
                wL_max=wL;
            else
                w_T=exp(wL-wL_max);
                w_sum=w_sum+w_T;
            end
        end
    end
end
% resulting Bayes factor
rateL=log(w_sum)-log(w_sum1)-log(w_sum2)+wL_max-w1L_max-w2L_max+...
            gammaln((2*n_data+2*alfa_s)/2)-2*gammaln((n_data+2*alfa_s)/2)+...
            log(nA*nB*(d_h-d_l+1))-alfa_s*log(beta_s)+gammaln(alfa_s);
```

Both functions call in their bodies other functions. For the completeness, all necessary functions are listed below.

```
function g=response_point(d,a,b,time);
% g=response_point(d,a,b,time);
% returns one point of impulse response (i.e. TAC) for model
% with parameters d,a,b at time time, see (4.8)

g=b*(nchoosek(time+d-1,time)*a^time);

%-----------------------------------------

function x=inner_response(d,a,b,t_h);
% x=inner_response(d,a,b,t_h);
% computes impulse response (4.9) (i.e. TAC) in time interval (0, t_h)
% for model with parameters d,a,b

x=zeros(1,do+1);
for i=0:do,
    x(i+1)=response_point(d,a,b,i);
end

%-----------------------------------------

function t=time_of_maximum(d,a);
% t=time_of_maximum(d,a);
% finds the time of the maximum (4.12) of impulse response (i.e. TAC) for model
% with parameters d,a

t=floor((d*a-1)/(1-a))+1;

%-----------------------------------------

function b=inverse_gain(d,a);
% b=inverse_gain(d,a);
% returns the inverse of the maximum of impulse response (TAC) for model
% with parameters d,a (b=1)

t=time_of_maximum(d,a);
b=(response_point(d,a,1,t))^(-1);

%-----------------------------------------

function ll=loglikelihood(data,time,d,a,b,alfa_s,beta_s);
% ll=loglikelihood(data,time,d,a,b,alfa_s,beta_s);
% computes loglikelihood f(data|a,d,b) (parameter s is integrated out).
% It corresponds to w_{a,d,b} in (5.5).
%
% Input:  data  column of data
%       time    column of samling times;
%               both data and time have to be of the same length
%       d,a,b   model parameters
%       alfa_s, beta_s hyperparameters of prior gamma distribution (4.17)
%               of noise precision parameter s
%
```

```
% Output: ll    loglikelihood log(f(data|a,d,b))

pom2=0;
n_data=length(data);
for i=1:n_data,
    pom2=pom2+(data(i)-response_point(d,a,b,time(i)))^2;
end
pom2=.5*pom2+beta_s;
ll=log(pom2)*(-(n_data+2*alfa_s)/2);

%-----------------------------------------

function a=find_A_with_maximum_time(d,t);
% a=find_A_with_maximum_time(d,t);
% finds model parameter a given model order d, for which impulse response (TAC) maximum
% is in time t;
% for the model of the first order (d=1) t is time where the response descents to
% its half

bound=1e-6;
la=.9;
ha=.9999999;
C=.5; %descent for the model of the first order
if d>1,
    while time_of_maximum(d,la)>=t
        la=la^2;
    end
    a=la;
    while abs(ha-la)>bound
        a=(la+ha)/2;
        if time_of_maximum(d,a)<t
            la=a;
        else
            ha=a;
        end
    end
else
    while la^t>=C
        la=la^2;
    end
    while ha^t<C
        ha=sqrt(ha);
    end
    a=la;
    while abs(ha-la)>bound
        a=(la+ha)/2;
        if a^t<C
            la=a;
        else
            ha=a;
        end
    end
end
```

Bibliography

[1] J. Altorfer and L. Clodius. Chronic experimental lymphedema of the extremities: pathological changes. *Experentia*, 32:823–824, 1976.

[2] M. Ambramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1964.

[3] K. J. Astrom and B. Wittenmark. Problems of identification and control. *J. Math. Anal. Appl.*, 34:90–113, 1971.

[4] R.C. Atkins, E.M. Briganti, P.Z. Zimmet, and S.J. Chadban. Association between albuminuria and proteinuria in the general population: the ausdiab study. *Nephrology Dialysis Transplantation*, 18(10):2170–2174, October 2003.

[5] H.H. Barret and W. Swidell. *Radiological Imaging. The Theory of Image Formation, Detection, and Processing*, volume 2. Academic Press, 1981.

[6] B.O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, 1985.

[7] J.M. Bernardo and A.F.M. Smith. *Bayesian theory*. John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1997. 2nd edition.

[8] G.E.P. Box and G.M. Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, 1970.

[9] P.J. Brockwell and R.A. Davis. *Introduction to Time Series and Forecasting*. Texts in Statistics. Springer-Verlag, Berlin, 1996.

[10] A. Butkovskiy and L.M. Pustylnikov. *Characteristics of Distributed-Parameter Systems: Handbook of Equations of Mathematical Physics and Distributed-Parameter Systems*. Kluwer Academic Publishers, Dordrecht, 1993.

[11] R.A. Cambria, P. Gloviczki, J.M. Naessens, and H.W. Wahner. Nonivasive evaluation of the lymphatic system with lymphoscintigraphy — a prospective, semiquantitative analysis in 386 extremities. *Journal of Vascular Surgery*, 18(5):773–782, 1993.

[12] M. Carena, C. Aprole, G.P. Zelaschi, G. Rossi, G. Paroni, and R. Campini. Lymphoscintigraphy in study of lymphedema of the arms. *Radiology in Medicine*, 73(4):310–312, 1987.

[13] M. Carena, R. Campini, G. Zelaschi, G. Rossi, C. Aprile, and G. Paroni. Quantitative lymphoscintigraphy. *European Journal of Nuclear Medicine*, 14(2):88–92, 1988.

[14] T.C. Case, C.L. Witte, and M.H. Witte at al. Magnetic resonance imaging in human lymphoedema: comparison with lymphangioscintigraphy. *Magnetic Resonance Imaging*, 10:549–558, 1982.

[15] E. Comets, K. Ikeda, P. Hoff, P. Fumoleau, J. Wanders, and Y. Tanigawara. Comparison of the pharmacokinetics of s-1, an oral anticancer agent, in western and japanese patients. *Journal of Pharmacokinetics and Pharmacodynamics*, 30(4):257–283, 2003.

[16] M.H. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill Company, New York, 1970.

[17] T.E. Djaferis and I.C. Schick. *System Theory: Modelling, Analysis and Control*. Kluwer Academic Publishers, Boston, 1999.

[18] S.B. Doldi, E. Lattuada, and M.A. Zappa et al. Ultrasonography of extremity lymphedema. *Lymphology*, 25:129–133, 1992.

[19] J. Egan. *Signal Detection Theory and ROC Analysis*. Academic Press, New York, 1975.

[20] M. Encinas, R. de Juan, A. Marcos, P. Gil, A. Barabash, C. Fernaandez, C. de Ugarte, and J.A. Cabranes. Regional cerebral blood flow assessed with tc-99m-ecd spet as a marker of progression of mild cognitive impairment to alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging*, 30(11):1473–1480, November 2003.

[21] P. Eykhoff. *System Identification, Parameter and State Estimation*. John Wiley, New York, 1974.

[22] E. Foldi, M. Foldi, and H. Weissleder. Conservative treatment of lymphoedema of the limbs. *Angiology*, 36:171–180, 1985.

[23] W.F. Ganong. *Review of Medical Physiology*. Apleton and Lange, Norwalk, Connecticut, 15th edition, 1991.

[24] P. Gebouský. On modelling & estimation problems with armax model. Technical Report 2041, ÚTIA AV ČR, Praha, 2001.

[25] P. Gebouský, M. Kárný, and H. Křížová. Contribution to quantitative evalution of lymphoscintigraphy of upper limbs. *ERCIM News*, 43:41, 2000.

[26] P. Gebouský, M. Kárný, and A. Quinn. Scintigraphic quantification of lymphatic function: Bayesian inference on diffusion dynamics. In A. Rakar, editor, *3rd International PhD Workshop: Advances in Supervision and Control Systems, Young Generation Viewpoint*, pages 31–38, Strunjan, Slovenia, October 1–4, 2002 2002. Department of Computer Automation and Control, Jožef Stefan Institute, Slovenia.

[27] P. Gebouský, M. Kárný, and A. Quinn. Lymphoscintigraphy of upper limbs: A bayesian framework. In J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. David, D. Heckerman, and A.F.M. Smith, editors, *Bayesian Statistics 7*, pages 543–552, Oxford, 2003. Clarendon Press. Proceedings of the Seventh Valencia International Meeting, June 2–6, 2002.

[28] P. Gebouský and H. Křížová. Bayesian identification from incomplete data with application to lymphedema evaluation. Technical Report 1991, ÚTIA AV ČR, Praha, 2000.

[29] A.E. Gelfand, A.F.M. Smith, and T.M. Lee. Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *Journal of American Statistical Association*, 87:523–532, 1992.

[30] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo in practice.* Chapman & Hall, London, 1997.

[31] P. Gloviczki, D. Calgano, and A. Schirger. Noninvasive evaluation of the swollen extremity: experiences with 190 lymphoscintigraphic examinations. *Journal of Vascular Surgery*, 9:683–689, 1989.

[32] E. Goltner, P. Gass, J.P. Haas, and P. Schneider. The importance of volumetry, lymphoscintigraphy, and computed tomography in the diagnosis of brachial oedema after mastectomy. *Lymphology*, 21:134–143, 1988.

[33] J.M. Gottman. *Time-series analysis: A comprehensive introduction for social scientists.* Cambridge University Press, Cambridge, 1981.

[34] M. De Groote, C. Jonnart, F. Puissant, J. Buisset, and E. Schlikker. Lymphoscintigraphic evaluation of the efficiency of manual lymphatic drainage. *European Journal of Lymphology*, 3(85-87), 1992.

[35] W.S. Halsted. The swelling of the arm after operations for cancer of the breast — elephantiasis chirurgica — its cause and prevention. *Bouletin of the John Hopkins Hospital*, 32:309–313, 1921.

[36] W.S. Handley. Lymphangioplasty: a new method for the relief of the brawny arm of breast-cancer and for similar conditions of lymphatic oedema. *Lancet*, 1:783–785, 1908.

[37] B. Harris and G. Heinel. The relation between statistical decision theory and approximation theory. In J.G. Rustagi, editor, *Optimizing Methods in Statistics*, pages 263–272. Academic Press, New York, San Francisco, London, 1979.

[38] A.E. Hawley, L. Illum, and S.S David. The effect of oedema on the uptake of colloid to the lymph nodes. *Biopharm Drug Dispos*, 19:193–197, 1998.

[39] J. Heřmanská, M. Kárný, L. Jirsa, and J. Němec. Evaluation of biophysical quantities related to the treatment of thyroid diseases. In *Quantitative Image Analysis in Functional Scintigraphic Imaging*, pages 39–42, Prague, September 1998. GAMAMED.

[40] J.K. Hwang, J.Y. Kwon, and K.W. Lee et al. Changes in lymphatic function after complex physical therapy for lympedema. *Lymphology*, 32:15–21, 1999.

[41] Petrek J.A., Pressman P.I., and Smith R.A. Lymphedema: current issues in research and management. *CA — Cancer Journal for Clinicians*, 50(5):292–307, September/October 2000.

[42] A.M. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 1970.

[43] H. Jeffreys. *Theory of Probability.* Clarendon Press, Oxford, 1985.

[44] H. Jeffreys and B.S. Jeffreys. *Methods of Mathematical Physics.* Cambridge:University Press, Cambridge, third edition, 1972.

[45] N.L. Johnson and S. Kotz. *Discrete Distributions.* Wiley, New York, 1969.

[46] N.L. Johnson and S. Kotz. *Continuous Univariate Distributions.* Wiley, New York, 1970.

[47] N.L. Johnson and S. Kotz. *Continuous Multivariate Distributions*. Wiley, New York, 1972.

[48] W.D. Kaplan, S.A. Slavin, J.A. Markisz, S.M. Laffin, and H.D. Royal. Qualitative and quantitative upper extremity radionuclide lymphoscintigraphy. *Journal of Nuclear Medicine*, 24:40, 1983.

[49] M. Kárný, J. Böhm, T.V. Guy, and P. Nedoma. Mixture-based adaptive probabilistic control. *International Journal of Adaptive Control and Signal Processing*, 17(2):119–132, 2003.

[50] M. Kárný and I. Nagy. Dynamic Bayesian Decision-making: Part I. Technical Report 1971, ÚTIA AV ČR, Praha, 1999.

[51] R.L. Kashyap and A.R. Rao. *Dynamic Stochastic Models from Empirical Data*. Academic Press, New York, 1976.

[52] J.B. Kinmonth. *The Lymphatics. Diseases, Lymphography and Surgery*, chapter Physiology of the tissue fluid balance and oedema. Contractility of lymphatics, pages 82–86. Arnold, London, 1972.

[53] E. Kleinhaus. Evaluation of transport kinetics in lymphpscintigraphy: follow up study in patient in transplant lymphatic vessels. *European Journal of Nuclear Medicine*, 10:349–352, 1985.

[54] S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–87, 1951.

[55] M. Legrand, E. Comets, G. Aymard, R. Tubiana, C. Katlama, and B. Diquet. An in vivo pharmacokinetic/pharmacodynamic model for antiretroviral combination. *HIV Clinical Trials*, 4(3):170–183, 2003.

[56] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, London, 1987.

[57] R. Matas. The surgical treatment of elephantiasis and elephantoid states dependent upon chronic obstruction of the lymphatic and venous channels. *American Journal of Tropical Diseases*, 1:60–85, 1913.

[58] The MathWorks, Inc. *MATLAB User's Guide*, 1992.

[59] C.E. Metz. Basic principles of roc analysis. *Seminars Nucl. Med.*, 8(4):283–298, 1978.

[60] P.S. Mortimer. Managing lymphoedema. *Clin Exp Dermatology*, 20:98–106, 1995.

[61] P.S. Mortimer. The pathophysiology of lymphoedema. *Cancer*, 83(12):2798–2802, 1998.

[62] P.S. Mortimer, C.M.A Young, and T.J. Ryan. Characterisation of lymph flow by simultaneous and continuous monitoring of depot clearance and nodal uptake using ^{99m}tc-labelled colloid. *Prog Lymphol*, 8:132–134, 1985.

[63] A. Mostbeck, H. Partsch, and P. Kahn. Quantitative radionuclide lymphography. *Prog Lymphol*, 8:135–137, 1985.

[64] M. Mozes, M.Z. Papa, A. Karasik, A. Reshef, and R. Adar. The role of infection in post-mastectomy lymphedema. *Annals of Surgery*, 14:73–83, 1982.

[65] J. Myles, S. Duffy, R. Nixon, C. Boggis, A. Howell, A. Shenton, and G. Evans. Initial results of a study into the effectiveness of breast cancer screening in a population identified to be at high risk. *Revue d Epidemiologie Et De Sante Publique*, 49(5):471–475, October 2001.

[66] S.D. Nathanson, L. Nelson, and K.C. Karvelis. Rates of flow of technecium 99m-labeled human serum albumin form peripheral injection sites to sentinel lymph nodes. *Annals of Surgical Oncology*, 3(4):329–335, 1996.

[67] L. Nieborg. *Vascular surgery*, chapter 29.2 Lymphatic-venous and Lymph-Nodal-Venous Anatomoses. Springer-Verlag, Berlin, New York, 1989.

[68] R.T. Ogden. *Essential wavelets for statistical application and data analysis.* Birkhăuser, Boston, Basel, Berlin, 1997.

[69] A.V. Oppenheim, R.W. Schafer, and J.R. Buck. *Discrete-Time Signal Processing.* Prentice Hall, 1999.

[70] A.V. Oppenheim and A.S. Wilsky. *Signals and systems.* Prentice Hall, New Yersey, 1983.

[71] O. Osterberg, L. Erichsen, S.H. Ingwersen, A. Plum, H.E. Poulsen, and P. Vicini. Pharmacokinetic and pharmacodynamic properties of insulin aspart and human insulin. *Journal of Pharmacokinetics and Pharmacodynamics*, 30(3):221–235, 2003.

[72] P.Van Overschee and B.De Moor. *Subspace Identification for Linear Systems – Theory, Implementation, Applications.* Kluwer Academic Publishers, Boston/London/Dordrecht, 1996.

[73] S.J. Pain. Lymphoedema following surgery for breast cancer. *British Journal of Surgery*, 87:1128–1141, 2000.

[74] S.J. Pain, R.S. Nicholas, R.W. Barber, J.R. Ballinger, A.D. Purushotham, P.S. Mortimer, and A.M. Peters. Quantification of lymphatic function for investigation of lymphedema: Depot clearance and rate of appeareance of soluble macromolecules in blood. *The Journal of Nuclear Medicine*, 43(3):318–324, 2002.

[75] H. Partsch. Assesment of abnormal lymph drainage for the diagnosis of lymphedema by isotopic lymphangiography and by indirect lymphography. *Clinics in Dermatology*, 13:445–450, 1995.

[76] A. Pecking, F. Firmin, and J.D. Rain et al. Lymphedema of the upper limb following surgery or radiotherapy: investigation by indirect radioactive lymphography. *Nouv Presse Med*, 9:3349–3351, 1980. in French.

[77] V. Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.

[78] N.B Piller. Lymphedema, macrophages and benzypyrones. *Lymphology*, 13:109–119, 1980.

[79] C.R. Rao. *Linear method of statistical inference and their applications.* Academia, Prague, 1987. in Czech.

[80] B.D. Ripley. *Stochastic simulation.* Wiley, Chichester, 1987.

[81] J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, editors. *Computer-Intensive Methods in Control and Data Processing. Can We Beat the Curse of Dimensionality? 3rd European IEEE Workshop CMP'98*, Prague, September 1998.

[82] D.B. Rubin. Using the sir algorithm to simulate posterior distributions. In J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, editors, *Bayesian Statistics 3*, pages 395–402, Oxford, 1988. University Press. (with discussion).

[83] N.K. Sinha and B. Kuszta. *Modelling and Identification of Dynamic Systems*. Van Nostrand-Reinhold Company, 1983.

[84] A.F.M. Smith and A.E. Gelfand. Bayesian statistics without tears: a sampling-resamplig perspective. *American Statistics*, 46:84–88, 1992.

[85] T. Söderström and P. Stoica. *System Identification*. Series in Systems and Control Engineering. Prentice-Hall, London, 1989.

[86] M.G. Stabin. Mirdose 3 software for internal dose assesment in nuclear medicine. *Journal of Nuclear Medicine*, 37:538–546, 1996.

[87] A.W.B. Stanton, W.E. Swensson, R.H. Mellor, A.M. Peters, JR Levick, and P.S. Mortimer. Differences in lymph drainage between swollen and non-swollen regions in arms with breast-cancer-related lymphoedema. *Clinical Science*, 101(2):131, 2001.

[88] F.W. Stewart and N. Treves. Lymphangisarcoma in postmastectomy lymphoedema. a report of six cases in elephantiasis chirurgica. *Cancer*, 1:64–81, 1948.

[89] W. Svensson, D.M. Glass, D. Bradley, and A.M. Peters. Measurement of lymphatic function with technecium-99-m-labelled polyclonal immunoglobulin. *European Journal of Nuclear Medicine*, 26(5):504–510, May 1999.

[90] A. Szuba and S.G. Rockson. Lymphedema: Classification, diagnosis and therapy. *Vascular Medicine*, 3:145–156, 1998.

[91] M.A. Tanner. *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. Springer Series in Statistics. Springer-Verlag, New York, 2nd edition, 1993.

[92] S.E. Ter, A. Alavi, C.K. Kim, and G. Merli. Lymphoscintigraphy. a reliable test for the diagnosis of lymphedema. *Clin. Nucl. Med.*, 18(8):646–654, 1993.

[93] D.M. Titterington, A.F.M. Smith, and U.E. Makov. *Statistical Analysis of Finite Mixtures*. John Willey & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1985.

[94] M. Vaqueiro, P. Gloviczki, J. Fisher, L.H. Hollier, A. Schirger, and H.W. Wahner. Lymphoscintigraphy in lymphedema: an aid to microsurgery. *Journal of Nuclear Medicine*, 27(7):1125–1130, 1986.

[95] H. Verlooy, J.P. Biscompte, L. Nieuborg, P. Drent, C. Schiepers, L. Mortelmans, and M. De Roo. Noninvasive evaluation of lympho-venous anastomosis in upper limb lymphedema. *The European Journal of Lymphology*, 6(21):27–33, 1997.

[96] A. Wald. *Statistical Decision Functions*. John Wiley&Sons, New York, London, 1950.

[97] M. Weiss, R.G.H. Baumaister, K. Tatsch, and K. Hahn. Lymfoscintigraphy and semi-quantitative evaluation of lymph drainage for long-term follow-up in patients with autogenous lymph vessel transplantation. *The European Journal of Lymphology*, 4:34–37, 1997.

[98] H. Weissleder and M. Weissleder. Lymphedema: Evaluation of qualitative and quantitative lymphoscintigraphy in 238 patients. *Radiology*, 167(3):729–735, 1988.

[99] M. Woods, M. Tobin, and P.S. Mortimer. The psychosocial morbidity of breast cancer patients with lymphoedema. *Cancer Nurse*, 18:467–471, 1995.

Index