Lucian Mihai Itu · Puneet Sharma
Constantin Suciu *Editors*

Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies

Springer

Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies

Lucian Mihai Itu • Puneet Sharma
Constantin Suciu
Editors

Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies

Springer

Editors
Lucian Mihai Itu
Transilvania University of Brasov
Brasov
Romania

Constantin Suciu
Siemens Corporate Technology
Brasov
Romania

Puneet Sharma
Medical Imaging Technologies
Siemens Healthcare
Princeton, New Jersey
USA

ISBN 978-3-319-56852-2 ISBN 978-3-319-56853-9 (eBook)
DOI 10.1007/978-3-319-56853-9

Library of Congress Control Number: 2017943812

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Cardiovascular disease is the leading cause of death, globally. Of these deaths, approximately 42% are caused by coronary artery disease.

Hemodynamic computations represent a state-of-the-art approach for patient-specific assessment of cardiovascular pathologies. Thus, blood flow computations, when used in conjunction with patient-specific anatomical models extracted from medical images, provide important insights into the structure and function of the cardiovascular system. In recent years, these techniques have been proposed for diagnosis, risk stratification, and surgical planning.

The aim of this book is to present modern approaches for truly patient-specific diagnosis of cardiovascular pathologies (e.g., coronary artery disease, aortic coarctation), based on personalized computational blood flow models. All personalized computational models introduced herein are validated against patient-specific measurements in relatively large patient cohorts (n = 4–90). High-performance computing solutions are further introduced, which reduce the execution time of the computational models, rendering them suitable for routine clinical workflows.

The foundation of the works presented herein has been the excellent collaboration over many years between Siemens, with its offices in Brasov, Romania (Corporate Technology, Siemens SRL), and Princeton, NJ, USA (Medical Imaging Technologies, Siemens Healthcare), the Transilvania University of Brasov, the University Politehnica of Bucharest, and the Clinical Emergency Hospital of Bucharest. Hence, the methods and results presented herein are the result of a unique synergy between academia, industry, and clinical research.

Specifically, the seven chapters of the book present the development of reduced-order multiscale hemodynamic models for coronary artery disease, aortic coarctation, and whole-body circulation, which can be applied in routine clinical settings for personalized diagnosis. Advanced parameter estimation frameworks are introduced for calibrating the parameters of the models.

The chapters are written for scientists in the field of biomedical engineering focusing on pathologies of the cardiovascular system, as well as for research-oriented physicians in cardiology and industrial players in the field of healthcare technologies.

We would further like to acknowledge that the work was partially supported by the program Partnerships in Priority Domains (PN II), financed by ANCS, CNDI – UEFISCDI, under the project no. 130/2012, and by the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 600932.

The editors express their gratitude to the contributors of the seven chapters and to other collaborators who put enormous effort in preparing the examples, running the software, and drawing the figures.

We are all deeply grateful to our families who have supported us over many years in this work.

Brasov, Romania	Lucian Mihai Itu
New Jersey, USA	Puneet Sharma
Brasov, Romania	Constantin Suciu

Contents

Contributors

Elisabeta Badila Internal Medicine, Clinical Emergency Hospital, Bucharest, Romania

Internal Medicine, "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania

Vlad Bataila Interventional Cardiology, Clinical Emergency Hospital, Bucharest, Romania

Lucian Calmac Interventional Cardiology, Clinical Emergency Hospital, Bucharest, Romania

Ulrich Foerster Corporate Technology, Siemens AG, Munich, Germany

Harz University of Applied Sciences, Wernigerode, Germany

Alina Gîrbea Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Lucian Itu Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Ali Kamen Medical Imaging Technologies, Siemens Healthcare, Princeton, NJ, USA

Laszlo Lazăr Corporate Technology, Siemens SRL, Brasov, Romania

Tommaso Mansi Medical Imaging Technologies, Siemens Healthcare, Princeton, NJ, USA

Viorel Mihalef Medical Imaging Technologies, Siemens Healthcare, Princeton, NJ, USA

Cosmin Nita Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Tiziano Passerini Medical Imaging Technologies, Siemens Healthcare, Princeton, NJ, USA

Daniela Penes Internal Medicine, Clinical Emergency Hospital, Bucharest, Romania

Puneet Sharma Medical Imaging Technologies, Siemens Healthcare, Princeton, NJ, USA

Iulian Stroia Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Constantin Suciu Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Irina Andra Tache Department of Automatic Control and Systems Engineering, University Politehnica of Bucharest, Bucharest, Romania

Anamaria Vizitiu Corporate Technology, Siemens SRL, Brasov, Romania

Automation and Information Technology, Transilvania University of Brasov, Brasov, Romania

Emma Weiss Internal Medicine, Clinical Emergency Hospital, Bucharest, Romania

Internal Medicine, “Carol Davila” University of Medicine and Pharmacy, Bucharest, Romania

Diana Zamfir Laboratory of Echocardiography, Clinical Emergency Hospital, Bucharest, Romania

Computational Modeling of the Human Cardiovascular System

1

Puneet Sharma, Lucian Itu, and Constantin Suciu

Abstract

Hemodynamic computations represent a modern approach for the patient-specific diagnosis of cardiovascular pathologies. These are based on data acquired through medical imaging, and Computational Fluid Dynamics (CFD), which is one of the major topic areas in the field of fluid mechanics, employing numerical methods and algorithms for solving and analyzing applications related to fluid movements. Since the cardiovascular system is a closed-loop system, the simulations performed for particular segment have to take into account the influence of the other cardiovascular components. However, since CFD based computations are computationally expensive, multiscale models have been proposed, which combine detailed three-dimensional modeling in the region of interest with one- or zero-dimensional modeling for the remaining components of the cardiovascular system. Finally, parallel processing techniques are typically employed to further reduce the execution time.

1.1 Introduction

The modeling of arterial hemodynamics is based on the theory of fluid mechanics. Computational Fluid Dynamics (CFD) is one of the major topic areas in the field of fluid mechanics, which employs numerical methods and algorithms for solving and

---

P. Sharma (✉)
Medical Imaging Technologies, Siemens Healthcare, 755 College Road,
Princeton, NJ 08540, USA
e-mail: sharma.puneet@siemens.com

L. Itu • C. Suciu
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

Automation and Information Technology, Transilvania University of Brasov,
Mihai Viteazu nr. 5, Brasov 5000174, Romania

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_1

analyzing applications related to fluid movements. The fundamental equations which govern fluid flow are the Navier-Stokes system of equations (conservation of mass, conservation of momentum and conservation of energy). This system of equations can be solved analytically only under specific simplified and idealized conditions. Hence, typically, the Navier-Stokes equations are solved numerically.

During the last decades several numerical methods for solving these equations have been developed: finite element method, finite volume method, finite difference method, Lattice-Boltzmann method, etc. (Wendt 2009).

In hemodynamics applications, which simulate the flow of blood in the cardiovascular system, the Navier-Stokes equations are solved inside of a closed domain, delimited by the walls of the arteries/heart chambers, etc. The incompressibility of blood leads to a simplification of the Navier-Stokes equations, which, typically are solved without considering the energy conservation equation (the temperature is considered to be constant).

Hemodynamic simulations generally focus on a particular segment/region of interest in the cardiovascular system, which is modeled in a three-dimensional pace given by an anatomical model reconstructed from medical imaging data acquired through Computer Tomography (CT), Magnetic Resonance Imaging (MRI), X-ray Angiography (XA), Ultrasound (US), etc. Due to the large execution times, only limited sections of the cardiovascular system can be considered for three-dimensional simulations. Since the cardiovascular system is a closed-loop system, however, the simulations performed for particular segment, without taking into account the influence of the other cardiovascular components, leads to erroneous results, especially in terms of pressure results (Quarteroni et al. 2001).

To mitigate this aspect, multiscale models have been proposed, which combine the detailed three-dimensional modeling in the region of interest with one- or zero-dimensional modeling for the remaining components of the cardiovascular system.

A crucial role in any fluid dynamics simulation is played by the boundary conditions. In hemodynamics simulations, the inlet boundary condition is typically given by patient-specific measurements (e.g. phase contrast MRI measurements) or by lumped parameter models representing the upstream cardiovascular components (e.g. the heart). The boundary condition at the arterial walls is given by the blood viscosity (the no-slip condition) (Perktold and Rappitsch 1995). The outlet boundary condition typically poses the largest difficulties, whereas several types of such conditions can be imposed (Vignon-Clementel et al. 2010): constant pressure, resistance, windkessel model with varying number of parameters, structured tree, etc.

Since the numerical simulation of arterial hemodynamics requires large execution times, there is a continuous need for faster computations. Parallel processors lead to an improvement of the execution time, since the majority of fluid dynamics algorithms can be parallelized efficiently. With the technological advancements of Graphics Processing Units (GPU), which contain a very large number of cores (more than 1000), and especially with the introduction of the CUDA (Compute Univied Device Architecture) language cost-efficient parallelization can be performed (Jesperson 2009). Similar to other applications where execution time is crucial, hemodynamic computations can profit from the porting of computationally intensive parts of the algorithms to the GPU in three ways: solve the same problem in a shorter amount of time, solve more complex problems in the same amount of time, and obtain better solutions for problems which need to be solved in a given amount of time (Kirk and Hwu 2010).

1.2 Multiscale Modeling of Arterial Hemodynamics

During the last years, the progress achieved in the field of computational fluid dynamics, together with the technological advancements in medical imaging, has enabled the quantitative analysis of blood flow (Vignon-Clementel et al. 2010).

As mentioned above, one of the main difficulties for the correct modeling of the cardiovascular system is its closed-loop nature, with significant inter-dependencies between the individual components. The local blood flow characteristics are closely correlated with global system dynamics (Vignon-Clementel et al. 2006). The distribution of blood in various vascular segments is a property of the entire system, but it influences the dynamics of each individual segment. The study of local hemodynamics is of paramount importance since several pathologies like the thickening of the arterial wall or the development of stenoses are influenced by local hemodynamics. On the other hand, certain local changes, like those of the vascular lumen, can lead to a change in the global blood flow distribution, triggering a compensatory mechanism which in turn affects the blood flow distribution. For example, it has been shown, that a significant reduction of the carotid artery lumen does not lead to a significant decrease of blood flow to the brain, due to the presence of these compensatory mechanisms (Formaggia et al. 1999).

This reciprocal influence between global or systemic hemodynamics and local hemodynamics has lead to the development of the geometrical multiscale modeling of the cardiovascular system.

The multiscale perspective is important even when the focus lies on local hemodynamics. Artificial boundary conditions (at the inlet and outlet of the region of interest) delimit the segment of interest: they are defined based on the global hemodynamics and modeled using one- or zero-dimensional models.

One-dimensional models are derived using a series of simplifying assumptions from the three-dimensional Navier-Stokes equations. A fundamental aspect of this is model is that it takes into account the compliance of the vessels, i.e. it represents a reduced-order fluid-structure interaction model, and thus allows for a proper modeling of the wave phenomena. Whereas one-dimensional models are represented through partial differential equations, zero-dimensional models are represented through ordinary differential equations. They model the time-varying spatially averaged pressures and flow rates in a compartment of the cardiovascular system.

Both models are of great importance for the multiscale modeling since they allow for a proper modeling of the fundamental aspects in the cardiovascular circulation in a reasonable amount of time.

Another aspect of interest is the coupling of the simplified models with the three-dimensional model (Formaggia et al. 2001; Quarteroni et al. 2001), as displayed in Fig. 1.1. The coupling is a challenging task due to the different nature of the coupled models. The Navier-Stokes equations are a system of partial differential equations

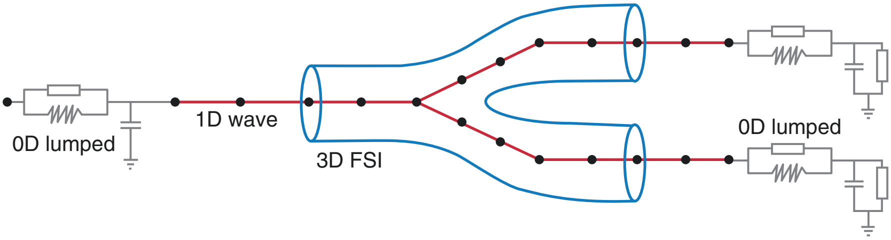


Fig. 1.1 Coupling of three-, one- and zero-dimensional models for a multiscale fluid-structure interaction simulation (van de Vosse and Stergiopulos 2011)

of parabolic type from the velocities' point of view; while the one-dimensional models are hyperbolic, and the zero-dimensional models have no spatial dependence (only time derivatives are present in these equations).

Similar to regular CFD based applications, in case of multiscale models of the arterial circulation the following workflow is employed (Chung 2002):

- Pre-processing: the geometry of the application is defined, the volume occupied by the fluid is divided into finite cells (the grid may be uniform or non-uniform) and the physical model is defined. Next, the boundary conditions are defined (the behavior of the fluid at the boundaries of the region of interest is defined);
- The simulation is started and the equations are solved iteratively in space and/or time to determine the steady-state of transient solution;
- Post-processing: the numerical results are analyzed and visualized.

1.3 Patient-Specific Multiscale Modeling of Arterial Hemodynamics

The prerequisite for generating a patient-specific multiscale model is the generation of the anatomical model. During the last decades tremendous progresses have been made in the field of medical imaging, and, hence, nowadays various technologies are available for acquiring the input data required for the reconstruction of anatomical models.

Steinman et al. underlined the fact that the opportunities generated by the state-of-the-art medical imaging techniques come with both advantages and disadvantages. If these techniques are employed correctly, they can help clinicians in understanding the cardiovascular pathologies and boost the development and evaluation of medical devices and therapeutic interventions. The disadvantage is that these tools may lead to large quantities of data, without leading to a similar increase in the understanding of the phenomena that cause the observed results (Steinman and Taylor 2005).

The first image-based CFD simulations were performed for understanding the role of hemodynamic forces in the development of carotid artery atherosclerosis (Milner et al. 1998) and coronary artery atherosclerosis (Krams et al. 1997).

Furthermore, hemodynamics in abdominal aortic aneurysms (Taylor et al. 1996) and the planning of surgical interventions have been of interest (Taylor and Draney 2004). The above mentioned compartments of the cardiovascular system have been chosen on one hand because of their clinical significance, and on the other hand due to the possibility of properly visualizing the arterial lumen and wall. Thus, the carotid bifurcation is ideal for ultrasound and MRI. The abdominal aorta can be visualized through MRI or CT, while coronary arteries are typically investigated through X-ray angiography.

Significant progress was made in the understanding of the role of hemodynamics forces in the development of coronary plaque (Slager et al. 2005). An MRI-based study has shown that the locations with large wall shear stress have a higher risk of rupture in case of coronary arteries (Groen et al. 2007). On the other hand, a hemodynamic study of carotid arteries has indicated that the locations with low wall shear stress tend to have the largest plaque growth (Tang et al. 2008).

Several studies have focused on cerebral aneurysms. Hence, it was concluded that low wall shear stress leads to progressions and rupture of cerebral aneurysms (Shojima et al. 2004). A subsequent study, performed on 62 patients (Cebral et al. 2005) has shown that the dimension of the aneurysms and the inlet flow rate are correlated. Due to the risks associated with surgical interventions for aneurysms, intact aneurysms are often only monitored clinically, providing thus an ideal setting for studying their evolution. Thus, the roles of low wall shear stress and stagnant blood flow for the evolution of aneurysms have been studies (Boussel et al. 2008).

Remarkable advancements have also been reported in the modeling of the abdominal aorta for diagnosing atherosclerosis and analyzing the evolution of aneurysms (Tang et al. 2006). Patient-specific hemodynamics have been analyzed in anatomical models reconstructed from MRA (Magnetic Resonance Angiography) medical images. Furthermore phase contrast MRI was employed to determine blood flow velocities at rest. Hemodynamic computations were then performed for the rest state and a state of moderate physical exercise. The latter ones proved that physical exercise is beneficial for the appearance of mechanical stimuli, which in turn trigger long lasting biological processes and protect the arteries from the onset and the evolution of atherosclerosis.

Hemodynamic simulations are not only useful for understanding the arterial circulation but also for planning a patient-specific treatment course. Wilson et al. have described the development of a software system for planning surgical interventions, proving the utility of a hemodynamic model in the planning of aorto-femoral bypasses for patients with total occlusions in peripheral arteries (Wilson et al. 2005). This technique was also employed to simulate and compare the effects of total occlusions in vertebral arteries (Hassan et al. 2004).

MRA imaging data was also employed to generate an anatomical model of an inoperable aneurysm in the basilar artery (Fig. 1.2). PC-MRI was used to determine the inlet boundary condition, and CFD techniques were employed to determine blood flow velocities and the wall shear stress distribution in case one of the supplying arteries was blocked (Acevedo-Bolton et al. 2006). The authors concluded that by modifying the cerebral blood flow, the sections with low wall shear stress can be

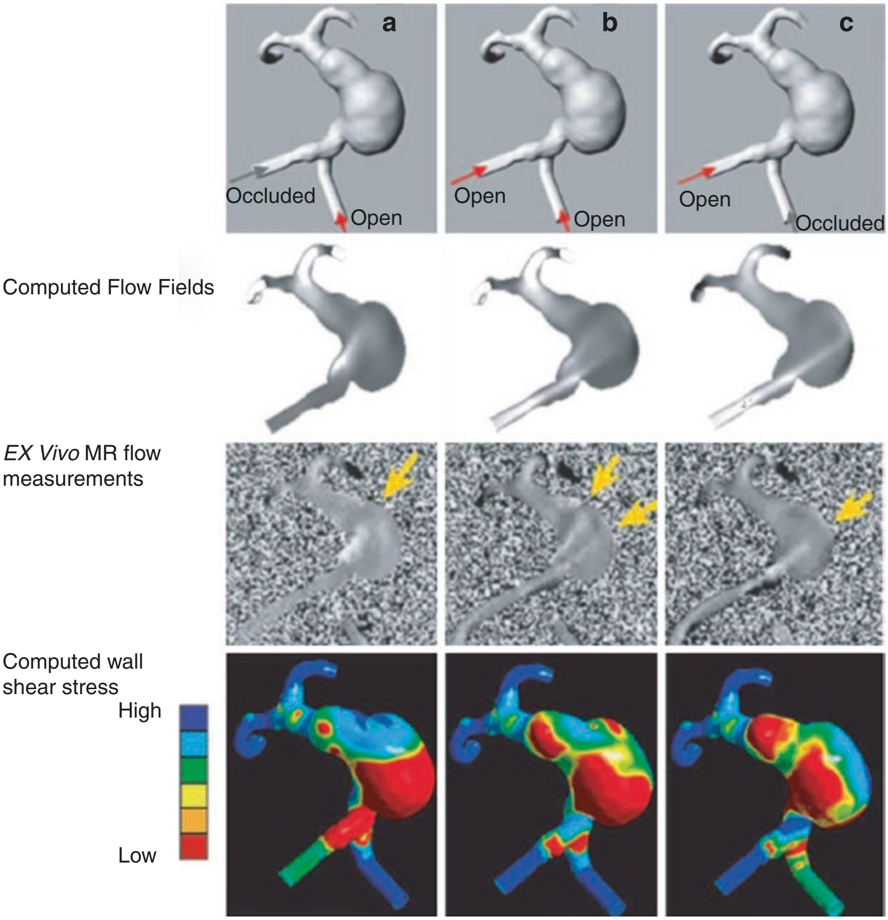


Fig. 1.2 Comparison between CFD simulations and experimental results for three different occlusion configurations (Acevedo-Bolton et al. 2006)

modified, and the sections with stagnant flow can be diminished (stopping thus the further evolution of the aneurysm).

The first step in the process of image-based patient-specific hemodynamic modeling is the segmentation of the arterial walls. Initially, the segmentation was performed manually on 2D images, whereas the 3D anatomical model was subsequently generated by combining the 2D images (Steinman 2002). During the last years a tremendous progress has been made in terms of spatial resolution and image quality. Coupled with stronger graphics processors, this has enabled a real-time reconstruction of 3D anatomical models. A preferred approach is the level-set method, which is typically employed in immersed boundary methods (Antiga et al. 2008). In case of volumetric images with well-defined arterial walls (as acquired through CT and MR with contrast agent) 3D segmentation is preferred due to the speed and robustness of the methodology. On the other hand, the modification of 3D surfaces is cumbersome, and, hence, in case of low contrast images, 2D segmentation is preferred (Wang et al. 1999). This technique is employed for example for IVUS (Intravascular Ultrasound) images, which are acquired at non-parallel planes (Slager et al. 2000). To generate a three-dimensional model of the vessel, a technique called *lofting* is employed, which connects the 2D rings. This technique does not pose any major difficulties in case of single vessels, but is cumbersome to apply in case of bifurcations (Gijsen et al. 2007). Finally, open-source platforms can be nowadays readily used to generate the three-dimensional models (vmtk, simtk, etc.).

Once the surface of the arterial segments of interest has been determined, a grid with a finite number of elements has to be generated for which the multiscale model will be solved. Commercial or open-source platforms may be employed at this stage. In hemodynamic simulations the quantities of interest may be global hemodynamic properties (like pressure or flow-rate) or local (like the wall shear stress). For the former, a relatively coarse grid may be generated (leading thus to a short execution time), while for the latter a denser grid is required (leading to a large execution time). For a reasonable compromise between accuracy and execution speed, adaptive grid refinement techniques are employed, so as to generate a fine grid in the regions of interest and a relatively coarse grid for the rest of the domain.

As described above, to run multiscale simulations in a reasonable amount of time, artificial boundaries are required. At the inlet boundary typically a velocity profile with a pre-defined shape is applied (e.g. flat profile for the ascending aorta, Womersley profile for the carotid artery, etc.). In case measurements are not available for the outlet boundary conditions, these are typically determined based on the dimensions of the vessels and scaling laws (Kassab 2006). If patient-specific measurements are available (PC-MRI, US) these may be applied directly at the inlet/outlet boundary conditions, but if rigid wall simulations are performed, a correction of phase differences and attenuations, which occur naturally in the cardiovascular system, need to be performed. Remarkable advancements have been made in the development of multiscale modeling, whereas the artificial boundary conditions of the three-dimensional models are given by one- or zero-dimensional models (Blanco et al. 2007; Lagana et al. 2005; Migliavacca et al. 2006). The usage of reduced-order models comes with the requirement of defining the parameters of these models. Thus, in case of lumped parameter models the resistances, compliances and inertances need to be determined. Subsequently, numerical optimization methods may be applied to adapt the values of the parameters in the reduced order models so to match patient-specific measurements (of pressure and/or flow rate).

An important aspect in the hemodynamic modeling is the extraction of relevant information from large quantities of data generated by the simulation. Most of the studies have focused on the quantities defined at the arterial walls, which are considered to be the most important ones for cardiovascular pathologies. The average wall shear stress and the oscillatory shear index still represent the most often analyzed quantities in hemodynamic studies. Besides these, volumetric quantities are also of great interest. Thus, arterial sections with increased blood flow velocities, with recirculation zones or with reversed flow can be determined. To be able to trust the results of hemodynamic simulations, the numerical methods need to be verified and the mathematical models need to be validated (are the equations solved right, and are the right equations being solved?). Very few studies compare the numerical results with well-known analytical solutions (like the Womersley solution in case of oscillating flow in rigid or elastic tubes (Womersley 1955a, b). A comparison between numerical and experimental results with very good results was described in (Ku et al. 2005). The *in vivo* validation of hemodynamic models is much more challenging, since the experimental setup is difficult to control and data acquisition is more cumbersome. A comparison of CFD results with *in vivo* acquired MRI data was performed in Boussel et al. (2009) for cerebral aneurysms.

Significant advancements have also been made in the multiscale modeling of the coronary circulation. A patient-specific detailed model has been introduced in Kim et al. (2010), whereas a time-varying elastance model is applied for determining the boundary condition at the inlet of the aorta. For the terminal branches of the coronary circulation specialized lumped parameter models are employed, which take into account the influence of the heart contractions on the coronary blood flow. These contractions lead to a low flow during systole and large flow during diastole. The flow rates, pressures and the wall shear stress in the coronary circulation have been analyzed in this study.

Another multiscale model of the coronary circulation has been introduced in Sankaran et al. (2012). The proposed model has been employed for studying coronary artery bypasses in case of severely stenosed arteries.

1.4 Parallel Processing in Multiscale Hemodynamic Modeling

The main goal of parallel processing is to reduce the execution time. An application is a good candidate for parallel processing if it processes large quantities of data, if it performs a large number of iterations, or both. The process of parallel programming is typically based on the following steps:

- Problem decomposition;
- Selecting the algorithm;
- Implementation;
- Optimization of execution speed.

Conceptually, the identification of parallelizable components in complex applications is simple, but in practice this may become a challenging task. One key aspect is to define the activities to be performed by each work unit (typically each thread), so that the inherent parallelism is properly exploited. The theoretical speed-up that can be achieved through parallelization is determined by the portion of the application which can be safely parallelized. The speed-up is given by Amdahl's law:

$$S = \frac{1}{\left(1 - P\right) + \frac{P}{N}}, \tag{1.1}$$

where *S* is the execution time speed-up, *P* is the fraction of the program which can be parallelized, and *N* is the number of processors or threads which perform the parallel operations. The higher the value of *N* the lower is the fraction *P*/*N*, and the higher will *S* be. Nevertheless, the most important factor for the parallelization is *P*, the value of which should as close as possible to one.

To perform a successful parallelization, the following aspects need to be clarified:

- The architecture of the hardware: memory organization, data locality, caching, memory bandwidth, and the execution architecture:
  - SIMT (Single Instruction Multiple Thread): a group of threads execute the same instruction, but different groups of threads can execute different instructions—architecture employed in case of graphics processors (Rodrigues et al. 2008);
  - SPMD (Single Program Multiple Data): the same program is executed in multiple instances on multiple processors—architecture used in case of classical multi-core processors;
  - SIMD (Single Instruction Multiple Data): all threads execute the same instructions at each time point—architecture used in case of vectorial processors;
- Programming models and compilers: parallel execution models, types of available memory, data arrangement in memory, etc.
- Implementation technqiues: the transformation/replacement of sequential algorithms with parallel algorithms, understanding the scalability and the implications of the bandwidth for each algorithm, etc.;
- Extensive domain know-how: numerical methods, models, accuracy requirements, etc.

The currently most used parallel processor is the Graphics Processing Unit (GPU). Its popularity is given on one hand by the relatively low cost and, on the other hand, by the ease of programming (Kirk and Hwu 2010).

References

Acevedo-Bolton G, Jou LD, Dispensa BP, Lawton MT, Higashida RT, Martin AJ, Young WL, Saloner D (2006) Estimating the hemodynamic impact of interventional treatments of aneurysms: numerical simulation with experimental validation: technical case report. Neurosurgery 59:429–430

Antiga L, Piccinelli M, Botti L, Ene-Iordache B, Remuzzi A, Steinman DA (2008) An image-based modeling framework for patient-specific computational hemodynamics. Med Biol Eng Comput 46:1097–1112

Blanco PJ, Feijoo RA, Urquiza SA (2007) A unified variational approach for coupling 3D–1D models and its blood flow applications. Comput Methods Appl Mech Eng 196:4391–4410

Boussel L, Rayz V, McCulloch C, Martin A, Acevedo-Bolton G, Lawton M, Higashida R, Smith WS, Young WL, Saloner D (2008) Aneurysm growth occurs at region of low wall shear stress: patient-specific correlation of hemodynamics and growth in a longitudinal study. Stroke 39:2997–3002

Boussel L, Rayz V, Martin A, Acevedo-Bolton G, Lawton MT, Higashida R, Smith WS, Young WL, Saloner D (2009) Phase-contrast magnetic resonance imaging measurements in intracranial aneurysms: in vivo of flow patterns, velocity fields, and wall shear stress: comparison with computational fluid dynamics. Magn Reson Med 61:409–417

Cebral JR, Castro MA, Burgess JE, Pergolizzi RS, Sheridan MJ, Putman CM (2005) Characterization of cerebral aneurysms for assessing risk of rupture by using patient-specific computational hemodynamics models. Am J Neuroradiol 26:2550–2559

Chung TJ (2002) Computational fluid dynamics. Cambridge University Press, Cambridge

Formaggia L, Nobile F, Quarteroni A, Veneziani A (1999) Multiscale modeling of the circulatory system: a preliminary analysis. Comput Vis Sci 2:75–83

Formaggia L, Gerbeau GF, Nobile F, Quarteroni A (2001) On the coupling of 3D and 1D navier-stokes equations for flow problems in compliant vessels. Comput Methods Appl Mech Eng 191:561–582

Gijsen FJ, Wentzel JJ, Thury A, Lamers B, Schuurbiers JC, Serruys PW, van der Steen AF (2007) A new imaging technique to study 3-d plaque and shear stress distribution in human coronary artery bifurcations in vivo, J Biomechanics 40:2349–2357

Groen HC, Gijsen FJ, van der Lugt A, Ferguson MS, Hatsukami TS, van der Steen AF, Yuan C, Wentzel JJ (2007) Plaque rupture in the carotid artery is localized at the high shear stress region: a case report. Stroke 38:2379–2381

Hassan T, Ezura M, Timofeev EV, Tominaga T, Saito T, Takahashi A, Takayama K, Yoshimoto T (2004) Computational simulation of therapeutic parent artery occlusion to treat giant vertebrobasilar aneurysm. Am J Neuroradiol 25:63–68

Jesperson DC (2009) Acceleration of a CFD code with a GPU. NAS technical report NAS-09-003

Kassab GS (2006) Scaling laws of vascular trees: of form and function. Am J Physiol Heart Circ Physiol 290:894–903

Kim HJ, Vignon-Clementel IE, Coogan JS, Figueroa CA, Jansen KE, Taylor CA (2010) Patient-specific modeling of blood flow and pressure in human coronary arteries. Ann Biomed Eng 38:3195–3209

Kirk D, Hwu WM (2010) Programming massively parallel processors: a hands-on approach. Elsevier, London

Krams R, Wentzel JJ, Oomen JA, Vinke R, Schuurbiers JC, de Feyter PJ, Serruys PW, Slager CJ (1997) Evaluation of endothelial shear stress and 3D geometry as factors determining the development of atherosclerosis and remodeling in human coronary arteries in vivo. combining 3D reconstruction from angiography and ivus (angus) with computational fluid dynamics. Arterioscler Thromb Vasc Biol 17:2061–2065

Ku JP, Elkins CJ, Taylor CA (2005) Comparison of CFD and MRI flow and velocities in an in vitro large artery bypass graft model. Ann Biomed Eng 33:257–269

Lagana K, Balossino R, Migliavacca F, Pennati G, Bove EL, de Leval MR, Dubini G (2005) Multiscale modeling of the cardiovascular system: application to the study of pulmonary and coronary perfusions in the univentricular circulation. J Biomech 38:1129–1141

Migliavacca F, Balossino R, Pennati G, Dubini G, Hsia TY, de Leval MR, Bove EL (2006) Multiscale modelling in biofluidynamics: application to reconstructive paediatric cardiac surgery. J Biomech 39:1010–1020

Milner JS, Moore JA, Rutt BK, Steinman DA (1998) Hemodynamics of human carotid artery bifurcations: computational studies with models reconstructed from magnetic resonance imaging of normal subjects. J Vasc Surg 28:143–156

Perktold K, Rappitsch G (1995) Computer simulation of local blood flow and vessel mechanics in a compliant carotid arterybifurcation model. J Biomech 28(7):845–856

Quarteroni A, Ragni S, Veneziani A (2001) Coupling between lumped and distributed models for blood flow problems. Comput Vis Sci 4:111–124

Rodrigues CI, Stone J, Hardy D, Hwu WW (2008) GPU acceleration of cutoff-based potential summation. In: Proceedings of the ACM international conference on computing frontiers, Ischia, Italy, pp 1–10

Sankaran S, Moghadam ME, Kahn A, Tseng E, Guccione J, Marsden A (2012) Patient-specific multiscale modeling of blood flow for coronary artery bypass graft surgery. Ann Biomed Eng 40:2228–2242

Shojima M, Oshima M, Takagi K, Torii R, Hayakawa M, Katada K, Morita A, Kirino T (2004) Magnitude and role of wall shear stress on cerebral aneurysm: computational fluid dynamic study of 20 middle cerebral artery aneurysms. Stroke 35:2500–2505

Slager CJ, Wentzel JJ, Schuurbiers JC, Oomen JA, Kloet J, Krams R, von Birgelen C, van der Giessen WJ, Serruys PW, de Feyter PJ (2000) True 3-dimensional reconstruction of coronary arteries in patients by fusion of angiography and IVUS (ANGUS) and its quantitative validation. Circulation 102:511–516

Slager CJ, Wentzel JJ, Gijsen FJ, Thury A, van der Wal AC, Schaar JA, Serruys PW (2005) The role of shear stress in the destabilization of vulnerable plaques and related therapeutic implications. Nat Clin Pract Cardiovasc Med 2:456–464

Steinman DA (2002) Image-based computational fluid dynamics modeling in realistic arterial geometries. Ann Biomed Eng 30:483–497

Steinman DA, Taylor CA (2005) Flow imaging and computing: large artery hemodynamics. Ann Biomed Eng 33:1704–1709

Tang BT, Cheng CP, Draney MT, Wilson NM, Tsao PS, Herfkens RJ, Taylor CA (2006) Abdominal aortic hemodynamics in young healthy adults at rest and during lower limb exercise: quantification using image based computer modeling. Am J Physiol Heart Circ Physiol 291:668–676

Tang D, Yang C, Mondal S, Liu F, Canton G, Hatsukami TS, Yuan C (2008) A negative correlation between human carotid atherosclerotic plaque progression and plaque wall stress: in vivo MRI-based 2D/3D FSI models. J Biomech 41:727–736

Taylor CA, Draney MT (2004) Experimental and computational methods in cardiovascular fluid mechanics. Annu Rev Fluid Mech 36:197–231

Taylor CA, Hughes TJR, Zarins CK (1996) Computational investigations in vascular disease. J Comput Phys 10:224–232

van de Vosse FN, Stergiopulos N (2011) Pulse wave propagation in the arterial tree. Annu Rev Fluid Mech 43:467–499

Vignon-Clementel I, Figueroa CA, Jansen KE, Taylor CA (2006) Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries. Comput Methods Appl Mech Eng 195:3776–3796

Vignon-Clementel I, Figueroa CA, Jansen KE, Taylor CA (2010) Outflow boundary conditions for 3D simulations of non-periodic blood flow and pressure fields in deformable arteries. Comput Methods Biomech Biomed Engin 13(5):1–16

Wang KC, Dutton RW, Taylor CA (1999) Improving Geometric model construction for blood flow modeling. J Med Biol Eng 18:33–39

Wendt JF (2009) Computational fluid dynamics: an introduction, 3rd edn. Springer, Berlin

Wilson NM, Arko FR, Taylor CA (2005) Predicting changes in blood flow in patient-specific operative plans for treating aorto-iliac occlusive disease. Comput Aided Surg 10:257–277

Womersley JR (1955a) Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. J Physiol 127:553–563

Womersley JR (1955b) Oscillatory motion of a viscous liquid in a thin-walled elastic tube. I. The linear approximation for long waves. Philos Mag 7:199–221

The Cardiovascular System and the Coronary Circulation

2

Elisabeta Badila, Lucian Calmac, Diana Zamfir,
Daniela Penes, Emma Weiss, and Vlad Bataila

Abstract

The cardiovascular system represents the pathway for the delivery and distribution of essential substances to the tissues, as well as for the removal of metabolic waste. Furthermore, it plays an important role in homeostatic mechanisms, such as regulation of body temperature, humoral communication throughout the body and adjustments of oxygen and nutrient supply under different physiologic states. Its main component is the heart, which is composed of two synchronous pumps. A crucial role is played by the coronary arteries, which supply the myocardium and the specialized electrical tissue of the heart with oxygenated blood. Coronary artery disease is the most common cause of morbidity and mortality in developed countries, and is diagnosed using medical imaging techniques and invasive functional assessment procedures. Finally, we also provide insights into aortic pathologies, which comprise a wide spectrum of arterial diseases, with a special focus on coarctation of the aorta.

E. Badila • E. Weiss (✉)
Internal Medicine, Clinical Emergency Hospital,
Calea Floreasca nr. 8, Bucharest 014461, Romania

Internal Medicine, "Carol Davila" University of Medicine and Pharmacy,
Dionisie Lupu nr. 37, Bucharest 020021, Romania
e-mail: emmaweissmd@gmail.com

L. Calmac • V. Bataila
Interventional Cardiology, Clinical Emergency Hospital,
Calea Floreasca nr. 8, Bucharest 014461, Romania

D. Zamfir
Laboratory of Echocardiography, Clinical Emergency Hospital,
Calea Floreasca nr. 8, Bucharest 014461, Romania

D. Penes
Internal Medicine, Clinical Emergency Hospital,
Calea Floreasca nr. 8, Bucharest 014461, Romania

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_2

2.1 The Cardiovascular System

The invasive evaluation of cardiovascular hemodynamics represents the final frontier when elucidating cardiovascular physiopathology in patients in whom noninvasively obtained data remain inconclusive or require firm confirmation of the clinical aspects.

The circulatory system represents the pathway for delivery and distribution of essential substances to the tissues, removing metabolic waste. The system also plays an important role in homeostatic mechanisms, such as regulation of body temperature, humoral communication throughout the body and adjustments of oxygen and nutrient supply in different physiologic states (Wier and Pappano 2013).

The cardiovascular system provides these functions through a pump (the heart), connected to a series of tubules (large and small arteries), capillaries (an extensive system of thin vessels that confer the capacity for rapid exchanges between different compartments) and veins. The heart consists functionally of two synchronized pumps: the right and the left side of the heart. The right side of the heart consists of right atrium, a chamber receiving deoxygenated blood from the veins (through the superior and inferior vena cava) and the right ventricle, whose function is to pump the blood through the lungs for the rapid exchange of O_2 and CO_2 (also called *pulmonary circulation*) through respiration. The left side of the heart is structured contains the left atrium (receiving oxygenated blood from the lungs) and the left ventricle, whose role is to propel blood and nutrients to the entire body through the aorta—the largest artery in the body (the *systemic circulation*). There are four valves inside the heart (one for each chamber), which ensure unidirectional flow inside and outside of the heart. The pumping role is supported by a muscular tissue (myocardium) which provides electrical and mechanical activation, generating the contractile force necessary for blood distribution to the various organs.

The amount of blood flow propelled by the left ventricle in 1 min is known as the cardiac output (CO). It represents an indicator for the heart's ability to support the needs of the body in a minute. Cardiac output is determined by the product of stroke volume (amount ejected by the heart) and heart rate (beats per minute). In resting conditions, the cardiac output measures approximately 5 L/min, depending on body size, and this value varies largely (being able to increase fivefold during exercise) to meet the body demand (Smith and Fernhall 2011).

Vessels are responsible for delivering oxygen and nutrients throughout the body. Therefore, the circulation systems of most major organs are parallel, but one major exception is the liver, that receives a large amount of blood from the venous circulation of the intestinal tract, which drains into the portal system and to the liver. The parallel arrangement of major vascular beds prevents blood flow changes in one organ from significantly affecting blood flow in other organs (Klabunde 2012).

It is important to note that the large distribution of blood to an organ or a system is controlled by hemodynamic rules. The degree of vasoconstriction or vasodilation in the supplying arterioles of organs is influenced by a series of local factors and not by the output of the heart itself. Therefore, vessels are dynamic structures that permanently vary in diameter to modify the amount of supplied blood in order to meet the demand of oxygen and nutrients (Smith and Fernhall 2011).

Systemic vascular resistance, also called total peripheral resistance represents the resistance to blood flow derived from all the systemic vasculature, except for the pulmonary vasculature. Systemic vascular resistance is determined by changes in vascular diameters and by the blood viscosity. An increase in systemic vascular resistance through vasoconstriction, in response to sympathetic stimulation, depends on the degree of the sympathetic activation, the responsiveness of the vasculature and the number of vascular beds being involved (Klabunde 2012).

During the cardiac cycle the diastolic ventricular filling is augmented by the atrial systole to achieve a certain end-diastolic volume. This is followed by an isovolumic contraction, leading to aortic valve opening (along with pulmonary valve opening in the right heart), followed by stroke volume ejection. Meanwhile, the atrial pressure increases progressively during ventricular systole, as blood continues to fill in the atria and the atrio-ventricular valves are closed. After reaching the point of end-systolic volume of both ventricles, there is a period of isovolumic relaxation, accompanied by the opening of the atrio-ventricular valves (mitral and tricuspid valve) and the passive filling of the heart chambers. In a normal resting state, the diastole occupies approximately two-thirds of the cardiac cycle, whereas both systole and diastole shorten with increased heart rate (Martin and Stephens 2013).

Arterial pressure varies during the cardiac cycle, from a minimum value, called the diastolic pressure, to a maximal one, known as systolic pressure. The difference between the two values is the pulse pressure. It is an important characteristics quantifying the distensibility of large arteries.

To ensure its contractile function, the myocardium requires a proper vascular network. The amount of oxygen used by the working myocardium is influenced by coronary flow and oxygen extraction rate. It is essential that the oxygen supply is continuous in order to ensure the metabolic substrate necessary to convert it to chemical energy (adenosine-triphosphate molecules). Coronary vasculature consists of the two main coronary arteries—right and left coronary arteries; the right coronary artery serves the right atrium and right ventricle; the left coronary artery divides shortly after its origin in the left anterior descending artery and the circumflex artery—ensuring blood supply to the rest of the myocardium. After passing the coronary bed, blood flows towards the coronary sinus through the coronary veins. The myocardial contraction delivers the necessary pressure (equal to the aortic pressure) needed for proper myocardial perfusion.

Coronary flow is influenced by the coronary vascular resistance. Sympathetic activity increases coronary flow, thus increasing rate and force of myocardium contraction. With the increased rate, diastolic filling of coronary arteries is shortened, thus limiting myocardial flow. Consequently, vasoconstriction increases oxygen extraction and thus myocardial metabolism is accelerated.

To function as a pump, the heart has intrinsic electrical activity generated via its excito-conductor tissue, which functions in a sequence that determines the uniform and rhythmical activation of myocardial fibers. This specialized conduction tissue is made of specific cells which have the capacity to generate electrical impulses which are further transmitted to the muscular cells. Under normal conditions, the electrical impulse originates in the sinoatrial node, connected to the atrial myocardium, and is rapidly distributed through specialized internodal pathways in the atria to the atrioventricular node, in order to further progressively stimulate the ventricles, cell by cell, through the bundle of Hiss and the Purkinje network. This electrical stimulation leads to the contraction of all muscle fibers, arranged so that during systole they will ensure longitudinal, circumferential and radial motion to efficiently pump the blood into the aorta (Smith and Fernhall 2011).

2.2 The Anatomy of the Coronary Arteries

The heart, an ovoid organ, situated between the lungs in an oblique position, is divided into four chambers: two atria (left atrium, the most superior and posterior, and right atrium right antero-inferior) and two ventricles (left ventricle positioned posteriorly, lined by the right ventricle which is the most anterior of the chambers). Each chamber has a different muscular wall and is separated by two valves: mitral valve (separating the left atrium by the left ventricle) and tricuspid valve (which separates right atrium by the right ventricle).

Two important arteries derive from the two ventricles: the pulmonary trunk, from the right ventricle, which transports deoxygenated blood to the lungs, and the aorta, transporting oxygenated blood to the whole body.

From the aorta, at the level of the aortic cusps, arise the coronary arteries, which supply the myocardium and the specialized electrical tissue of the heart with oxygenated blood. In a normal coronary anatomy, the coronary arteries arise perpendicularly from the aorta below the sino-tubular ridge, at the point between the sinuses of Valsalva and the tubular aorta. At this level, usually two coronary ostia are formed, one of the right coronary artery (RCA), originating from the right aortic sinus, and one of the left main coronary artery (LM), which originates from the left aortic sinus. The latter is subdivided in the left anterior descending artery (LAD) and left circumflex artery (LCX) (Fig. 2.1).

Due to the high variability of the coronary anatomy and distribution (Oudkerk and Reiser 2008) that varies in presence, size, shape and length, a general anatomy will be described first (Dewey 2009). Both the RCA and the LCA (along with its branches) decrease in diameter along their lengths (Anderson et al. 2013).

Coronary heart disease (CHD) is the leading cause of death in the world, according to a 2011 report of the World Health Organization (WHO 2011). Even though in some developed countries (Arciero et al. 2004), the CHD mortality decreased, it still remains responsible as the main cause of death for about a third of all deaths in adults (Ford et al. 2007; Lyoyd-Jones et al. 2010). In the past decades great efforts towards population awareness have been made, concentrating on finding, diagnosing and lowering all risk factors for CHD. Primary prevention, defined as lowering risk factors such as smoking, hypertension, high cholesterol, diabetes and a sedentary lifestyle, plays a leading role in lowering the incidence of CHD. Great research, with perhaps the most significant impact in lowering the morbi-mortality in the world after the invention of antibiotics, has been done towards the acute phase treatment and secondary prevention. Besides an extensive array of medication, great progress has been made in the interventional, device-oriented and surgical treatment.

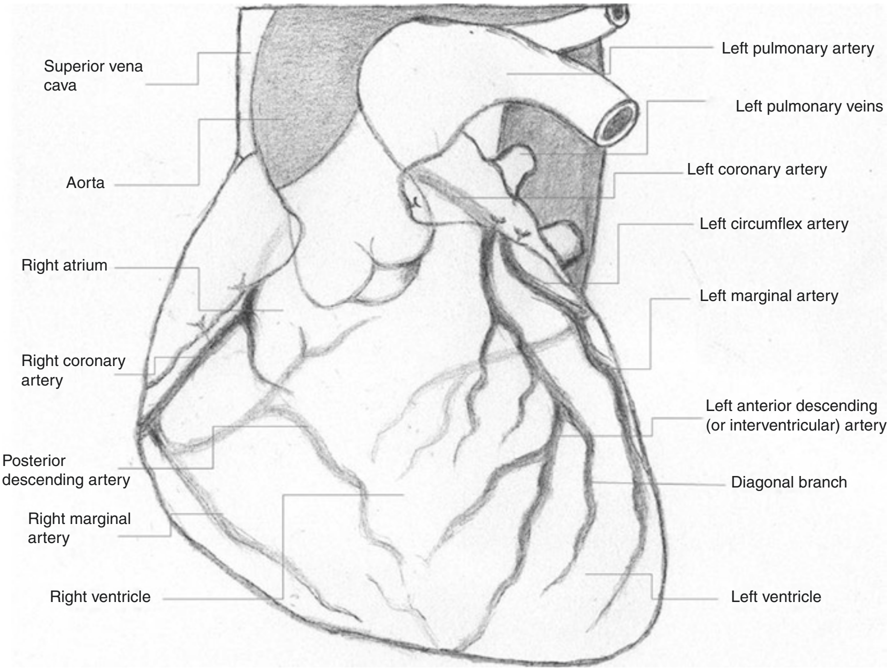


Fig. 2.1 The coronary anatomy

2.2.1 Right Coronary Anatomy

The RCA originates from the right sinus of Valsalva, below the plane of the left coronary artery ostium. It travels between the right atrial appendage and right ventricular outflow tract (RVOT) and continues straight along the right atrioventricular groove. Along its course, the proximal branches of the RCA are the conus artery, the right ventricular branch and the acute marginal branch (AM). The conus branch, often the first branch, supplies the infundibulum. In half of the population, the conus artery arises from the right coronary ostium or from the first millimeters of the RCA. The conus artery describes an anterior and upward route over the RVOT towards the LAD. In patients with LAD occlusion, it may serve as a source of collateral blood supply.

In the other half of the population, the conus artery arises from its own ostium in the right aortic sinus, directly from the aorta, just above the right coronary ostium. Furthermore, in 60% of the cases, RCA provides the sinoatrial node branch.

The RCA supplies branches to the right atrium. Furthermore, the acute marginal branches that arise from the mid-segment of the RCA follow the right ventricle and supply the anterior free wall. They branch off the RCA, detach in an acute angle, vary in size and number and are labeled from proximal to distal: AM1, AM2, AM3 and so forth.

In cases of LAD occlusion, the AM may serve as sources of collateral circulation. The distal portion of the RCA starts just below the origin of the AM branch and continues straight along the diaphragmatic surface of the heart, generating the posterior descending artery (PDA) in most individuals (85%). The PDA supplies the posterior third of the interventricular septum with septal perforators. After the origin of the PDA, in case of right dominancy (50–60% of cases), the RCA goes beyond crux cordis in the left atrioventricular sulcus and terminates into the postero-lateral artery. The last branch supplies the posterior part of the left ventricle.

When RCA supplies the atrioventricular (AV) nodal branch, it branches off at the crux cordis going upward towards the AV node.

The posteromedial papillary muscle is often supplied only by the PDA, being vulnerable to ischemic dysfunction. A myocardial infarction involving the PDA is more likely to cause mitral regurgitation.

2.2.2 Left Coronary Anatomy

2.2.2.1 Left Main Coronary Artery

The LM originates just above the left sinus of Valsalva and continues behind the pulmonary trunk. Its diameter varies among individuals, between 3 and 6 mm (Oudkerk and Reiser 2008) in diameter and 0–15 mm in length (Dewey 2009).

The LM bifurcates into two large branches: the left anterior descending artery and the left circumflex artery. In one third of the population, the LM ends as a trifurcation generating an intermediate branch (ramus intermedius) that arises between the LAD and the LCX, suppling the myocardium between diagonal and marginal territories (Oudkerk and Reiser 2008). This ramus intermedius can be considered as an obtuse marginal branch or as a diagonal branch, depending on its course along the left ventricle (Dewey 2009). Also, the LM may be absent, in which case the LAD and the LCX then have separate origins (Oudkerk and Reiser 2008).

2.2.2.2 Left Anterior Descending Artery

The LAD courses behind the pulmonary trunk and goes forward between the left atrial appendage and the pulmonary trunk to the anterior interventricular groove (Oudkerk and Reiser 2008). It represents the direct continuation of the left main artery. The LAD and the LM form a curve that resembles a reversed "S" shape (Mukherjee et al. 2010). In its trajectory, it gives rise to diagonal and septal perforating branches.

The first branch that emerges from the LAD is the left conal artery, which supplies the infundibulum. The septal perforating branches arise from the LAD in a 90° angle (Oudkerk and Reiser 2008). In most individuals, the blood supply to the anterior interventricular septum comes from the LAD (Mukherjee et al. 2010). These branches are perforating the interventricular septum displaying a large variation in number, distribution and size; they connect together and with the branches derived from the RCA—posterior descending branch, creating a network of potential collateral channels. All septal branches of the LAD supply 2/3 of the anterior interventricular septum.

Diagonal branches that supply the anterolateral free wall of the left ventricle and the anterolateral papillary muscle also vary widely in number and size. In more than 90% of cases, one to three diagonal branches originate from the LAD and less than

1% have no diagonal branches at all. When diagonal branches are not visible, acquired atherosclerotic occlusion of the diagonal branches is highly probable, especially among patients with echocardiographic movement abnormalities of the anterolateral left ventricle.

Furthermore, LAD goes straight to the apex and terminates along the diaphragmatic aspect of the left ventricle in two thirds of cases. In one third of cases, the distal segment of the LAD ends before reaching the cardiac apex, without reaching the diaphragmatic surface, in which case a dominant RCA is cited, having a posterior descending branch larger and longer than usual which then supplies the apex (Oudkerk and Reiser 2008).

2.2.2.3 The Circumflex Artery

LCX arises from the LM posteriorly, passing by the left atrial appendage and coursing posteriorly to the left AV groove, on the obtuse margin of the heart (Oudkerk and Reiser 2008). It originates from the LM at an approximately right angle, but may have a greater or lesser degree of angulation from the left main at its origin in some patients. As it passes by the atrioventricular groove, it circumscribes the mitral valve annulus, giving rise to the obtuse marginal branches (Mukherjee et al. 2010).

Left atrial branches may arise from the LCX and supply in 40% of the patients the sinoatrial node (Mukherjee et al. 2010). Usually, three obtuse marginal branches are described: the anterolateral branch, the obtuse marginal branch—being the largest of them all—and the posterolateral branch (Oudkerk and Reiser 2008). They supply the posterolateral wall (lateral free wall) of the left ventricle (Dewey 2009). The LCX also generates one or two left atrial circumflex branches that supply the lateral and posterior walls of the left atrium.

In few cases (10%), in left dominancy, the LCX gives rise to the posterior descending artery and the AV nodal arterial branch.

2.2.3 Determination of Dominance

The artery that crosses the crux cordis and gives rise to the PDA, the AV nodal branch and the posterolateral artery is considered to be the dominant coronary artery. A co-dominancy or balanced circulation can be described, when the PDA arises from the RCA and the postero-lateral artery arises from the LCX. Seventy percent of patients are right-dominant, 20% are co-dominant and 10% are left-dominant (Oudkerk and Reiser 2008; Mukherjee et al. 2010). In case of a left-dominant circulation, the RCA is small and does not supply blood to the left ventricular myocardium (Dewey 2009).

2.2.4 Frequent Coronary Artery Variants

Apart from variations of normal anatomy, as those presented above, other variations may be encountered, such as myocardial bridging or variability in the course and origin of the coronary arteries.

Myocardial bridging, which describes the situation when an arterial segment of the coronary artery passes through the myocardium, may be seen in almost 5% of angiography interventions. The most frequent cases of bridging are confined to the LAD or diagonal branches. It is a condition in which, during systole, the muscular bridge running over the artery contracts, leading to the temporary compression of the coronary artery involved; during diastole, the caliber goes back to normal and, in most cases, this anatomic variant does not cause ischemia. In clinical practice, there are cases associated with ischemia (Dewey 2009) secondary to myocardial bridging.

Intracoronary artery continuity or “coronary arcade” describes a rare condition of the coronary circulation in which communications develop between the distal segment of the LCX and the RCA in the posterior atrioventricular sulcus, and between the LAD and the PDA in the interventricular groove. These connections occur in the absence of obstructive coronary disease and they are distinct from coronary collaterals (Mukherjee et al. 2010).

Two most frequent anomalies are an RCA abnormally originating from the LM or the left sinus of Valsalva, or an LCX with an anomalous origin from the RCA or from the right sinus of Valsalva. In the first variant, the course of the RCA is anterior, between the aorta and the pulmonary trunk. This anatomic variant is called “malignant course”, because these patients have a high risk for exercise-induced ischemia and sudden death, as at the time of the effort, the arterial segments contain more blood which causes compression and induces ischemia (Dewey 2009). Another anomaly is that of a LM arising from the right sinus of Valsalva which then courses between the aorta and the pulmonary trunk, this being the most dangerous variety of an anatomic anomaly (Mukherjee et al. 2010; Anderson et al. 2013). On the other hand, an LCX originating from the RCA or right sinus of Valsalva, coursing posteriorly to the aorta and entering in its proper location, in the left atrioventricular groove, is actually a benign condition which never causes ischemia (Dewey 2009).

2.2.5 Angiography of the Coronary Vessels

The most frequent indication for a coronary evaluation is diagnosing coronary atherosclerosis, based on which the clinician can decide upon the most appropriate treatment strategy (bypass surgery, angioplasty or medical therapy). The indications for angiography are well established and should be based on clinical guidelines. Coronary angiography is to be performed in symptomatic patients with noninvasive evidence of myocardial ischemia.

In patients with unstable angina, coronary angiography will be performed after initiation of intensive drug therapy with the aim of stabilizing the atherosclerotic plaque. The intervention will usually be performed within the first 6 weeks from the acute episode.

Acute myocardial infarction patients currently undergo primary angioplasty. It is currently unclear whether asymptomatic patients benefit from coronary angiography post-myocardial infarction.

Other indications for a coronary angiography are atypical chest pain syndromes in patients with cardiovascular risk factors, ventricular arrhythmias, heart failure of unclear etiology (Moscucci 2014).

2.3 Coronary Artery Disease

2.3.1 Pathophysiology of Coronary Artery Disease

Coronary artery disease (CAD) is the most common cause of morbidity and mortality in developed countries. It represents the final common pathway after combining genetic predisposition with lifestyle factors. This concept includes diseases that result from atherosclerotic changes of the coronary vessels, generating a whole spectrum of coronary plaques, from stable plaques (lipid-poor, thick fibrous cap) to unstable ones (lipid-rich, thin fibrous cap) (Ashley and Niuebauer 2004).

The atherosclerotic process develops early in life, but it presents large variations among individuals, with an unpredictable speed of progression (Theroux 2011). The severity of events correlates with the type of coronary plaques: the more unstable, the more likely to rupture, generate pro-thrombotic events, release vasoconstrictive factors and lead to partial or complete occlusion of the affected coronary artery (Ashley and Niuebauer 2004). Temporary occlusion leads to (possibly reversible) ischemia (insufficient blood supply to the affected tissue), but permanent occlusion is associated with transmural myocardial infarction (irreversible ischemia followed by death of affected tissue).

The atherosclerotic process often develops in multiple vascular beds (coronary arteries but also cerebral, carotid and peripheral arteries). Patients affected by atherosclerosis therefore manifest symptoms characteristic of systemic atherosclerotic disease. The REACH Registry included 68,000 patients from 44 countries across six major regions; the study enrolled patients aged over 45 with at least three atherosclerotic risk factors or documented coronary artery disease, cerebrovascular disease or peripheral artery disease. This research had found that as much as 30% of coronary artery disease patients had systemic atherosclerotic disease (disease in more than one arterial bed) (Ohman et al. 2006).

The most important risk factors for developing and enhancing atherosclerosis are dyslipidemia, smoking, arterial hypertension, diabetes mellitus, obesity, immune mechanisms or hemodynamic factors, all of them generating injury and consequently dysfunctionality of the endothelium—the interior lining of the arteries.

2.3.1.1 Atherosclerotic Risk Factors

Dyslipidemia is a condition characterized by an abnormally high level of circulating lipids and represents a major risk factor for developing atherosclerosis. Many studies have found that in countries with a high consumption of saturated fat and a high prevalence of hypercholesterolemia, there is a strong association with coronary heart disease, as opposed to populations whose diet is based on low saturated fat intake and have low serum cholesterol levels (Mediterranean areas) (Estruch et al. 2013).

Data from the Framingham Study and other cohorts have shown that the risk of coronary heart disease increases with higher total serum cholesterol levels. Particularly, elevated levels of LDL-cholesterol (low density lipoprotein-cholesterol, commonly known "bad cholesterol") are associated with an increased incidence of atherosclerosis and subsequent complications. In opposition, HDL-cholesterol (high-density lipoprotein) particles seem to protect against this process, due to their ability to transport cholesterol away from the peripheral tissues back to the liver, having antioxidant and anti-inflammatory properties (Lilly 2011).

High levels of cholesterol and its fractions (LDL-C and similar) are usually the consequence of a high fat intake diet. However, they may be secondary to an abnormal lipid metabolism, as in the case of genetic anomalies of the genes encoding LDL-receptors. These conditions are known as familial hypercholesterolemias, characterized by high plasma LDL levels from very early ages, leading to premature atherosclerosis during the first decades of life and eventually major cardiac events (myocardial infarction, stroke) in subjects as young as 30 years old.

Strategies that improve plasma lipid levels can limit the effects of atherosclerosis. Lifestyle changes have shown positive effects on the reduction of LDL levels, slowing the progression of atherosclerotic plaque (Barry et al. 2014). The current European guidelines on Cardiovascular Prevention recommend an optimal LDL cholesterol level as low as <100 mg/dL for those with high cardiovascular risk, or an even lower goal of <70 mg/dL for those at higher risk of future cardiovascular events (established cardiovascular disease, diabetes mellitus type 2 with target organ damage, moderate to severe chronic kidney disease) or other associating risk factors (Reiner et al. 2011). Diet and physical exercise comprise two important tools in the overall risk reduction.

When lifestyle modifications fail to achieve target values, pharmacologic agents should be used to improve abnormal lipid levels. The main categories of lipid-altering agents are: HMG-CoA (3-hydroxy-3-methyl-glutaryl-coenzyme A) reductase, also known as statins, niacin, fibric acid derivatives, cholesterol intestinal absorbtion inhibitors, bile acid-binding agents and, the more recently approved PCSK9 (proprotein convertase subtilisin/kexin type 9) inhibitors. The currently most effective and widely used LDL lowering therapy is statin therapy (Lilly 2011).

Smoking increases the risk of atherosclerotic disease and promotes ischemic events. The heaviest smokers carry on the highest risk for cardiovascular events. The act of smoking enhances oxidative modification of LDL-cholesterol particles, decreases the circulating level of HDL-cholesterol, promoting endothelial dysfunction and tissue hypoxia and increases platelets adhesiveness (Lilly 2011). In a recent multiethnic study, current smoking and cumulative exposure remain important modifiable determinants of cardiovascular disease (McEvoy et al. 2015). Smoking cessation can reverse or slow some of the adverse outcomes.

Arterial hypertension is a condition characterized by high blood pressure values (>140/90 mmHg) which acts as a risk factor both for the development and the progression and acceleration of atherosclerosis, coronary heart disease and major cardiovascular events. The risk increases progressively with higher pressure values (especially systolic). Hypertension may produce acceleration of the atherosclerotic process by increasing the permeability of the vessel wall to lipoproteins, augmenting hemodynamic stress and enhancing the development of foam cells (fat-laden immune cells indicative of atherosclerotic deposition building-up in plaques) (Hollander 1976). As in the case of dyslipidemias, the treatment of arterial hypertension should start with lifestyle changes (diet and exercise), and progress to pharmacological therapy when blood pressure values remain uncontrolled.

Diabetes mellitus influences atherosclerosis in such great extent that the condition is considered to be equivalent to the development of the atherosclerotic process. This is due partly to the fact that diabetes mellitus leads to the non-enzymatic glycation of lipoproteins (enhancing the uptake of LDL cholesterol by foam cells) and to a pro-thrombotic state which becomes permanent (Lilly 2011). Intense control of serum glucose levels in diabetic patients reduces the risk of vascular complications, such as myocardial infarction and stroke, and controlling associated dyslipidemia and/or hypertension lowers the overall cardiovascular risk in these patients.

Metabolic syndrome includes multiple risk factors, as: hypertension, hypertriglyceridemia, reduced HDL-cholesterol, insulin resistance syndrome and visceral obesity. These multiple conditions associate higher risk for developing atherosclerotic disease.

Sedentarism or low level of physical activity may mitigate atherogenesis in multiple ways. Exercise improves insulin sensitivity and endothelial production of nitric oxide, an endogenous vasodilator, lipid profiles and blood pressure values. Studies suggest that moderate-level activities, such as 30-minutes-a-day walking can lower cardiovascular mortality (Lilly 2011).

2.3.1.2 Pathophysiology of Coronary Artery Disease

Regulation of Coronary Blood Flow

The main physical factor influencing myocardial perfusion is the aortic pressure, which causes parallel changes in the coronary hemodynamics. In addition to this, changes in vessel caliber in response to the metabolic demands of the heart lead to changes in coronary blood flow.

Opposing this pressure is the coronary vascular resistance, a factor that influences the hemodynamics of coronary flow. In normal arteries, the resistance depends on external forces that compress the coronary artery and on factors that alter intrinsic vascular tone (Lilly 2011).

Coronary flow reserve represents the difference between maximal flow as caused by vasodilators and the flow in the physiological range.

Unlike other arterial beds where the highest blood flow occurs during systole, coronary perfusion takes place during diastole, when the small coronary arteries are no longer compressed, as they are during systole. In the case of the coronary arteries, the perfusion pressure is close to equal to the aortic diastolic pressure; conditions that decrease the aortic diastolic pressure, such as hypotension or aortic regurgitation, also decrease coronary artery perfusion pressure and therefore lower oxygen supply to the myocardium (Lilly 2011).

Sympathetic stimulation produces an increase in coronary blood flow through an increased metabolic demand and a predominance of β receptor activation. Alpha stimulation may play a role in the distribution of blood flow within the myocardium by restricting metabolically mediated flow increase and exerting an anti-steal effect (Ramanathan and Skinner 2005).

The supply of oxygen to the myocardium is determined by the oxygen content in the blood and the rate of coronary blood flow. Oxygen content is determined by the hemoglobin concentration and systemic oxygenation; the rate of coronary blood flow is proportional with the vessel's perfusion pressure and inversely correlated to coronary vascular resistance. Alterations of the balance between oxygen-supply and oxygen-demand generate areas of ischemia and eventually death of the involved tissue (infarction).

A decrease in the oxygen supply to oxygen-demand ratio (produced by a reduction in oxygen supply or an increase of oxygen demand) causing ischemia, leads to a metabolic regulation of blood flow with the release of vasodilator substances from the myocardium into the interstitial fluids, causing relaxation of the coronary vessels. The most important vasodilator agents are adenosine and nitric oxide (Pappano and Wier 2013).

Atherosclerotic plaque deposition decreases vessel lumen, thus reducing blood flow, which is the most important cause of ischemia at this site. If ischemia is present and prolonged, necrosis of affected cardiac cells develops. Myocardial ischemia produces alterations in the mechanical and electrical behavior of the heart. The mechanical impairment is not only due to lack of oxygen supply but also to the accumulation of harmful substances from the necrotic process, such as lactic acid, H+ or K+ (Pappano and Wier 2013).

Atherosclerotic Plaque Formation

The arterial wall is composed by three layers: the tunica intima (the inner layer), the tunica media (the middle layer) and the tunica adventitia (the outer layer). The inner layer consists of a thin layer of simple cells called the endothelium which forms an interface between circulating blood and its two more profound layers—the subendotelial proteoglycan-rich layer and the deeper musculo-elastic layer made of smooth muscle cells and elastic fibers (Theroux 2011).

Atherosclerosis is a chronic inflammatory proliferative disease of the intima of large and medium sized arteries. It is characterized by early retention and structural modification of atherogenic lipoproteins, recruitment of immune cells and accumulation of abundant fibrous tissue (Theroux 2011). An atherosclerotic plaque consists of a core of foam cells (immune cells which come to the site of plaque formation to address and solve what seems to them to be an inflammatory process but end-up integrating lipids and thus become lipid-engorged and die), covered by a fibrous cap (a region of the intimal layer that has become thickened and hardened as a result of collagen and elastin fibers deposition) (Ashley and Niuebauer 2004).

The first step in the atherosclerotic process is represented by the retention of apo-B containing lipoproteins, primarily LDL, to extracellular proteoglycans of the tunica intima. The accumulation of subendothelial lipoproteins is followed by aggregation and oxidation in enzymatic reactions and by oxidative radicals (leading to oxidized LDL), generating the signaling-pathway to stimulate the recruitment, differentiation and replication of monocyte-derived macrophages (immune cells), acting as a pro-inflammatory trigger (Theroux 2011).

Inflammation plays a major role in the pathogenesis and clinical expression of atherosclerosis to the extent that the disease is now considered an inflammatory one (Buja and McAllister 2007). The recruited macrophages engulf LDL-particles through scavenger receptors, giving them the above mentioned name—foam cells. Cytokines produced by the macrophages create a pro-inflammatory environment that facilitates the recruitment of more and more inflammatory cells.

At this point, the process of foam cells production is reversible if the stimuli causing their formation disappear. If the accumulation of intracellular lipids continues, it is followed by accumulation of lipids in the musculo-elastic layer, beneath the foam cells, generating intermediate lesions, also reversible. Moreover, if the conversion of the isolated lipid areas into one or more confluent lipid-rich cores goes further, they are generating the necrotic core, a permanent lesion that disrupts the structure of the inner layer.

The lipid-rich core is separated from blood flow by a fibrous cap, a structure that plays an important role in the pathology and clinical consequences of the atherosclerotic process. This structure consists of fibrous tissue and calcifications, without any lipid particles (Theroux 2011). An excess of degradative enzymes leads to degradation and thinning of the fibrous capsule.

The two components of plaque, the inner lipid layer and the fibrous cap, are extremely important as they determine the risk for plaque complications such as rupture or erosion which lead to thrombus formation and complete artery occlusion. A stable plaque (poor lipid core, rich calcified fibrous cap) is less likely to rupture, while an unstable plaque (rich lipid core, thin fibrous cap) is highly likely to rupture or erode.

In the initial phase, atherosclerosis is a focal disease, generating plaques adjacent to branch points in areas of low-velocity flow and low shear stress areas, near high shear stress ones; this flow pattern promotes endothelial dysfunction (Gimbrone 2010). Areas of severe atherosclerosis of the coronary system include the proximal left anterior descending coronary artery and the proximal and distal right coronary arteries.

Atherosclerotic disease leads to extensive remodeling of the arteries, causing dilatation of the vessels initially without influencing lumen diameter; luminal narrowing occurs, however, but only in more advanced stages of the disease (Willerson and Holmes 2015).

2.3.1.3 Pathophysiology of Stable Angina (Stress Induced Angina)

With time plaque grows continuously, reducing the arterial lumen progressively and leading to a point where ischemia sets in, causing what is clinically known as angina pectoris—typical chest pain with exertion. If a stenosis never produces flow-limitation, it will never produce symptoms (Theroux 2011). Flow obstruction is determined by a series of factors such as: plaque size, vasospasm (vasoconstriction) and vessel wall remodeling. Expansive or positive remodeling refers to the initial preservation of vessel lumen and is more common than the opposite constrictive remodeling (negative remodeling), that reduces vessel lumen area. Coronary artery remodeling is unrelated to sex or smoking and it is plaque specific (Varnava and Davies 2001). Also, multiple plaques at different stages of progression and of variable morphology coexist within the same patient or even artery (Koskinas et al. 2010). In addition to arterial remodeling, local paradoxical exercise-induced vasospasm may influence the blood flow, causing temporary obstruction and symptoms specific to stable angina (Gage et al. 1986).

The cut-off point of vessel lumen reduction at which coronary blood flow at rest becomes significantly influenced is 80%. The types of stenosis which cause such lumen reductions are highly calcified fibroatheromas or fibrous plaques (Ehara et al. 2004). A small percent of culprit plaques in patients with stable angina present residual thrombi that suffer recanalization with time (Mann et al. 1998). The severity of the stenosis correlates with the number of plaque events (Burke et al. 2001) and majority of plaque ruptures are not lethal, being clinically silent (Davies et al. 1989).

2.3.1.4 Pathophysiology Behind the Acute Coronary Syndromes

Acute coronary syndromes (unstable angina and acute myocardial infarction with or without ST segment elevation) are produced by a luminal thrombus attached to an atherosclerotic plaque, with possible concomitant vasospasm. Rare causes of acute coronary syndromes are embolic, subsequent arterial dissection, vasculitis, cocaine abuse and posttraumatic lesions (Theroux 2011).

Unstable angina is the clinical condition in which a patient manifests symptoms of myocardial ischemia (severe chest pain, sweating, anxiety) without an increase in serum levels of myocardial damage enzymes. When the myocardial ischemia leads to an infarction involving part of the heart wall, such enzymes are released in the circulation and the death of heart cells lead to electrical changes that can be observed on the electrocardiogram (ECG). When the lesion is more severe and crosses the entire heart wall, all the above mentioned conditions are met and the specific ECG changes called “ST-elevation” gives the name of the condition—ST Elevation Myocardial Infarction or STEMI. The other type of myocardial infarction (affecting only part of the heart wall) is thus called a non-STEMI (non-STEMI), as in this case all other conditions are met except the ST-elevation on the ECG tracing.

The thrombus formed on a ruptured or eroded plaque is usually occlusive and sustained in STEMI, whereas in non-STEMI, the thrombus is frequently non-occlusive and dynamic.

Coronary thrombosis is associated with plaque rupture, in which a structural defect of the fibrous cap is exposing the thrombogenic lipid-rich core to the blood flow; when no rupture is identified, the pathological insight is based on plaque erosion (Schaar et al. 2004).

The time between plaque rupture and syndrome onset is not known, because the first event is totally asymptomatic and the thrombotic process is highly unpredictable;

some studies suggest that severe thrombosis occurs immediately after plaque rupture (Falk 1983), others describe a dynamic response to a thrombotic event characterized by subsequent thrombolysis and vasospasm, with the tendency of creating a layered thrombus over days (Falk et al. 2004).

2.3.1.5 Plaque Vulnerability

On rupture-prone plaques, the rupture tends to occur at the cap margin, its thinnest and weakest area (Falk et al. 1995). Rupture of a thin cap and subsequent thrombosis may occur spontaneously, but in some cases, a temporary increase in emotional or physical stress provides the final triggering of the event. Recognized triggers include physical and sexual activity, heavy meals, psychological status (anger, anxiety, stress), earthquakes, war and terror attacks, temperature changes (cold exposures), acute infections, cocaine use (Mittleman and Mostofsky 2011) or just simple activities (Servoss et al. 2002). A specific circadian rhythm in the frequency of onset of acute coronary syndromes was detected, with a peak of events in the first ours of the morning. This statement was not characteristic in patients receiving beta-adrenergic blocking agents before myocardial infarction onset, but was present in those not receiving such therapy (Muller et al. 1985).

Very thin fibrous caps are at risk of rupturing and are influenced by the composition and gradual loss of smooth muscle cells, ruptured caps containing lower levels of smooth muscle cells and less collagen than intact caps (Kolodgie et al. 2001).

The lipid-rich core is essential for defining plaque rupture and its complications. A larger lipid-rich core promotes a higher risk, the influence on size being somewhat intuitive, as a large lipid-rich core may erode the fibrous cap and promote thrombosis after plaque rupture (Theroux 2011).

It has been considered that plaque size does not play an important role. Emerging evidence suggests that microcalcifications play a critical role in determining atherosclerotic plaque vulnerability (Hutcheson et al. 2014).

Multiple studies suggest that plaques causing significant stenosis are at a higher risk for clinical events. Conversely, it has been demonstrated that most acute coronary syndromes are induced by plaques that were not causing significant stenosis on angiography evaluations performed months before, a finding which may be that these lesions are much more prevalent than stenotic ones (Stone et al. 2011).

Lesions responsible for acute coronary syndromes are less calcified than plaques responsible for stable angina (Burke et al. 1997). In ruptured plaques there is more neovascularization and inflammation of the adventitia than in intact plaques (Virmani et al. 2005).

Erosion-prone plaques promote coronary thrombosis without rupture, but the mechanisms of erosion are not well known, and it is currently considered to involve multiple pathological ways. They generally involve endothelial denudation over pathological intimal thickening and fibroatheromas with thick cap (Virmani et al. 2000). A potential trigger of endothelial damage is the presence of vasospasm (Virmani et al. 2006).

2.3.1.6 Clinical Significance of Atherosclerotic Disease

Significant narrowing of the coronary arteries produces angina pectoris, its symptomatology consisting of acute chest pain during exercise, cold temperatures or emotional stress, lasting for 2–10 min. It may be described as retrosternal chest pressure, burning, squizzing, heaviness, radiating occasionally to the neck, jaw, epigastrium, shoulder or left arm (Mann et al. 2015).

It is usually produced by fixed stenosis on one or more coronary arteries. The pattern of symptoms is correlated to the degree of stenosis. During exertion, the sympathetic system increases heart rate, blood pressure and contractility, all these mechanisms eventually increasing the demand of myocardial oxygen; when exceeding the supply of oxygen, myocardial ischemia results, causing chest discomfort called angina pectoris. The symptomatology persists until oxygen balance is restored (Barry et al. 2014).

Endothelial dysfunction prevents appropriate vasoconstriction thus aggravating the already insufficient oxygen supply and the capacity to release local metabolites which induce vasodilatation in normal vessels. Therefore, the altered coronary vascular tone, which varies in different patients, contributes to the pre-existing physical arterial narrowing.

Symptoms are extremely variable, and “angina equivalents” have been described in different patient populations (women, elderly people, diabetics) who experience angina, in the absence of chest pain, as jaw or shoulder pain, dyspnea, nausea or vomiting and diaphoresis (Mann et al. 2015).

Unstable angina presents under three main forms: (a) rest angina or angina with minimal exertion usually defined as occurring within the last month; (b) new-onset angina, usually defined as occurring within the last month; (c) crescendo angina, defined as previously diagnosed angina that has become more frequent, longer in duration or more severe in nature (Betriu et al. 1992).

Usually, the pain in unstable angina frequently waxes and wanes, lasting for a few minutes to as long as, but usually less than, 20 min (Theroux 2011).

Non ST-segment elevation and ST-segment elevation myocardial infarction is diagnosed based on typical ischemic chest pain, typical ECG changes, including development of Q waves and elevated serum markers of myocardial injury, usually creatine-kinase myocardial band. In the recent years, the availability in clinical practice of serum troponin levels, being much more sensitive and specific to myonecrosis than creatine-kinase myocardial levels, has been associated with higher rates of rapid diagnostic and assessment (Fuster et al. 2011).

STEMI refers to the most lethal form of acute coronary syndromes, and develops when a completely occlusive thrombus interrupts coronary blood flow in the territory of the occluded artery. Specific changes on the ECG, such as ST segment elevation and development of Q waves, highly suggest full or nearly full thickness necrosis of the myocardium wall (Fuster et al. 2011). The fast and accurate diagnosis dictates the necessity for immediate therapeutic measures, such as reperfusion therapy, mechanical revascularization or thrombolysis. Mortality was proven to significantly decrease in those who received reperfusion therapy within 12 h of symptom onset (Califf 1998).

Patients with acute myocardial infarction (both STEMI and NSTEMI) present ischemic pain that develops abruptly, is steady and lasts for more than 30 min. Patients frequently describe the pain as a pressure, burning, grawing, tightness, heaviness or crushing, when severe.

Typically patients describe center or left chest pain, radiating to classical areas (left shoulder, left elbow, jaw). Less commonly, they present with epigastric pain causing confusion and indigestion, which may lead to misdiagnosis. On the contrary, patients with acute coronary syndromes may have other accompanying symptoms, most commonly dyspnea, diaphoresis, nausea, vomiting and palpitations. Gastrointestinal symptoms are usually associated with inferior wall infarction. Other atypical presentations include syncope, acute heart failure with pulmonary edema, generalized weakness and acute mental status changes (Theroux 2011).

Clinical manifestations of atherosclerosis become evident in the middle-aged and elderly individuals, but atherosclerosis is a lifelong disease. Fatty streaks are present in most individuals arteries even after puberty, intermediate lesions occur in most people from 20 to 30 years of age, and fibroatheromas are frequent in those aged 30 years and older.

When coronary atherosclerosis begins to cause symptoms, it has already become a systemic disease affecting most of the epicardial portion of the coronary arteries.

The invasive approach (e.g. percutaneous coronary intervention) of acute coronary syndromes ensures, in many cases, rapid, complete, and sustained recanalization of the culprit lesion. The possibility to prevent future events by addressing, during the same intervention, other coexisting asymptomatic lesions which display characteristics of vulnerable lesions is currently being explored.

The optimal strategy to prevent and slow the progression of atherosclerosis is based on prevention and treatment of lifelong risk factors, which involves both population-wide risk reduction strategies and individual lifestyle changes. Currently available medical therapy is highly efficient in affected individuals and many other novel therapies are still in development for those who still remain at high risk of future cardiovascular events.

2.3.2 Diagnostic Imaging in Coronary Artery Disease: Echocardiography Studies

In the past 20 years, echocardiography has become an established and powerful tool for diagnosing the presence of CAD and defining its consequences in patients with acute ischemic syndromes or in those with chronic coronary atherosclerosis. There is little to no risk in performing a cardiac echo in a patient with suspected CAD whereas other modalities that can be used to assess ischemia have restrictions and higher risk factors. In addition, its bedside availability, made this method the most used imagistic noninvasive method for diagnostic and risk stratification in CAD. With little exceptions, echocardiography represents the third step in the algorithm of management of a patient with CAD after clinical evaluation and resting ECG tracings.

It can provide vital information of the pathophysiology of CAD by using 2D, M-Mode, and Doppler applications, this techniques being generally sufficient for evaluating patients with suspected or documented cardiac ischemia. However, transesophageal echocardiography (TEE) may be needed in some patients, particularly those with serious hemodynamic compromise but non-diagnostic transthoracic echocardiography (TTE) studies. In these circumstances TEE can distinguish among extensive infarction with pump failure, mechanical complications of infarction, or hypovolemia and can guide prompt therapy. Stress echocardiography is useful for evaluating the presence, location, and severity of inducible myocardial ischemia as well as risk stratification and prognostic information. Newer echocardiographic techniques like speckle tracking echocardiography and three dimensional echocardiography aren't used yet in clinical practice, but only for research purposes.

2.3.2.1 Basic Principles of Echocardiography in Coronary Artery Disease

Echocardiography is not able to visualize directly the coronary blood flow through the coronary arteries, but can depict indirectly consequences of the imbalance between flow and consumption. The mainstay of this phenomenon are the changes in left ventricular wall motion characteristics. The changes are usually regional and may be detected by echocardiography considering that ultrasound beam traverses the affected area and image quality is satisfactorily. Unfortunately many ischemic patients are heavy smokers, obese, with hyperinflated lungs, making echocardiographic exam difficult.

History begun in 1935, when Tennant and Wiggers demonstrated that coronary occlusion resulted almost immediately in abnormalities of wall motion (Tennant and Wiggers 1935). Forty years later, experimental studies on canine models using ultrasonic crystals, demonstrated that during acute ischemia (Theroux et al. 1974) and infarction (Kerber and Abboud 1973), reduction in regional flow are followed closely by a reduction in contractile function, and set the use of ultrasound in CAD.

Myocardial ischemia results in a cascade of events in which various changes are hierarchically ranked in a well-defined time sequence (Picano 1989). Ischemic cascade starts in aerobic tissues after seconds to minutes of inadequate blood supply and begins with a series of biochemical reactions followed by changes in left ventricular function and only afterwards ECG changes and clinical manifestations (Fig. 2.2). Hence, echocardiography can detect ischemic heart disease before electrical and clinical signs occur, although in clinical practice assessing a patient with suspected cardiac ischemia begins rarely with ultrasound examination.

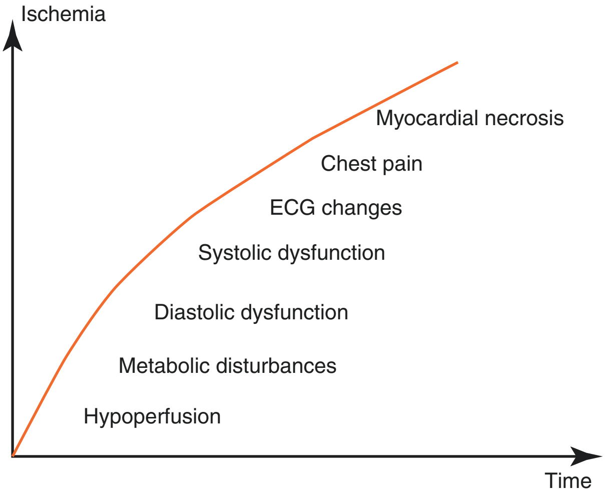


Fig. 2.2 Schematic depiction of ischemic cascade, showing that echocardiography can detect ischemia before ECG changes or clinical manifestation occur (Baskot 2011)

2.3.2.2 The Importance of Echocardiography in Patients with Coronary Artery Disease

In the light of this, echocardiography is very useful for assessing patients with chest pain in order to make a differential diagnosis, evaluating global and regional left ventricular function, determining the coronary artery that could be probably involved in the ischemic process, stratifying prognosis and guiding therapy.

Differential Diagnosis of Chest Pain

Echocardiography is very important in confirming the ischemic origin of chest pain by excluding other cardiac and noncardiac causes of chest pain (Wang and Fleischmann 2001), helping differentiate angina pain from aortic dissection, acute pericarditis, pulmonary embolism, pulmonary arterial hypertension, spontaneous pneumothorax, also life threatening conditions such as unstable angina or acute myocardial infarction. Echocardiographic findings like dilated aorta with intimal flap, pericardial effusion, and dilated right heart with elevated pulmonary pressure make acute coronary syndromes less probable. On the other hand, detecting left ventricular wall regional motion abnormalities and sometimes right ventricular wall kinetic changes confirms the suspicion of acute coronary syndromes and, together with ECG changes, allows for a more rapid diagnosis allowing the rapid establishment of invasive and noninvasive therapeutic measures. In addition, echocardiography has a high positive predictive value in patients with left bundle branch block (Cortigiano et al. 2001; Bouzas-Mosquera et al. 2009) or pacing in which ECG findings are not diagnostic for an ischemic process.

Diastolic Dysfunction in Ischemic Heart Disease

In the ischemic cascade, after the occurrence of metabolic changes, the next step represents the appearance of diastolic dysfunction. Diastole is the cardiac cycle phase during which the heart is relaxing and filling with incoming blood that is being returned from the body through the inferior and superior venae cavae to the right atrium and from lungs through pulmonary veins to the left atrium. Diastolic failure occurs when the ventricle cannot be filled properly because it cannot relax or because its wall is thick or rigid. Diastolic failure is characterized by an elevated diastolic pressure in the left ventricle, despite an essentially normal/physiologic end diastolic volume (Lenihan et al. 1995).

To evaluate left ventricular diastolic function (Sherif et al. 2009), a PW Doppler sample volume is placed at the mitral valve leaflet tips and the following measurements are recorded: E wave representing early diastolic velocity, meaning the filling of the left ventricle due to left atrium-left ventricle gradient in early diastole; A wave representing late diastolic filling due to atrial systole; E/A ratio (normal values 1.1–1.5); E wave deceleration time (normal values 160–240 ms), isovolumic relaxation time (normal values 76 ± 13 ms) (Fig. 2.3). For further evaluation of left ventricular diastolic function, one should evaluate pulmonary vein flow by placing PW sample into the left upper pulmonary vein ostia, recording S and D wave velocities and the pulmonary vein "a" flow reversal (normal values <25 cm/s) (Fig. 2.4), together with left atrium size, preferably by measuring LA volume (normal values <20 mL/mp) (Fig. 2.5).

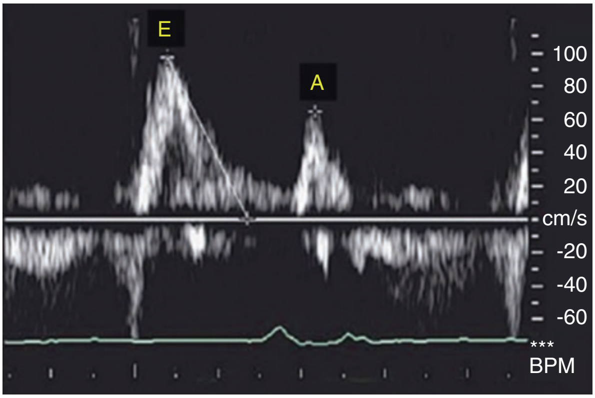


Fig. 2.3 Application of PW Doppler to the mitral valve for studying LV diastolic function

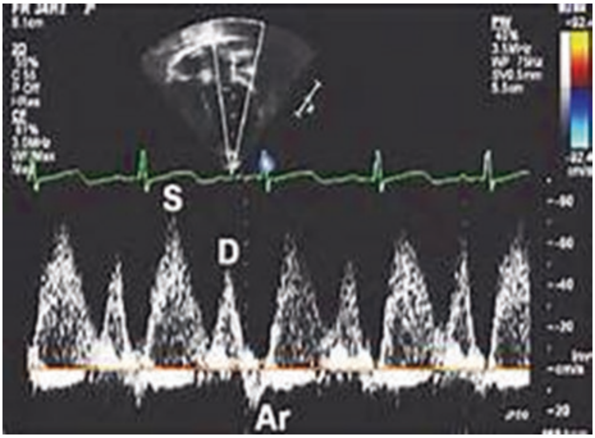


Fig. 2.4 Placing of a PW sample in the upper left pulmonary vein

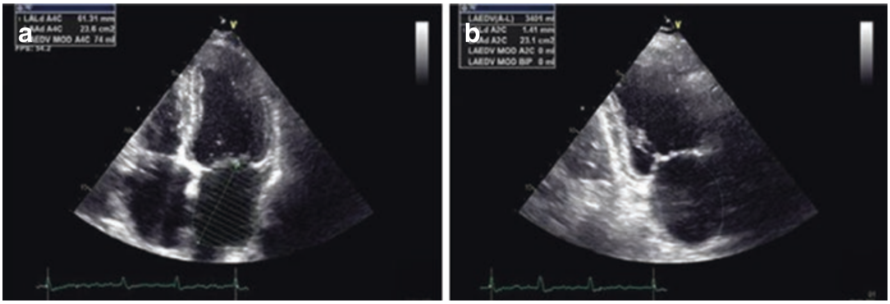


Fig. 2.5 Measuring left atrium size using biplane method, by measuring left atrium volume in apical 4 and 2 chamber views

Table 2.1 Stages of abnormal diastolic filling

Stage of diastolic dysfunction	Impaired early left ventricular relaxation	Pseudonormalization	Restricted filling pattern
Significance	Impaired (slow) early left ventricular relaxation	Suggests impaired (slow) early left ventricular with decreased left ventricular compliance	Severe decrease in left ventricular compliance and impaired (slow) early left ventricular relaxation
Signs and symptoms	None at rest	Exertional dyspnea	Dyspnea with minimal exertion
Functional status	Mild impairment	Moderate impairment	Marked impairment
Left atrium	Normal dimension (may be hypercontractile)	Enlarged and hypocontractile	Enlarged and hypocontractile
Filling pressures	Normal	Increased	Markedly increased
PW Doppler findings	E/A ratio < 1.0	E/A ratio: 1.0–1.5	E/A ratio: >1.5
	Deceleration time > 240 ms	Deceleration time: 160–240 ms	Deceleration time: <160 ms
	IVRT >90 ms	IVRT: 76 ± 13 (>40 years), 69 ± 12 (<40 years)	IVRT: <60 ms
	Pulmonary Vein “a” wave flow reversal <25 cm/s	Pulmonary vein “a” wave flow reversal: >25 cm/s	Pulmonary’ vein “a” wave flow reversal: >25 cm/s (Variable: Depending on atrial systolic function)

In acute or chronic ischemia, three patterns or stages can indicate abnormal diastolic filling (Table 2.1).

In CAD, as well as in other pathologies, it is important to distinguish between the normal and the pseudonormal diastolic pattern, as therapeutic interventions can be appropriate in reversing patients' condition. In this setting, methods of making this differentiation consist of applying the Valsalva maneuver to the Doppler transmitral inflow pattern together with using the pulmonary venous flow (a difference of >30 ms between pulmonary atrial reverse velocity and transmitral A velocity indicates increased left ventricle filling pressure).

During acute, transient ischemia, different stages of diastolic dysfunction can occur, from type I to type III. In case of a pre-existing diastolic dysfunction one can note its exacerbation instead. Once the ischemic process stops, diastolic dysfunction can reverse to normal or improve. However, it is characteristic for the ischemic cascade that diastolic dysfunction precedes systolic dysfunction.

After an acute myocardial infarction, the Doppler assessment of left ventricle diastolic function plays an important prognostic role regardless of other prognostic markers (e.g. NYHA class). A short transmitral Doppler deceleration in E velocity (<130 ms) is an early, independent marker of late left ventricle remodeling and confers an adverse prognosis; the predischarge persistency of a restrictive pattern is also prognostic (Temporelli et al. 2004). Left atrium dilation was also proved to hold prognostic value after acute myocardial infarction. The GISSI study found that that left atrium dilatation contributes to left ventricle remodeling (Popescu et al. 2004).

Global and Regional Systolic Left Ventricular Functional Changes in Coronary Artery Disease

Echocardiography can be used to assess global systolic left ventricle function and the regional wall motion abnormalities caused by transient ischemia or myocardial necrosis and identify their extent and site with great precision within just a few minutes of coronary occlusion and before onset of symptoms.

During an acute ischemic process in case of unstable angina or during the acute phase of a myocardial infarction evaluating global left ventricle function is mandatory, by assessing left ventricle ejection fraction which represents the ratio of left ventricle stroke volume (the difference between left ventricle end-diastolic volume and end-systolic volume) and left ventricle end-diastolic volume. End-diastole is defined as the first frame after mitral valve closure or the frame in the cardiac cycle in which the left ventricle linear dimensions or volume are the largest. End-systole is best defined as the frame after aortic valve closure or the frame in which L left ventricle V dimensions or volume are the smallest (Lang et al. 2015). In patients with a regular heart rhythm, measurements of the timing of openings and closures derived from M-mode echocardiography, pulsed wave (PW) or continuous wave Doppler (CW) may be used for accurate definitions of ventricular time intervals. Left ventricle volume estimates may be derived from two dimensional echocardiography (2D) or three dimensional echocardiography (3D). Volumetric measurements are based on tracing of the interface between the compacted myocardium and the left ventricle cavity. Left ventricle ejection fraction should be measured by means of 2D echocardiography using modified Simpsons method in apical four and two chamber view (Lang et al. 2015) (Fig. 2.6). 2D echocardiographic acquisition should aim to maximize left ventricle areas to avoid foreshortening and therefore

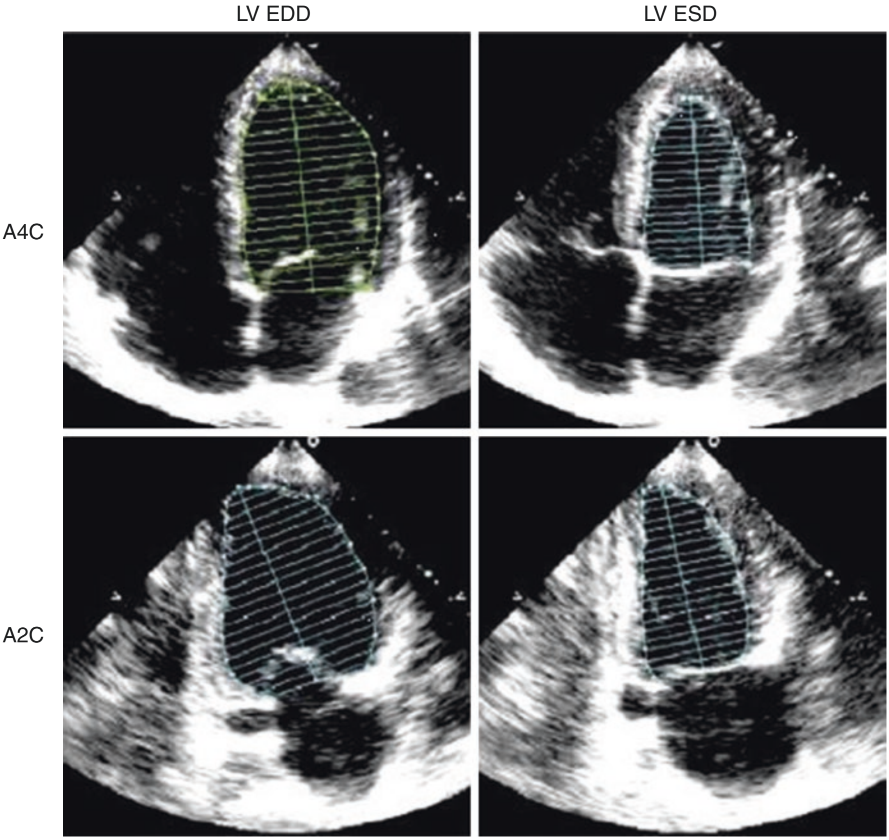


Fig. 2.6 Measuring left ventricle volumes through 2D echocardiography by modified Simpson method

underestimation of left ventricle volumes. This shortfall was avoided by introducing real-time 3D echocardiography which has led to very interesting perspectives for measuring left ventricle volumes and ejection fraction in ischemic heart disease, even when MRI is used as the gold standard (Jenkins et al. 2007). Thus, 3D echocardiography measurements of ejection fraction are accurate and reproducible and should be used when available and feasible (Fig. 2.7). Left ventricle ejection fraction <52% for men and <54% for women are suggestive of abnormal left ventricle global systolic function. The Teicholtz method for calculating left ventricle volumes from left ventricle linear dimensions is no longer recommended (Lang et al. 2015).

A number of disorders mimicking acute coronary syndromes (myocarditis/pericarditis, valvular heart diseases) or acute coronary syndromes caused by extensive ischemia (multicoronary vessel disease) can be associated with a globally reduced ejection fraction.

By means of 2D echocardiography regional myocardial function can be evaluated using the segmentation schemes that reflect the coronary perfusion territories.

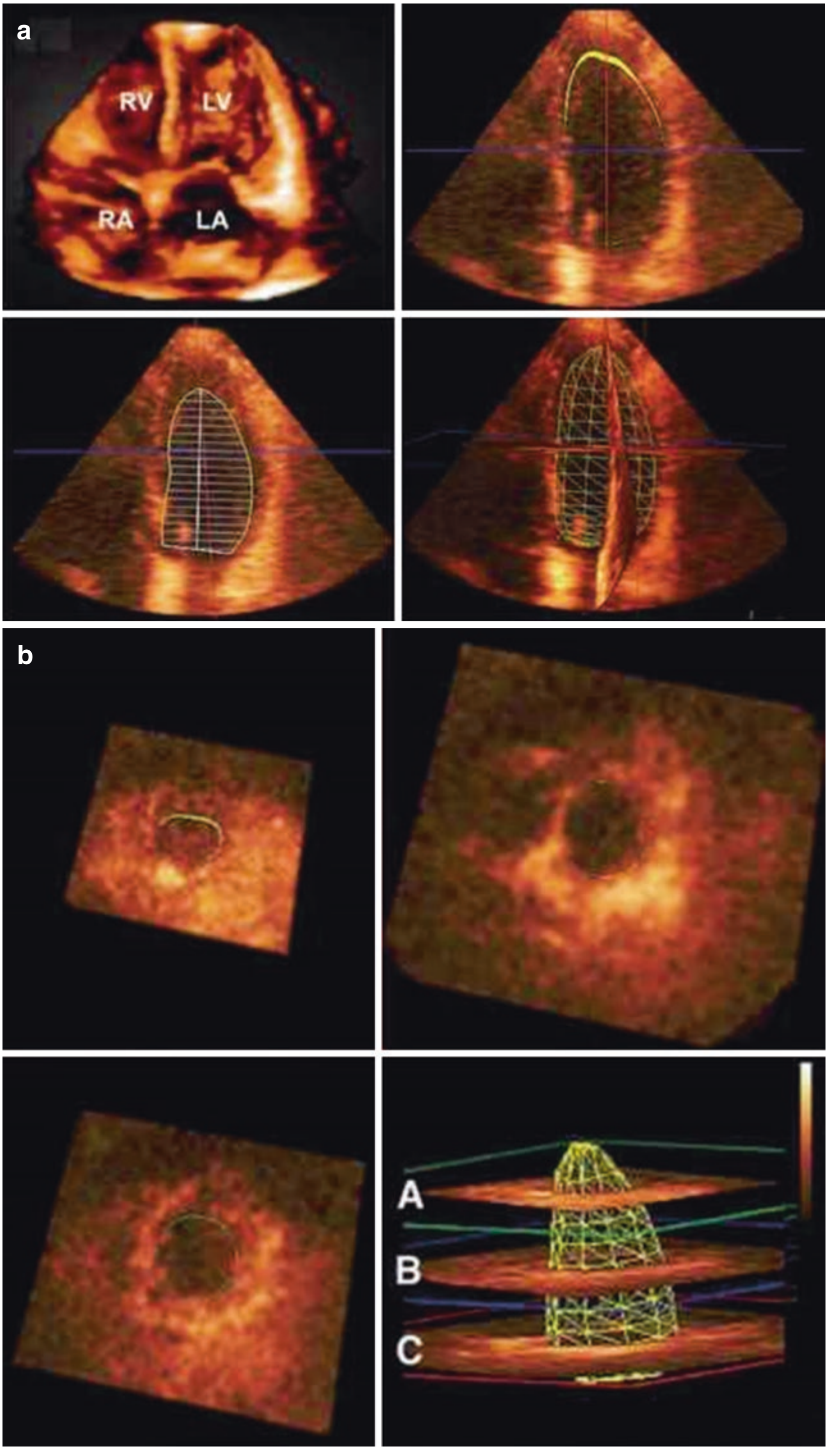
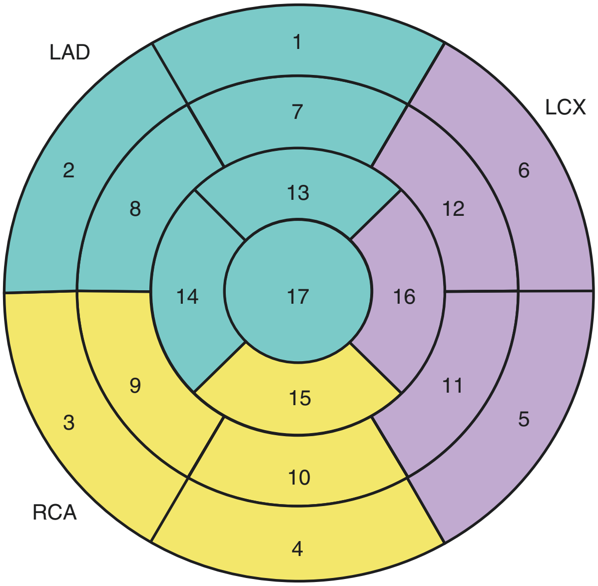


Fig. 2.7 Measuring LV volumes through real time 3D echocardiography

Fig. 2.8 Coronary artery distribution



In an attempt to establish a standard segmentation applicable to all types of cardiac imaging, a 17 segment model (including the apical cap) has been introduced and endorsed by the recent guidelines (Cerqueira et al. 2002). Despite the great variability in coronary artery supply, the assignment of individual segments to specific coronary arterial territories has been standardized (Fig. 2.8) (ASNCI 1999).

In echocardiography, regional myocardial function is assessed on the basis of the observed myocardial wall thickening and endocardial motion of the myocardial segment. It is recommended that each of the 17 segments be analyzed individually in multiple views. At least three wall motion abnormalities have been defined: hypokinezia—meaning reduced thickening, akinesia—meaning absent or negligible thickening, and dyskinesia—which refers to systolic thinning or stretching, such as in the case of an aneurysm. An aneurysm is characterized morphologically by focal dilatation and thinning with either akinezia or systolic deformation (Voigt et al. 2003b) (Fig. 2.9).

In clinical practice and for research purposes, a semiquantitative wall motion score index is used and is calculated as average of the scores of all segments visualized. Every normal or hyperkinetic segment receives one point, every hypokinetic segment receives two points, whereas akinetic myocardial regions receive three points and dyskinetic ones four points. The calculated sum is then divided by the number of segments. The wall motion score index has prognostic value after an acute myocardial infarction (Moller et al. 2006).

A regional wall motion abnormality at rest cannot be visualized until the stenosis exceeds 80–85% of the coronary lumen; with exercise or pharmacological stress, a 50% coronary artery lesion can lead to regional dysfunction. Studies show that in case of an acute ST elevation myocardial infarction, up to 90–100% of patients

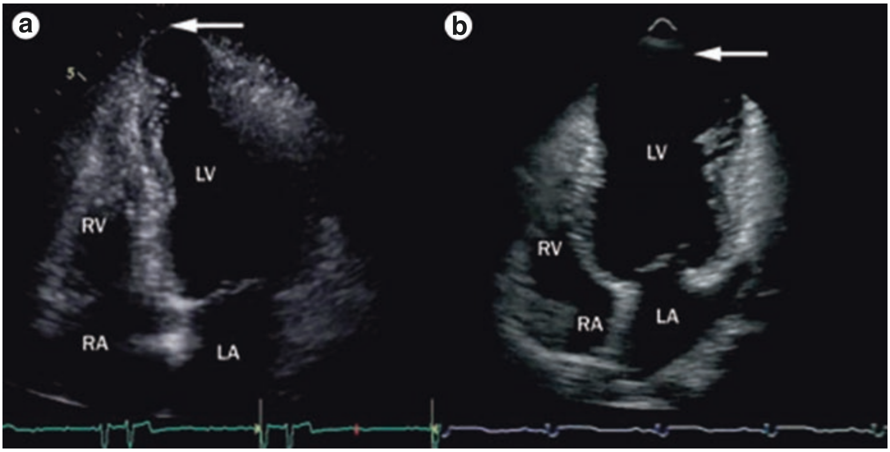


Fig. 2.9 Left ventricular aneurysm

show wall motion abnormalities, whereas in non ST elevation myocardial infarction, about 86% show wall motion disorders (Pierard et al. 1986; Ogawa et al. 1985). In general, wall thickening discriminates normal and infarcted myocardium better than wall motion, and transmural necrosis is better detected than subendocardial infarcts involving <20% of the whole myocardial thickness. In case of an acute myocardial infarction, the location of myocardial wall motion abnormalities is in accordance with the localization of the infarct deduced from ECG changes in 93–100% of patients, but discrepancies can appear in the context of previous myocardial scar/infarction, multivessel coronary artery disease or an overlap between the perfusion of the right and circumflex coronary arteries.

Besides 2D echocardiography, other echocardiographic methods derived from special techniques can be used for quantifying regional left ventricle dysfunction in ischemic heart disease. Doppler Tissue Imaging (DTI) and speckle tracking echocardiography are providing comparable data quality (Heimdal et al. 1998) (Leitman et al. 2004), although DTI is angle dependent and can underestimate motion that is not properly aligned to the ultrasound beam. Commonly used parameters are velocity, motion, deformation and deformation rate. Because velocity and motion are measured relative to the transducer, measurements may be influenced by overall heart motion. Therefore, the use of deformation parameters such as strain and strain rate are preferable. DTI assessment of mitral annulus can detect subclinical alterations in longitudinal systolic function (Sm) or in early myocardial relaxation (Em) (Fig. 2.10), even when conventional echocardiographic methods do not reveal changes. Such findings could be signs of mild subendocardial damage due to myocardial ischemia.

The term “deformation” refers to changes in shape and dimension of the myocardium during the cardiac cycle. If there is myocardial ischemia, or if a myocardial infarction had occurred in the past, this part of the heart muscle is weakened and shows reduced and altered systolic function. The most commonly used parameter to asses deformation is the ***longitudinal strain*** during LV systole. All four components

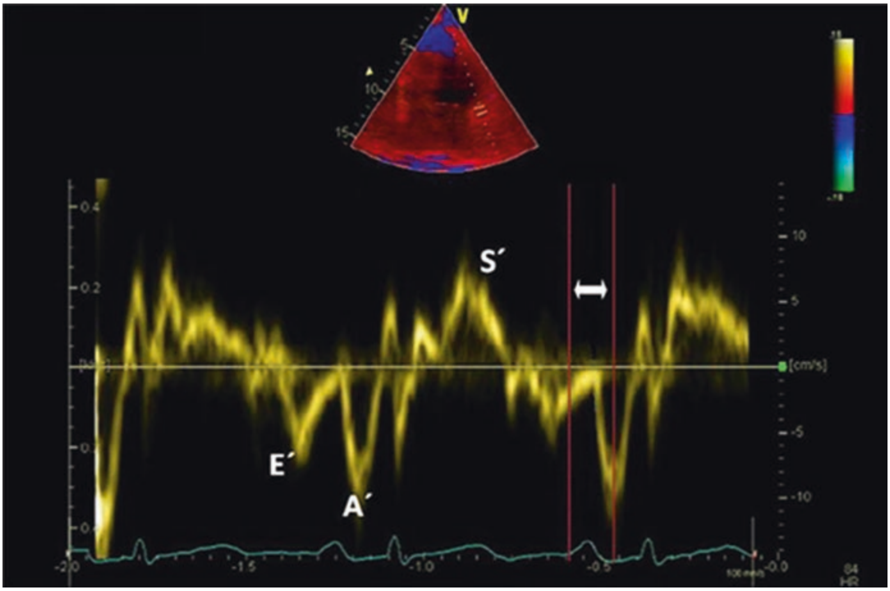


Fig. 2.10 DTI of the mitral annuls showing systolic and diastolic velocities

of strain can be expressed using three dimensional echocardiography, e.g. three dimensional strain, currently used only for research purposes, without having yet defined normal ranges. With currently available software, deformation parameters can vary in amplitude depending on the region, on the vendor or the methodology used. The temporal relationship of myocardial deformation is also important, while independent of the values obtained for strain. For example, longitudinal shortening or radial thickening of a myocardial region, occurring after aortic valve closure (sometimes called tardokinesia) is a sign of regional dysfunction, suggesting ongoing ischemia or the presence of an irreversible process (scar formation) (Voigt et al. 2009). The development of postsystolic shortening during a stress test has been proposed as an indicator of regional ischemia (Voigt et al. 2003a). Strain rate imaging is an imagistic method for measuring regional or global deformation of the myocardium. Strain rate imaging serves to assess longitudinal myocardial function in patients with acute myocardial infarction, where strain rate seems to predict regional wall motion recovery after percutaneous coronary intervention.

In a recent study, 3D real time color-flow Doppler transthoracic echocardiography was employed for the automated quantification of stroke volumes and mitral inlet flows (Thavendiranathan et al. 2012). Results were compared against measurements extracted from cardiac magnetic resonance imaging, and indicated the proposed method was accurate and reproducible. Due to its real-time capabilities it is well suited for clinical workflows.

Furthermore, since left ventricular stroke volume plays a significant role for the assessment of the circulation, an in vivo animal study was conducted for determining the accuracy of 3D real-time color-flow Doppler transthoracic echocardiography for the evaluation of volumetric flow rates through the aortic and mitral valves (Shimada et al. 2015). The study has shown that real-time 3D color-flow Doppler is an accurate method for the evaluation of both mitral and aortic valve volumetric flow rates.

Another echocardiographic method, frequently used in ischemic heart disease, is stress echocardiography, the combination of 2D echocardiography or 3D echocardiography with the use of an ischemic stressor: physical, pharmacological or electrical (Rosa et al. 2008). The diagnostic endpoint of myocardial ischemia is the induction of a transient worsening in regional function during stress in a segment that is contracting normally at rest. In parallel, the stress echo sign of myocardial viability is a stress induced improvement of function during low levels of stress in a region that is contracting abnormally at rest.

Regardless of the kind of stress used, myocardial ischemia propagates centrifugally, involving first the subendocardial region and afterwards the subepicardial layer (Gallagher et al. 1984). In patients with normal epicardial coronary arteries, coronary flow reserve can be reduced in the presence of microvascular disease (microvascular angina or left ventricular hypertrophy). In this setting, angina can occur typically along with signs of subendocardial lesion on the ECG, but without wall motion abnormalities on stress testing. Therefore, wall motion abnormalities are more specific for the detection of epicardial coronary artery disease (Nihoyannopoulos et al. 1991).

The choice of stressors depends on the experience of the individual operator and laboratory with the agent and/or protocol. Exercise is the prototype of ischemic stress and the most widely used, combining echocardiography with pedaling on a supine bicycle. In patients who cannot exercise maximally one could use pharmacological stressors like dobutamine or dypiridamol. The use of pharmacological stressors is more convenient as it lacks the drawbacks imposed by factors hampering the image quality during exercise stress testing, such as hyperventilation and excessive chest wall motion movement frequent in physical stress testing. Dobutamine and dipyridamole act on different receptors: dobutamine stimulates adrenoreceptors, while dipyridamole stimulates adenosine receptors, therefore inducing ischemia through different mechanisms. Dobutamine increases myocardial oxygen demand, while dipyridamole decreases subendocardial flow. Dipyridamole provokes “vertical steal“, which suggests the presence of an epicardial artery coronary stenosis while the subepicardium steals blood from the subendocardium (Rosa et al. 2008).

In terms of diagnostic criteria, in stress echocardiography, the fundamental stress response patterns are normal, ischemic, viabile and necrotic myocardium. In the normal model, a segment is normokinetic at rest and becomes hyperkinetic at stress. The ischemic segment’s function deteriorates during stress (a stress test is usually positive when these findings are present in two adjacent segments). The necrotic segment has resting dysfunction which remains fixed during stress. The viable segment has resting dysfunction and during stress can show either a sustained improvement indicating a stunned myocardium, or first improve during early stress and then subsequently deteriorate at peak stress (biphasic response)

(Senior and Lahiri 1995), indicating a hibernating myocardium. The hibernating myocardium requires therapeutic revascularization to improve contractility. A resting akinetic segment becoming dyskinetic during stress usually indicates a passive, mechanical consequence of increased intraventricular pressure and is not a true sign of active ischemia.

Given the fact that the interpretation of contractile function is subjective, an improved image quality could reduce inter-operator variability. For this purpose, harmonic imaging and ultrasound contrast agents for LV opacification are recommended to enhance endocardial border detection and hence accuracy of the examination (Kirkpatrick et al. 2005).

More recently, it 3D echocardiography was validated for the assessment of left ventricular volumes. Specifically, automated methods for real-time border detection have been proposed, thus contributing to the integration of 3D echocardiography in routine clinical workflows (Barbosa et al. 2013).

In conclusion, echocardiography is the most commonly used imagistic method to assess myocardium functionality in coronary heart disease. It is crucial for diagnostic and risk stratification, and has become so because of its bedside availability, relatively low cost, repetabilty and lack of adverse effects. Although the method does not image the coronary arteries directly, it allows accurate visualization of wall motion abnormalities, allowing precise localization of culprit coronary artery in most cases of chronic ischemia or acute coronary syndromes. Newer echocardiographic techniques, such as deformation imaging or three dimensional echocardiography, can detect subtle changes in myocardial mechanics or function, before classical methods show any alterations. Stress echocardiography is a very valuable method for increasing sensibility and specificity in diagnostic and risk stratification in ischemic heart disease.

2.3.3 Invasive Diagnosis of Coronary Artery Disease: Angiography

The invasive coronary evaluation—coronary angiography and angioplasty—is nowadays the cornerstone in the complex treatment management of CHD. From the right heart catheterization of Forssmann in 1929 to the accidental injection of contrast dye in the coronary arteries of Sones in 1960 and to Gruentzig in 1977, coronary angiography and angioplasty have come a long way to become a leading worldwide method of diagnosis and treatment. With the invention of coronary stents and extensive research for minimizing the technology for better and easier access, the method has become readily available all around the world. The past 25 years have witnessed a great leap of technology in the field of interventional cardiology technology and methodology, allowing physicians to treat CHD less invasively and to treat acute myocardial infarctions before complete damage of the myocardium occurs. Moreover, the past decades have brought novel research and breakthroughs to understand the pathology of atherosclerosis and describe the intimate processes of intravascular coronary physiology.

2.3.3.1 The Coronary Angiography

In order to perform a coronary angiogram or percutaneous coronary intervention (PCI), first we must have an arterial access, usually through the femoral or radial artery. As the most common side effects of coronary catheterization are access-site related, recent years have brought a shift made towards the radial access, as this method has almost no access-site complications. The artery access is performed via an arterial introducer, usually 1.7 or 2 mm (5 or 6 French gauge) in diameter. For very complex PCIs a 2.3 mm introducer can be used (7 Fr gauge). Through the introducer, specific pre-shaped catheters are introduced under X-ray guidance all the way to the ascending aorta and the tip of the catheters are gently introduced in the coronary ostia. Either manually, or via an automated syringe, contrast dye is injected and the entire coronary artery circulation can be visualized (Serruys et al. 2010) (Figs. 2.11 and 2.12). Even to this day, despite the availability of many incredibly accurate methods of assessment of both intravascular physiology and anatomy, the decision to revascularize is still performed via a visual assessment of the coronary artery stenosis (Hannawi et al. 2014).

Owing to technological progress and miniaturization, the coronary angiography is considered a safe technique, with low rates of complications that vary greatly accordingly to patient comorbidities. Most complications are access-site related, the femoral artery having been the main access artery for more than 20 years. These problems include most often local haematoma, the formation of a pseudoaneurysm and an arterio-venous fistula and occur in 0.3–5% of patients (Alonso et al. 2003; Baim and Grossman 1996). These complications are minor, often self-limiting and easily treated via manual, interventional or local minimal surgery (Krueger et al. 2005; Paulson et al. 2000). These access related complications are almost absent in

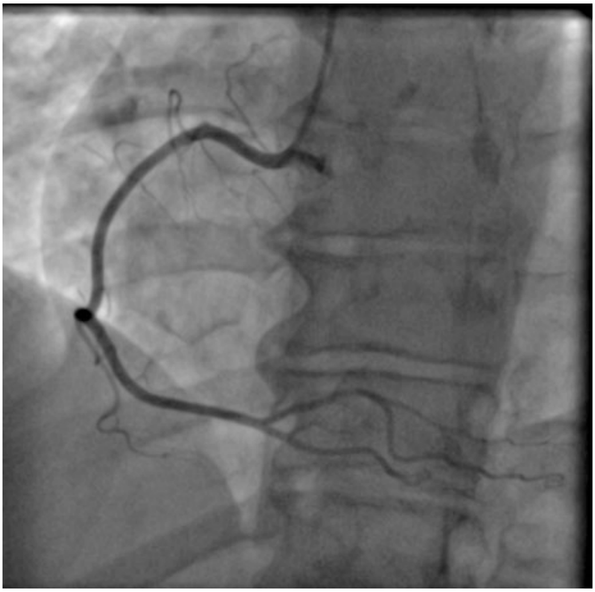


Fig. 2.11 Right coronary artery angiography—normal appearance

Fig. 2.12 Left coronary artery angiography with left anterior descending artery and left circumflex artery—normal appearance

the case of a radial access, but with slightly more radiation exposure for the operator and a longer learning curve. Other, more severe complications, include arrhythmias—the most severe being ventricular tachycardia or ventricular fibrillation and cardiac arrest, which is usually reversible—and complete atrioventricular block or asystole, also usually reversible. In addition, arterial embolism or dissection may occur as a severe complication followed by acute stroke (0.2–0.4%) (Kwok et al. 2015), or acute limb or renal ischemia, or even aortic dissection. Another significant complication is contrast-related nephropathy, with more than 5% of patients having a transient increase of serum creatinine, but usually less than 1% actually requiring haemodialysis. Myocardial infarction or death after a diagnostic procedure is quite rare, both being reported in less than 0.1% (Wyman et al. 1988; Johnson et al. 1989).

2.3.3.2 Percutaneous Coronary Interventions

The PCI is inextricably linked to the diagnostic coronary angiogram (Fig. 2.13). It is the cornerstone in the treatment of CHD because nowadays it is the main method of restoring normal coronary blood flow. The PCI follows after the decision to revascularize usually taken within the same procedure (Fig. 2.14).

If one can define the coronarography as a diagnostic method, the PCI is the treatment method that completes the coronary catheterisation spectrum. It is performed using the same type of vascular access, femoral or radial. Specific pre-shaped guiding catheters are used, with larger inner diameters and metallic wire reinforcements to compensate for thinner walls. These catheters are much stiffer than the diagnostic ones in order to provide the much needed support for the PCI. All PCIs commence with the gentle advance of a thin metallic wire called a guidewire. It has a diameter of 0.014 in. and its insertion is a main stage during the PCI. There are numerous

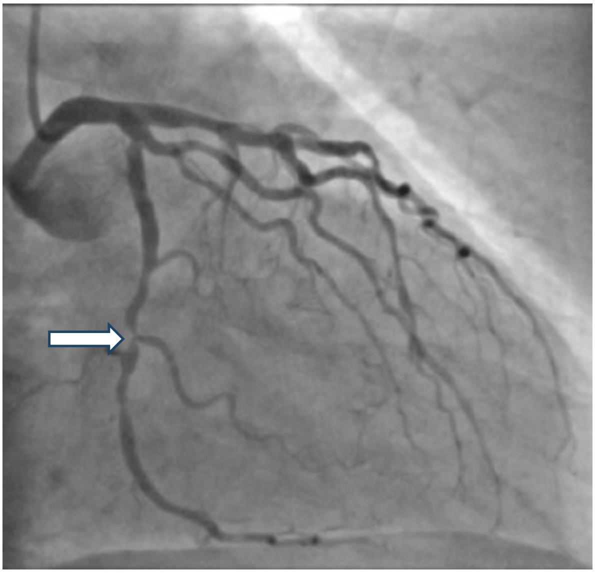


Fig. 2.13 Left circumflex artery severe stenosis—initial aspect

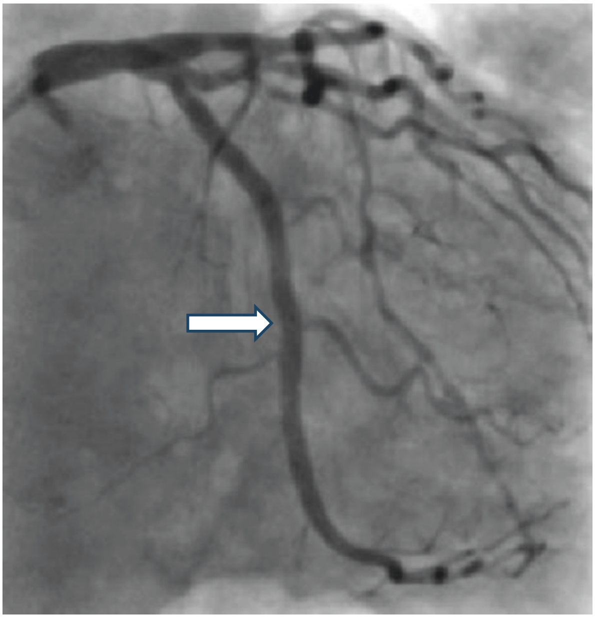


Fig. 2.14 Left circumflex artery—after PCI with stent implantation

types of guidewires, as they are classified by the guidewire coating (hydrophilic or non-hydrophilic), core material and structure, tip diameter and most importantly tip loading, steerability and trackability, torque and support. They act as a rail for subsequent material to be introduced in the coronary artery. After a wire is chosen and the target stenosis is crossed, usually a balloon catheter is then introduced on to the wire to the stenosis. The balloon catheter is connected to a manual pump and is inflated with pressures from 6–8 to at least 20 atm, with a mixed liquid formed from saline and contrast media. This stage is called predilation and it has the purpose of preparing the lesion for subsequent stent placement. The balloon is then deflated and extracted from the guiding catheter. The balloon is usually provided with radio-opaque markers in order to have a landmark for a correct stent length placement. Each of these stages are performed using intermittent injection of contrast media (Serruys et al. 2010).

Once the lesion is prepared and the length and diameter of the vessel stenosis is established, usually a stent implantation follows. Coronary stents are advanced tube-shaped devices that help maintain a normal blood supply to the myocardium. All stents are basically formed of a metallic framework made from a very thin metal alloy strut of around 80 μm that is intertwined to form the tube. The stent comes folded onto another balloon catheter. After the stent is positioned, the balloon is inflated to at least 12 atm in order to fully expand the stent struts into the vessel wall. Sometimes, a postdilation is required, with a special fixed-diameter non-compliant balloon catheter in order to fully apposition the stent to the vessel wall.

After some time, usually 1 month to several years, the intracoronary stent becomes part of the vessel wall via endothelialization. This process influences the two main adverse effects of the stents: in-stent restenosis and acute thrombosis. Restenosis is a relatively slow-acting process (around 3 months to 1 year), by which the endothelial cells react abnormally to the vessel injury and the foreign body, and leads to neointimal hyperplasia. This process is a proliferation of vascular smooth cells inside the new intima and leads to the thickening of the arterial walls and a decrease in the arterial lumen. Restenosis presence varies from 30% to less than 3–5%, has numerous risk factors and is clinically relevant because of the need, in such cases, of repeated revascularization (Dangas et al. 2010). The other complication, acute thrombosis, is defined as sudden vessel closure by a locally formed thrombus, initiated by the exposure of the metallic stent struts to the arterial lumen. Usually, all patients must receive strong antiplatelet therapy for various periods of time after a stent is placed in order to prevent thrombosis, until the stent is covered by the newly formed endothelium. Sometimes the treatment is stopped before full endothelialization (usually 1 month to 1 year) and the patient is exposed to the risk of thrombosis. The in-stent acute thrombosis is a serious complication, with a mortality rate as high as 45%. The incidence of this complication varies nowadays from 0.5 to 2%, with higher numbers in the past (Stone et al. 2007).

There are different types of coronary stents and great effort has been put into research and development of the technology. The type of stent used is often linked to the patient outcome and adverse cardiovascular events. Historically, the first stent types used were the bare metal stents (BMS). These stents were plagued with high rates of restenosis in selected patients, as high as around 30%, making repeat revascularization a frequent intervention. Still, they were extensively used for many years, with significant better results regarding CAD outcome than medication, and often similar in terms of mortality with coronary arterial by-pass graft surgery (CABG) (Serruys et al. 2001). Today, the bare metal stents are still used in patients who can receive a limited duration of antiplatelet treatment, because BMS are quickly endothelialized in 1 month, the risk of stent thrombosis decreasing after this period.

The next great leap in terms of stent design was the introduction of the drug eluting stents. These are regular bare metal stents coated with a drug that blocks cell proliferation and therefore restenosis. The drugs are slowly released over one to 6 months, in the time span when the restenosis occurs (Halkin et al. 2005). The first antiproliferative agents were paclitaxel and sirolimus, but nowadays zotarolimus, biolimus and everolimus are most often used. Thanks to the decreased need of revascularization and significant better outcomes, drug eluting stents now form the bulk of the coronary stents implanted worldwide (Sarno et al. 2012). Still, because of impeded endothelialization, drug eluting stents are prone to late and very late stent thrombosis (Lüscher et al. 2007).

The newest type of coronary stents are the bioresorbable vascular scaffolds (BVS). These are stents made not from metal alloys but from fully bioresorbable materials. The most used BVS is formed from a poly-L lactide material, coated with everolimus, in order to prevent restenosis. Over time, the material is broken down and it is dissolved by the organic processes in the vessel wall. The advantages of the BVS include vasomotion restoration, reduction in plaque area and late lumen gain, but with a more difficult intracoronary deployment and emergent data regarding late and very late in-stent thrombosis (Cassese and Kastrati 2016).

2.3.3.3 Indications of Coronary Angiography and PCI

The coronary heart disease comprises a wide spectrum of clinical presentations that can be divided in three clinical entities: stable coronary disease, non-ST-segment elevation acute coronary syndromes (non STEMI ACS—including unstable angina and non-ST-segment elevation myocardial infarctions) and ST-segment elevation acute myocardial infarctions (STEMI).

Nowadays, the introduction of the PCI as the main revascularization method in STEMI, has lowered the in-hospital mortality almost four-fold, to around 4%. The European Society of Cardiology (ESC) Guidelines for STEMI indicates the primary PCI is the recommended reperfusion therapy and must be performed as soon as possible, in the first 12 h or in selected cases, in the first 24 h. In order to obtain the maximum advantage of the primary PCI, a complex network involving patient education, the emergency service, transit hospitals and PCI-capable hospitals, must be put in place to bring the patient in the cath lab as soon as possible. The primary PCI is the best treatment of STEMI in terms of outcome because of the high rate of success in experienced centres, the prompt revascularization leading to a greater salvage of myocardium and extremely low reinfarction rates (Steg et al. 2012).

The non-STEMI ACS is a heterogeneous clinical entity, comprising from aggravating exertion angina to an abrupt coronary occlusion without the ECG manifestation of ST-segment elevation. The ESC Guidelines of non-STEMI ACS has established a risk stratification score with the invasive evaluation as the centre of the treatment strategy. In the highest risk group, the coronary angiogram must be performed immediately, in the second risk group it must be performed in the first 24 h, in the third risk group in the first 72 h, and in the least risk group a non-invasive evaluation must be performed first (Roffi et al. 2015).

All acute coronary syndromes are great medical emergencies that pose a real threat to human lives. Therefore, the invasive coronary evaluation is mandatory and its benefits, and that of subsequent revascularization, overwhelm the procedure and stent-associated risks. But the patients with acute coronary syndromes represent only a part of the great coronary atherosclerotic burden. In stable angina patients, the benefit of revascularisation compared to the optimal medical treatment is less pronounced, and in fact some clinical trials have showed that there is no benefit on mortality between the two groups. In all clinical studies comparing the two groups though, there was a pronounced and consistent benefit regarding the need for revascularisation.

The ESC Guidelines for stable angina argue that a coronary invasive evaluation should be performed only in patients with severe stable angina or in patients with a clinical profile suggesting a high clinical risk, in patients with a prior non-invasive test that suggest a high adverse event risk, or in patients with inconclusive or conflicting tests.

The guidelines also have some strict indications regarding revascularization. Because of the heterogeneity of the coronary anatomy and disease, prognostic indications have been made for the unstable angina. Revascularization is suited in patients with a left main disease stenosis of more than 50%, any proximal LAD stenosis >50%, two- or three-vessel disease associating an impaired left ventricle function, a large area of ischemia in the left ventricle (>10%) as defined by non-invasive studies, and a single remaining patent artery with a >50% stenosis. Besides anatomic prognostic criteria, there is also another criterion, regarding symptoms, the presence of a coronary stenosis of more than 50% with angina, unresponsive to medical therapy (Montalescot et al. 2013).

The non-invasive testing for coronary ischemia have their rationale in filtering out and also guiding patients with stable symptoms towards coronary invasive evaluation and subsequent revascularisation. Although a relatively safe procedure as shown above, there is no need to expose numerous patients to an invasive procedure if they do not need it. Non-invasive tests are grouped into three categories: the exercise stress ECG, imaging studies and coronary computed tomography angiography. The ECG stress test is widely available and cheap, but with the lowest sensitivity and specificity (Morise and Diamond 1995). The imaging studies comprise the exercise or pharmacological stress echocardiography, pharmacological stress magnetic resonance imaging, pharmacological stress PET and exercise or pharmacological SPECT (single photon emission computed tomography). These tests are highly accurate but expensive, they are not widely available, difficult to perform and necessitate highly trained personnel (Heijenbrok-Kal et al. 2007; de Jong et al. 2012). The coronary computer tomography angiography has difficulties in assessing the coronary anatomy if the vessel are calcified and has a high radiation dose (Miller et al. 2008).

2.3.3.4 Adjunctive Invasive Diagnostic Tools

Even nowadays, in most coronary procedures, angiograms and PCIs, visual assessment of the stenosis, the vessel diameter and the size of the stent is still the most widely used method. Numerous studies showed that visual assessment is not adequate and is not well correlated with actual measurements and with a large inter-observer variability (Fischer et al. 2002; D'Ascenzo and Barbero 2015; Tonino et al. 2009).

Regarding coronary anatomy, there are two methods employed in correctly assessing the vessel diameter, extent of the lesion, stent deployment and apposition as well as important information regarding atherosclerotic plaque physiopathology.

The first method is the intravascular ultrasound (IVUS). A small probe is inserted in the coronary artery and using ultrasound technology the vessel wall is visualised. It has a penetrance of 4–8 mm and a resolution of 150 μm. The IVUS is the gold standard for accurate measurement of plaque burden and is used to determine the evolution of the plaque under different drugs. It is also used for stent implantation optimization and many clinical trials have assessed the clinical outcome of stent placement under IVUS guidance with mixed results. The high price of the probe limits its use and the routine IVUS evaluation is not recommended (Windecker et al. 2014; Mintz et al. 2001).

The other method uses optical coherence tomography (OCT). A probe is inserted in the coronary artery and using a light based technology, it accurately detects intraluminal structures, as it has a resolution of 15 μm, but a penetrance of only 2–3 mm. Because of such a high resolution, structures like lipid pools, fibrous plaque, white or red thrombi and even the thin fibrous cap of the vulnerable atherosclerotic plaque can be seen. It is the standard method of evaluating the BVS placement and optimization, as the BVS struts are echolucent (Prati et al. 2012).

Regarding coronary physiology, the most important breakthrough is the determination of the fractional flow reserve (FFR). It is the gold standard method for evaluation of the functional assessment of a lesion severity. The visual assessment on the angiogram is a poorly correlated method to the FFR, as many >70% (the classical threshold for revascularization) stenosis have a non-significant FFR and also many <70% stenosis on visual assessment have a positive FFR and a need to treat (Fischer et al. 2002). The method implies the introduction of a special coronary 0.0014″ guidewire, with a pressure sensor at its tip, distally to the stenosis to be assessed. Different drugs (papaverine, adenozine, ATP, isosorbid dinitrate) are then administered either intravenously or intracoronary in order to obtain a full microcirculation hyperemia, to mimic the maximal blood flow distally to the stenosis. The ratio between the pressure distally to the lesion and the aortic pressure is then calculated. A ratio bellow 0.75–0.80 shows that it is a functionally significant lesion and must be treated, while a ratio over 0.80 may be left untreated. Randomized clinical trials with hard endpoints established the safety and effectiveness of the 0.80 threshold (Tonino et al. 2009; Pijls et al. 2007; De Bruyne et al. 2012).

2.4 Aortic Pathologies

2.4.1 Background

Aortic pathologies comprise a wide spectrum of arterial diseases: aortic aneurysms, acute aortic syndromes including aortic dissection, intramural hematoma, penetrating atherosclerotic ulcer and traumatic aortic injury, pseudoaneurysm, aortic rupture, atherosclerotic and inflammatory affections, as well as genetic diseases (e.g. Marfan syndrome) and congenital abnormalities, including the coarctation of the aorta (CoA) (Erbel et al. 2014).

This large spectrum of pathologies may be diagnosed after a long period of subclinical development or they may be discovered under an acute presentation, as acute aortic syndrome, a clinical condition that needs rapid diagnosis and decision-making to reduce major complications or even death.

Coarctation of the aorta, a congenital abnormality considered to be a complex disease of the vasculature, occurs in approximately 5–8% of patients with congenital heart disease. The prevalence of isolated forms is 3 per 10,000 live births (Erbel et al. 2014).

It occurs at least 1.5–2 times more often in males than in females (Warnes 2009). Although the majority of the cases are not inherited, there have been identified possible genetic associations with left-sided obstructive lesions that include coarctation of the aorta, left ventricular outflow obstruction, subaortic stenosis, bicuspid aortic valve and supravalvular aortic stenosis (Aboulhosn and Child 2006).

CoA covers multiple aspects of anatomy and physiology and the severity of the disease depends not only on the degree of aortic narrowing but also on associated heart lesions. Timely repair of coarctation with surgery or catheter intervention relieves symptoms and improves prognosis (Crawford 2010).

This chapter focuses on coarctation of the aorta, and, before presenting the evaluation strategies of this pathology, several physiopathological aspects are discussed.

2.4.2 Anatomy, Physiology and Physiopathology

The aorta, the largest artery in the body, is anatomically divided into thoracic and abdominal segments. The former, the thoracic segment, is subdivided into the ascending aorta, the aortic arch and the descending aorta. The latter has two distinct portions, in relation to the renal arteries, called supra- and infrarenal segments (Mann et al. 2015).

The aortic wall is histologically composed of three layers, playing an important role in the physiopathology of aortic diseases: tunica intima, a thin layer lined by the endothelium, tunica media, arranged as concentric sheets of elastic and collagen fibers with the border zone of the lamina elastica interna and externa, and also smooth muscle cells. The outer layer, called tunica adventitia, contains collagen fibers, vasa vasorum and lymphatics (Erbel et al. 2014).

The media gives the aorta its circumferential elasticity, necessary to resist hemodynamic stress. The adventitial collagen fibers ultimately govern the tensile strength of the aortic wall. Oxygen and nutrients are supplied to the aortic wall by diffusion from the lumen (Mann et al. 2015).

Coarctation of the aorta represents a narrowing of the aorta that causes a restriction of blood flow to the more distal vessels and organs, and hypertension proximal to the CoA (Warnes 2009).

In the typical form of coarctation of the aorta, the narrowing is located near the thoracic end of patent ductus arteriosus (PDA), usually in a distal point to the left subclavian artery, depending on the length of the isthmus (Crawford 2010). In patients with coarctation the ascending aorta is usually dilated and the severity of the narrowing is variable. The dilatation may become aneurysmal and also involve sinuses of Valsalva. In many cases, coarctation of the aorta is often associated with other cardiac malformations, of which the most common is a bicuspid aortic valve, giving rise to another associated pathology: aortic stenosis or aortic regurgitation. It may associate shunt lesions, such as patent ductus arteriosus or ventricular septal defect (Willerson et al. 2007). It is also important to mention the possible coexistence of a noncardiac anomaly in this category of patients, the aneurysm of the circle of Willis (Hodes 1959).

Aortic coarctation can rarely occur in the ascending intrathoracic or abdominal aorta and it may be associated with systemic disease, such as Takayasu's aortitis, Williams syndrome and middle aortic syndrome (Crawford 2010).

Based on the elastic properties, the aorta transmits pulsatile arterial blood pressure to all points in the vascular tree (Mann et al. 2015). Through its elasticity, the aorta has the role of a second pump—Windkessel function/effect—during diastole, being essential to ensure coronary perfusion, and also perfusion for the rest of the organs (Erbel et al. 2014).

Coarctation of the aorta is a diffuse disease, an affirmation based on histologic examinations that show intimal and medial lesions consisting of a thickened shelf that protrudes posteriorly and laterally into the aortic lumen. This shelf is related to the wall of the ductus, and encircles the vessel. These histological aspects reveal that the coarctation develops as a result of migration of ductus smooth muscle cells into the periductal aorta, resulting in a subsequent constriction and narrowing of the aortic lumen (Crawford 2010).

The impedance of aortic narrowing in coarctation causes a chronic increase of afterload for the left ventricle. Blood flow to the head and upper extremities is preserved because the vessels supplying these areas branch off from the proximal segment of the aorta. Blood flow to the descending aorta and lower extremities is usually diminished. Compensatory mechanisms, such as left ventricular hypertrophy, dilatation of collateral blood vessels from the intercostal arteries, to provide blood flow to the descending aorta, will appear if the primary cause is not corrected (Lilly 2011).

2.4.3 Natural History and Clinical Presentation

Coarctation of the aorta is likely to produce consequences during two periods of life: in early infancy (if the lesion is severe) and between 20 and 30 years of age. If the coarctation is not corrected only 10% of the patients live beyond age 50 (Willerson et al. 2007).

The clinical aspects of coarctation depend largely on the nature of the associated cardiac lesions and the severity of the narrowing. In early infancy, severe coarctation and cardiac-associated anomalies may lead to congestive heart failure and cardiac shock, often occurring after the closure of PDA, which before had the role of bypassing the narrowed area. Those cases must be managed under administration of intravenous prostaglandin to maintain patency of the ductus arteriosus, preserving systemic blood flow until surgery can be performed (Warnes 2009).

Isolated coarctation may be mostly asymptomatic and may have subtle clinical findings. In these cases, it is frequently misdiagnosed and treatment is delayed. In older children and adolescents, with less severe types of coarctation, systolic hypertension and heart murmur (most prominent in the left paravertebral region, near the left scapula) are the most common presentations (Crawford 2010).

A complete cardiovascular examination may strongly suggest the correct diagnosis. Normally, the lower extremity blood pressure is higher than that of the upper extremities. The classical coarctation patient presents lower blood pressure in the legs than in the right upper extremity. This diminishes the femoral pulse, combined with delayed timing, when compared with the right upper extremity pulses. The left arm may be variable depending on the area of the left subclavian artery relative to coarctation site (Warnes 2009). A significant blood pressure gradient (above 20 mmHg) between the upper and lower extremities indicates significant coarctation of the aorta (Erbel et al. 2014).

Another clinical finding is increased left ventricular impulse, that may be displaced, secondary to left ventricular hypertrophy, signs of collateral circulation and additional murmurs (Warnes 2009).

Aortic coarctation may also be associated with Marfan or Turner's syndrome. The somatic features in those cases should be correctly evaluated (Willerson et al. 2007).

2.4.4 Diagnostic Work-Up

If coarctation of the aorta is suspected, various tests enable a definitive diagnosis.

2.4.4.1 Chest Radiography

In infants with coarctation of the aorta, enlargement of the heart may be observed on the chest X-ray examination, but the outline may vary. If the child has an intracardiac lesion with left-to-right shunting, moderate to severe cardiomegaly with increased pulmonary vascular drawing may be observed (Crawford 2010). Rib notching can be seen as irregular areas on the undersurfaces of the posterior ribs. Those marks seldom appear before patients reach 7 years of age. The bilateral rib notching, usually confined to the third to eighth posterior rib, indicates that the area of coarctation is located just distal to the subclavian artery (Willerson et al. 2007). An indented aorta at the site of coarctation may also be visualized (Lilly 2011).

2.4.4.2 Electrocardiography

In general, right axis deviation and right ventricular hypertrophy can be seen in an infant. In adults, left ventricular hypertrophy is usually evident because of the long-term effects of left ventricular pressure overload. In case of associated right ventricular hypertrophy in adults, it is required to evaluate for another cardiac lesion (ventricular septal defect and pulmonary hypertension) (Crawford 2010).

2.4.4.3 Echocardiography

Ultrasound diagnosis for coarctation of the aorta represents a non-invasive, inexpensive investigation, readily accepted by every patient and it can be repeated as frequently as required. It is an important tool for quantification of left ventricular mass and for searching other possible intracardiac abnormalities (Sun et al. 2015).

In neonates and infants, echocardiography from the suprasternal axis view provides a useful method for evaluating the aortic arch, the branching arteries and possible cardiac anomalies. Color-flow mapping can assist the severity of the aortic coarctation. A turbulent flow which travels downstream is usually associated with the coarctation and studies on both pulse-wave Doppler and continuous-wave Doppler can provide a calculation of the pressure gradient across the narrowing, using the modified Bernoulli equation. This method is influenced by a series of variable parameters, such as aortic arch geometry, complex flow dynamics in the aortic arch, difficulty in aligning the Doppler beam with flow, or inability to measure flow velocity under certain circumstances, and the use of the Bernoulli simplified equation is not always reliable. In older adults, the suprasternal long-axis view may not offer as much information, due to poor imaging windows, but it can be useful to measure ascending aorta and detect aortic aneurysm (Crawford 2010). The peak and mean velocity detected with continuous wave Doppler through the coarctation site may help define the severity of the coarctation but may overestimate the gradient with long-segment stenosis or increased arterial stiffness (Warnes 2009).

2.4.4.4 Magnetic Resonance Imaging

The possibility of employing MRI to delineate the intrinsic contrast between flow and vessel wall makes an optimal choice for diagnosis of the aortic disease (Erbel et al. 2014). However, it is an expensive investigation for many countries, and it requires valuable experience in cardiac and vascular studies (Warnes 2009).

MRI in sagittal and parasagittal projections can appreciate the location and the severity of coarctation of the aorta and the anatomy of the transverse aortic arch, isthmus and poststenotic dilatation. It may also highlight the presence of a patent ductus arteriosus and the extent of collateral arteries (Crawford 2010).

2.4.4.5 Diagnostic Cardiac Catheterization and Angiography

Cardiac catheterization and angiography are invasive procedures, used for clinical evaluation and for cases in which clinical questions arise when referring to the severity of coarctation or associated cardiac lesions (Crawford 2010). For the majority of cases these investigations are rarely required because noninvasive imaging often ensures accurate evaluation. Through catheterization, the true pressure gradient across the arch is obtained; when the peak gradient exceeds 20 mmHg in noninvasive imaging, catheterization is reserved for therapeutic interventions (Warnes 2009).

2.4.5 Management of Coarctation of the Aorta

Once a diagnosis of coarctation is performed, both surgery and catheter interventions (balloon angioplasty and stent placement) can be considered for resolving the narrowing of the aorta. It is important to know that uncorrected aortic coarctation is associated with a poor clinical course and short life expectancy—reaching an average of 30 years—and is usually secondary to clinical consequences of chronic arterial hypertension (Calderón-Colmenero and Attie 2008). Other common causes of death in the case of unrepaired coarctation cases are congestive heart failure, endocarditis and aortic rupture, as well as intracranial hemorrhage (due to rupture of aneurysm of Willis circle) (Warnes 2009).

Medical treatment implies drugs for controlling arterial hypertension with the inhibitors of the renin-angiotensin-aldosterone system (ACE inhibitors), angiotensin II type 1 receptor antagonists, or aldosterone antagonists (such as spironolactone, eplerenone), betablockers or centrally acting α-adrenergic agonists (clonidine) (Willerson et al. 2007). Neonates and infants with aortic coarctation should be admitted in the hospital for intravenous administration of prostaglandin E1 as soon as the diagnosis is confirmed, in order to reopen the ductus arteriosus for improving lower body perfusion (Moodie 2014).

The mortality rate of surgical repair for isolated coarctation of the aorta is very low in infants and older children, but it may be higher in the presence of associated lesions. Also, the morbidity after the correction may imply a rebound of arterial hypertension, neurological disorders, as paralysis due to spinal cord lesion or vocal cord paralysis, due to recurrent laryngeal nerve injury or immediate postoperative complications, like hemorrhage caused by the suture line or infections (Crawford 2010).

In a recent multicenter study, endovascular repair represented a safe and effective treatment for de novo coarctation and post-surgical complications of coarctation repair in adults, together with the lowest rates of complications and long-term clinical success (Erben et al. 2015).

In native coarctation of the aorta with appropriate anatomy, stenting has become the first choice treatment in adults in many centers. In these cases, appropriate antihypertensive treatment may still be necessary for controlling hypertension (Erbel et al. 2014).

2.4.6 Long Term Prognosis

Long term prognosis after CoA repair depends on age, associated lesions and systemic complications. Despite successful repair, the prevalence of hypertension in adults under 20 years varies from 20 to 40% and it may be higher in some cases. In addition to persisting arterial hypertension, a considerable number of patients with repaired coarctation are normotensive at rest, but develop a hypertensive response to exercise (Crawford 2010).

Clearly, the coarctation of the aorta represents a lifelong disorder consisting of more than a narrowing of an aortic segment. After surgical or catheter-based intervention these patients require persistent follow-up, preferably in a center that specializes in adults with congenital heart disease. Along with the regular follow-up and specific medical therapy, it is recommended that such patients undergo physical exercise and cardiovascular prevention training in order to lower their cardiovascular risk.

References

Aboulhosn J, Child JS (2006) Left ventricular outflow obstruction: subaortic stenosis, bicuspid aortic valve, supravalvar aortic stenosis, and coarctation of the aorta. Circulation 114(22):2412–2422. http://circ.ahajournals.org/cgi/pmidlookup?view=long&pmid=17130357

Alonso M, Tascón J, Hernández F, Andreu J, Albarrán A, Velázquez MT (2003) Complications with femoral access in cardiac cathetization. Impact of previous systematic femoral angiography and hemostasis with VasoSeal-ES® collagen plug. Rev Esp Cardiol 56(6):569–577

American Society of Nuclear Cardiology Images (1999) Imaging guidelines form nuclear cardiology procedures part II. J Nucl Cardiol 6:G47–G84

Anderson RH, Baker EJ et al (2010) Paediatric cardiology, 3rd edn. Churchill Livingstone, London

Anderson RH, Spicer DE et al (2013) Wilcox's surgical anatomy of the heart, 4th edn. Cambridge University Press, Cambridge

Arciero TJ, Jacobsen SJ, Reeder GS et al (2004) Temporal trends in the incidence of coronary disease. Am J Med 117:228

Ashley EA, Niuebauer J (2004) Cardiology explained. Remedica, London

Baim DS, Grossman W (1996) Complications of cardiac catheterization. In: Cardiac catheterization, angiography and intervention. Williams & Wilkins, Baltimore, p 17

Barbosa D, Dietenbeck T, Heyde B, Houle H, Friboulet D, D'hooge J, Bernard O (2013) Fast and fully automatic 3-d echocardiographic segmentation using B-spline explicit active surfaces: feasibility study and validation in a clinical setting. Ultrasound Med Biol. 39(1):89–101. doi:10.1016/j.ultrasmedbio.2012.08.008

Barry FA et al (2014) Impact of diet and exercise on lipid management in the modern era. Best Pract Res Clin Endocrinol Metab 28(3):405–421. http://www.bprcem.com/article/S1521-690X(14)00021-9/abstract

Baskot B (2011) Coronary angiography—advances in noninvasive imaging approach for evaluation of coronary artery disease. ISBN 978-953-307-675-1. https://www.intechopen.com/books/coronary-angiography-advances-innoninvasive-imaging-approach-for-evaluation-of-coronary-artery-disease/contrast-echocardiography-incoronary-artery-disease. Published 15 Sept 2011 under CC BY-NC-SA 3.0 license

Betriu A, Heras M, Cohen M et al (1992) Unstable angina: outcome according to clinical presentation. J Am Coll Cardiol 19:1659–1663

Bouzas-Mosquera A, Peteiro J, Alvarez-García N et al (2009) Prognostic value of exercise echocardiography in patients with left bundle branch block. JACC Cardiovasc Imaging 2:251–259

Buja LM, McAllister HA Jr (2007) Atherosclerosis: pathologic anatomy and pathogenesis. In: Willerson JT, Cohn JN, Wellens HJJ, Holmes DR Jr (eds) Cardiovascular medicine, 3rd edn. Springer, London, pp 1581–1591

Burke AP, Farb A, Malcom GT et al (1997) Coronary risk factors and plaque morphology in men with coronary disease who died suddenly. N Engl J Med 336:1276–1282

Burke AP, Kolodgie FD, Farb A et al (2001) Healed plaque ruptures and sudden coronary death: evidence that subclinical rupture has a role in plaque progression. Circulation 103:934–940

Calderón-Colmenero J, Attie F (2008) Aortic coarctation. important considerations in long-term follow-up after correction. Rev Esp Cardiol 61(11):1117–1119. Available from http://linkinghub.elsevier.com/retrieve/pii/S1885585709600232

Califf RM (1998) Ten years of benefit from a one-hour intervention. Circulation 98:2649–2651

Cassese S, Kastrati A (2016) Bioresorbable vascular scaffold technology benefits from healthy skepticism. J Am Coll Cardiol 67(8):932–935. doi:10.1016/j.jacc.2015.12.020

Cerqueira MD, Weissman NJ, Dilsizian J et al (2002) Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals form the Cardiac Imaging Comitee of the Council of Clinical Cardiology of the American Heart Association. Circulation 105:539–542

Cortigiano L, Picano E, Vigna C et al (2001) Prognostic value of pharmacologic stress echocardiography in patients with left bundle branch block. Am J Med 110:361–389

Crawford MH (ed) (2010) Cardiology. Elsevier Saunders, Philadelphia

D'Ascenzo F, Barbero U (2015) Accuracy of intravascular ultrasound and optical coherence tomography in identifying functionally significant coronary stenosis according to vessel diameter. Am Heart J 169(5):663–673

Dangas GD, Claessen BE, Caixeta A, Sanidas EA, Mintz GS, Mehran R (2010) In-stent restenosis in the drug-eluting stent era. J Am Coll Cardiol 56(23):1897

Davies MJ, Bland JM, Hangartner JR et al (1989) Factors influencing the presence or absence of acute coronary artery thrombi in sudden ischaemic death. Eur Heart J 10:203–208

De Bruyne B, Pijls NH, Kalesan B, Barbato E, Tonino PA, Piroth Z, Jagic N, Mobius-Winkler S, Rioufol G, Witt N, Kala P, MacCarthy P, Engstrom T, Oldroyd KG, Mavromatis K, Manoharan G, Verlee P, Frobert O, Curzen N, Johnson JB, Juni P, Fearon WF, Investigators FT (2012) Fractional flow reserve-guided PCI vs. medical therapy in stable coronary disease. N Engl J Med 367(11):991–1001

Dewey M (2009) Coronary CT angiography. Springer, Heidelberg

Ehara S, Kobayashi Y, Yoshiyama M et al (2004) Spotty calcification typifies the culprit plaque in patients with acute myocardial infarction: an intravascular ultrasound study. Circulation 110:3424–3429

Erbel R, Aboyans V, Boileau C, Bossone E, Di Bartolomeo R, Eggebrecht, et al (2014) 2014 ESC guidelines on the diagnosis and treatment of aortic diseases. Eur Heart J 35(41):2873–2926. https://academic.oup.com/eurheartj/article-lookup/doi/10.1093/eurheartj/ehu281

Erben Y, Oderich GS, Verhagen HJ, Witsenburg M, Van den Hoven AT, Debus ES et al (2015) Abstract 10340: multi-center experience with endovascular treatment of aortic coarctation and post-surgical complications of coarctation repair in adults. Circulation 132(suppl 3):A10340–A10340

Estruch R et al (2013) Primary prevention of cardiovascular disease with a mediterranean diet. N Engl J Med 368:1279–1290

Falk E (1983) Plaque rupture with severe pre-existing stenosis precipitating coronary thrombosis. Characteristics of coronary atherosclerotic plaques underlying fatal occlusive thrombi. Br Heart J 50:127–134

Falk E, Shah PK, Fuster V (1995) Coronary plaque disruption. Circulation 923:657–671

Falk E, Shah PK, Fuster V (2004) Atherothrombosis and thrombosis-prone plaques. In: Fuster V, Alexander RW, O'Rourke RA (eds) Hurst's the heart, 11th edn. McGraw-Hill, New York, pp 1123–1139

Fischer JJ, Samady H, McPherson JA, Sarembock IJ, Powers ER, Gimple LW, Ragosta M (2002) Comparison between visual assessment and quantitative angiography versus fractional flow reserve for native coronary narrowings of moderate severity. Am J Cardiol 90(3):210–215

Ford ES, Ajani UA, Croft JB et al (2007) Explaining the decrease in U.S. deaths from coronary disease, 1980–2000. N Engl J Med 356:2388

Fuster V, Walsh RA, Harrington RA (eds) (2011) Hurst's the heart, 13th edn. McGraw-Hill, New York

Gage JE, Hess OM, Murakami T et al (1986) Vasoconstriction of stenotic coronary arteries during dynamic exercise in patients with classic angina pectoris: reversibility by nitroglycerin. Circulation 73:865–876

Gallagher KP, Matsuzaki M, Koziol JA et al (1984) Regional myocardial perfusion and wall thickening during ischaemia in conscious dogs. Am J Phys 247:H727–H738

Gimbrone MA (2010) The Gordon Wilson lecture. Understanding vascular endothelium: a pilgrim's progress. Endothelial dysfunction, biomechanical forces and the pathobiology of atherosclerosis. Trans Am Clin Climatol Assoc 121:115–127

Halkin A, Mehran R, Casey CW et al (2005) Impact of moderate renal insufficiency on restenosis and adverse clinical events after paclitaxel-eluting and bare metal stent implantation: results from the TAXUS-IV Trial. Am Heart J 150:1163

Hannawi B, Lam WW, Wang S, Younis GA (2014) Current use of fractional flow reserve: a nationwide survey. Tex Heart Inst J 41(6):579–584. doi:10.14503/THIJ-13-3917

Heijenbrok-Kal MH, Fleischmann KE, Hunink MG (2007) Stress echocardiography, stress single-photon-emission computed tomography and electron beam computed tomography for the assessment of coronary artery disease: a meta-analysis of diagnostic performance. Am Heart J 154:415–423

Heimdal A, Stoylen A, Storp H et al (1998) Real time strai rate imaging of the left ventricle by ultrasound. J Am Soc Echocardiogr 11:1013–1019

Hodes HL, Steinfeld L, Blumenthal S (1959) Congenital cerebral aneurysms and coarctation of the aorta. Arch Pediatr 76(1):28–43. Available from http://www.ncbi.nlm.nih.gov/pubmed/13618092

Hollander W (1976) Role of hypertension in atherosclerosis and cardiovascular disease. Am J Cardiol 38(6):786–800

Hutcheson JD, Maldonado N, Aikawa E (2014) Small entities with large impact: microcalcifications and atherosclerotic plaque vulnerability. Curr Opin Lipidol 25(5):327–332

Jenkins C, Bricknell K, Chan J et al (2007) Comparision of two and three dimensional echocardiography with sequential magnetic resonance imaging for evaluating left ventricular volume and ejection fraction over time in patients with healed myocardial infarction. Am J Cardiol 99:300–306

Johnson LW, Lozner EC, Johnson S, Krone R, Pichard AD, Vetrovec GW, Noto TJ (1989) Coronary arteriography 1984–1987: a report of the registry of the society for cardiac angiography and interventions. I. Results and complications. Catheter Cardiovasc Diagn 17(1):5

de Jong MC, Genders TS, van Geuns RJ, Moelker A, Hunink MG (2012) Diagnostic performance of stress myocardial perfusion imaging for coronary artery disease: a systematic review and meta-analysis. Eur Radiol 22:1881–1895

Kerber RE, Abboud FM (1973) Echocardiographic detection of regional myocardial infarction. An experimental study. Circulation 4:997–1005

Kirkpatrick JN, Lang RM, Savitri E (2005) Automated border detection on contrast enhanced echocardiographic images. Int J Cardiol 103:164–167

Klabunde RE (2012) Cardiovascular physiology concepts. Lippincot Williams & Wilkins, Baltimore, p 257

Kolodgie FD, Burke AP, Farb A et al (2001) The thin-cap fibroatheroma: a type of vulnerable plaque: the major precursor lesion to acute coronary syndromes. Curr Opin Cardiol 16:285–292

Koskinas KC et al (2010) Natural history of experimental coronary atherosclerosis and vascular remodeling in relation to endothelial shear stress: a serial, in vivo intravascular ultrasound study. Circulation 121(19):2092–2101

Krueger K, Zaehringer M, Strohe D, Stuetzer H, Boecker J, Lackner K (2005) Postcatheterization pseudoaneurysm: results of US-guided percutaneous thrombin injection in 240 patients. Radiology 236(3):1104

Kwok CS, Kontopantelis E, Myint PK, Zaman A, Berry C, Keavney B, Nolan J, Ludman PF, de Belder MA, Buchan I, Mamas MA (2015) Stroke following percutaneous coronary intervention: type-specific incidence, outcomes and determinants seen by the British Cardiovascular Intervention Society 2007–12. Eur Heart J 36(25):1618–1628. doi:10.1093/eurheartj/ehv113. Epub 2015 Apr 20

Lang RM, Badano LP, Mor Avi V et al (2015) Recommendation for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr 28:1–39

Leitman M, Lysysansky P, Sidenko S et al (2004) Two-dimensional strain – a novel software for real time quantitative echocardiographic assessment of myocardial function. J Am Soc Echocardiogr 17:1021–1029

Lenihan DJ, Gerson MC, Hoit BD et al (1995) Mechanisms, diagnosis, and treatment of diastolic heart failure. Am Heart J 130:153–166

Lilly LS (ed) (2011) Pathophysiology of heart disease: a collaborative project of medical students and faculty, 5th edn. Lippincott Williams & Wilkins, Baltimore

Lüscher TF, Steffel J, Eberli FR et al (2007) Drug-eluting stent and coronary thrombosis: biological mechanisms and clinical implications. Circulation 115:1051

Lyoyd-Jones D, Adams RJ, Brown TM et al (2010) Executeve summary: heart disease and stroke statistics-2010 update: a report from American Heart Association. Circulation 121:948

Mann JM, Kaski JC, Pereira WI et al (1998) Histological patterns of atherosclerotic plaques in unstable angina patients vary according to clinical presentation. Heart 80:19–22

Mann DL, Zipes DP, Libby P, Bonow RO, Braunwald E (2015) Braunwald's heart disease: a textbook of cardiovascular medicine, 10th edn. Elsevier-Saunders, Philadelphia. https://www.amazon.com/Braunwalds-Heart-Disease-E-Book-Cardiovascular-ebook/dp/B00MEIYKXY/ref=mt_kindle?_encoding=UTF8&me=

Martin JM, Stephens JC (2013) Cardiovascular hemodynamics [Internet]. 204:742. Available from http://link.springer.com/10.1007/978-1-60761-195-0

McEvoy JW, Blaha MJ et al (2015) Cigarette smoking and cardiovascular events. role of inflammation and subclinical atherosclerosis from the multiethnic study of atherosclerosis. Arterioscler Thromb Vasc Biol 35:700–709

Miller JM, Rochitte CE, Dewey M, Arbab-Zadeh A, Niinuma H, Gottlieb I, Paul N, Clouse ME, Shapiro EP, Hoe J, Lardo AC, Bush DE, de Roos A, Cox C, Brinker J, Lima JA (2008) Diagnostic performance of coronary angiography by 64-row CT. N Engl J Med 359:2324–2233

Mintz GS, Nissen SE, Anderson WD, Bailey SR, Erbel R, Fitzgerald PJ, Pinto FJ, Rosenfield K, Siegel RJ, Tuzcu EM, Yock PG (2001) American College of Cardiology Clinical Expert Consensus Document on Standards for Acquisition, Measurement, Reporting of Intravascular Ultrasound Studies (IVUS). A report of the American College of Cardiology Task Force on Clinical Expert Consensus Documents. J Am Coll Cardiol 37(5):1478–1492

Mittleman MA, Mostofsky E (2011) Physical, psychological and chemical triggers of acute cardiovascular events: preventive strategies. Circulation 124:346–354

Moller JE, Hillis GS, Jae KO (2006) Wall motion score index and ejection fraction for risk stratification after acute myocardial infarction. Am Heart J 151:419–425

Montalescot G, Sechtem U, Achenbach S et al (2013) ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 34:2949

Moodie DS (ed) (2014) Clinical management of congenital heart disease from infancy to adult. Cardiotext Publishing, Minneapolis

Morise AP, Diamond GA (1995) Comparison of the sensitivity and specificity of exercise electrocardiography in biased and unbiased populations of men and women. Am Heart J 130:741–747

Moscucci M (ed) (2014) Grossman & Baim's cardiac catheterization, angiography and intervention, 8th edn. Lippincot Williams & Wilkins, Baltimore

Mukherjee D, Bates ER, Roffi M, Moliterno DJ (2010) Cardiovascular catheterisation and intervention. In: A textbook of coronary, peripheral and structural heart disease. Informa Healthcare, New York

Muller JE, Stone PH, Turi ZG, Rutherford JD, Czeisler CA, Parker C, Poole WK, Passamani E, Roberts R, Robertson T (1985) Circadian variation in the frequency of onset of acute myocardial infarction. N Engl J Med 313:1315–1322

Nihoyannopoulos P, Kaski J-C, Crake T et al (1991) Absence of myocardial dysfunction during stress in patients with syndrome X. J Am Coll Cardiol 18:1463–1470

Ogawa S, Fujii I, Yoshino H et al (1985) Values of electrocardiography and two dimensional echocardiography to identify myocardial infarction due to left circumflex and right coronary artery disease. Clin Cardiol 8:269–275

Ohman EM et al, REACH Registry Investigators (2006) The REduction of atherothrombosis for continued health (REACH) Registry: an international, prospective, observational investigation in subjects at risk for atherothrombotic events-study design. Am Heart J 151(4):786.e1–10

Oudkerk M, Reiser MF (2008) Coronary radiology, 2nd revised edn. Springer, New York

Pappano AJ, Wier WG (2013) Cardiovascular physiology, 10th edn. Elsevier-Mosby, Philadelphia

Paulson EK, Sheafor DH, Kliewer MA, Nelson RC, Eisenberg LB, Sebastian MW, Sketch MH Jr (2000) Treatment of iatrogenic femoral arterial pseudoaneurysms: comparison of US-guided thrombin injection with compression repair. Radiology 215(2):403–408

Picano E (1989) Dipyridamole-echocardiography test: historical background and physiologic basis. Eur Heart J 10:365–376

Pierard LA, Sprynger M, Gilis F (1986) Significance of precordial ST segment depression in inferioa acute myocardial infarction as determined by echocardiography. Am J Cardiol 57:82–85

Pijls NH, van Schaardenburgh P, Manoharan G, Boersma E, Bech JW, van't Veer M, Bar F, Hoorntje J, Koolen J, Wijns W, de Bruyne B (2007) Percutaneous coronary intervention of functionally nonsignificant stenosis: 5-year follow-up of the DEFER Study. J Am Coll Cardiol 49(21):2105–2111

Popescu BA, Antonini Canterin F, Gianuzii P et al (2004) Left atrium remodeling after acute myocardial infarction: results of the GISSI-3 echo sub-study. J Am Coll Cardiol 43(9):1156

Prati F, Guagliumi G, Mintz GS, Costa M, Regar E, Akasaka T, Barlis P, Tearney GJ, Jang IK, Arbustini E, Bezerra HG, Ozaki Y, Bruining N, Dudek D, Radu M, Erglis A, Motreff P, Alfonso F, Toutouzas K, Gonzalo N, Tamburino C, Adriaenssens T, Pinto F, Serruys PW, Di Mario C (2012) Expert review document part 2: methodology, terminology and clinical applications of optical coherence tomography for the assessment of interventional procedures. Eur Heart J 33(20):2513–2520

Ramanathan T, Skinner H (2005) Coronary blood flow. Contin Educ Anaesth Crit Care Pain 5(2):61–64

Reiner Z et al (2011) ESC/EAS Guidelines for the management of dyslipidemias. Eur Heart J 32(14):1769–1818

Roffi M, Patrono C, Collet JP et al (2015) 2015 ESC guidelines for the management of acute coronary syndrome in patients presenting without persistent ST-segment elevation. Eur Heart J. doi:10.1093/eurheartj/ehv320

Rosa S, Nihoyannopoulos P, Evangelista A et al (2008) Stress echocardiography expert consensus statement. Eur J Echocardiogr 9:415–437

Sarno G, Lagerqvist B, Fröbert O et al (2012) Lower risk of stent thrombosis and restenosis with unrestricted use of ‘new-generation’ drug-eluting stents: a report from the nationwide Swedish Coronary Angiography and Angioplasty Registry (SCAAR). Eur Heart J 33:606

Schaar JA, Muller JE, Falk E et al (2004) Terminology for high-risk and vulnerable coronary artery plaques. Report of a meeting on the vulnerable plaque, June 17 and 18, 2003, Santorini, Greece. Eur Heart J 25:1077–1082

Senior R, Lahiri A (1995) Enhanced detection of myocardial ischaemia by stress dobutamine echocardiography utilizing the ‘biphasic’ response of wall thickening during low and high dose dobutamine infusion. J Am Coll Cardiol 26:26–32

Serruys PW, Unger F, Sousa JE, Jatene A, Bonnier HJ, Schonberger JP, Buller N, Bonser R, van den Brand MJ, van Herwerden LA, Morel MA, van Hout BA (2001) Comparison of coronary-artery bypass surgery and stenting for the treatment of multivessel disease. N Engl J Med 344(15):1117–1124

Serruys PW, Wijns W, Vahanian A, van Sambeek A, Rodney de Palma Eric Eeckhout (2010) Percutaneous interventional cardiovascular medicine—The PCR-EAPCI textbook. https://www.pcronline.com/eurointervention/textbook/pcr-textbook/table-of-contents/

Servoss SJ, Januzzi JL, Muller JE (2002) Triggers of acute coronary syndromes. Prog Cardiovasc Dis 44:369–380

Sherif FN, Christopher PA, Thierry CG et al (2009) Al recommendations for the evaluation of left ventricular diastolic function by echocardiography. J Am Soc Echocardiogr 22:107–133

Shimada E, Zhu M, Kimura S, Streiff C, Houle H, Datta S, Sahn DJ, Ashraf M (2015) Quantitative assessment of mitral inflow and aortic outflow stroke volumes by 3-dimensional real-time full-volume color flow doppler transthoracic echocardiography: an in vivo study. J Ultra Med 34(1):95–103

Smith DL, Fernhall B (2011) Advanced cardiovascular exercise physiology [Internet], p 227. Available from http://books.google.com/books?id=gBExM6v-XPcC&pgis=1

Steg G, James SK, Atar D, Badano LP, Lundqvist C, Borger MA et al (2012) ESC guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J 33:2569–2619

Stone GW, Moses JW, Ellis SG, Schofer J, Dawkins KD, Morice MC, Colombo A, Schampaert E, Grube E, Kirtane AJ, Cutlip DE, Fahy M, Pocock SJ, Mehran R, Leon MB (2007) Safety and efficacy of sirolimus- and paclitaxel-eluting coronary stents. N Engl J Med 356(10):998

Stone GW et al, PROSPECT Investigators (2011) A prospective natural-history study of coronary atherosclerosis. N Engl J Med 364:226–235

Sun Z, Cheng TO, Li L, Zhang L, Wang X, Dong N et al (2015) Diagnostic value of transthoracic echocardiography in patients with coarctation of aorta: the Chinese experience in 53 patients studied between 2008 and 2012 in one major medical center. PLoS One 10(6):e0127399

Temporelli PL, Gianuzii P, Nicolosi GL et al (2004) Doppler derived deceleration time as a strong prognostic marker of left ventricular remodeling and survival aftare acute myocardial infarction: results of the GISSI-3 echo sub-study. J Am Coll Cardiol 43:1646–1653

Tennant R, Wiggers CJ (1935) The effect of coronary occlusion on myocardial contraction. Am J Phys 112:351–361

Thavendiranathan P, Liu S, Datta S, Walls M, Nitinunu A, Van Houten T, Tomson NA, Vidmar L, Georgescu B, Wang Y, Srinivasan S, De Michelis N, Raman SV, Ryan T, Vannan MA (2012) Automated quantification of mitral inflow and aortic outflow stroke volumes by three-dimensional real-time volume color-flow Doppler transthoracic echocardiography: comparison with pulsed-wave Doppler and cardiac magnetic resonance imaging. J Am Soc Echocardiogr 25(1):56–65. doi:10.1016/j.echo.2011.10.004. Epub 2011 Nov 21

Theroux P (2011) Acute coronary syndromes. Companion to Braunwald's heart disease, 2nd edn. Elsevier-Saunders, Philadelphia

Theroux P, Franklin D, Ross J Jr, Lemper WS (1974) Regional myocardial function during acute coronary artery occlusion and its modification by pharmacological agent in the dog. Circ Res 34:896–908

Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, van't Veer M, Klauss V, Manoharan G, Engstrom T, Oldroyd KG, Ver Lee PN, MacCarthy PA, Fearon WF, Investigators FS (2009) Fractional flow reserve vs. angiography for guiding percutaneous coronary intervention. N Engl J Med 360(3):213–224

Varnava AM, Davies MJ (2001) Relation between coronary artery remodelling (compensatory dilatation) and stenosis in human native coronary arteries. Heart 86:207–211

Virmani R, Kolodgie FD, Burke AP, Farb A, Schwartz SM (2000) Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. Arterioscler Thromb Vasc Biol 20:1262–1275

Virmani R, Kolodgie FD, Burke AP, Finn AV, Gold HK, Tulenko TN, Wrenn SP, Narula J (2005) Atherosclerotic plaque progression and vulnerability to rupture: angiogenesis as a source of intraplaque hemorrhage. Arterioscler Thromb Vasc Biol 25:2054–2061

Virmani R, Burke AP, Farb A, Kolodgie FD (2006) Pathology of the vulnerable plaque. J Am Coll Cardiol 47:C13–C18

Voigt JU, Exner B, Schmiedehauser K et al (2003a) Strain rate imaging during dobutamine stress echocardiography provides objective of inducible ischemia. Circulation 107:2120–2126

Voigt JU, Lindenmeier G, Exner B et al (2003b) Incidence an characteristics of segmental postsystolic longitudinal shortening in normal, acutely ischemic and scarred myocardium. J Am Soc Echocardiogr 16:415–423

Voigt JU, Scheider TM, Korder S et al (2009) Apical transverse motion as surrogate parameter to determine left ventricular function inhomogeneities: a new integrative approach to left ventricular asynchrony assessment. Eur Heart J 30:959–968

Wang S, Fleischmann EK (2001) The role of echocardiographic evaluation in patients with acute chest pain in the emergency room. In: Otto C (ed) The practice of clinical echocardiography. WB Saunders, Philadelphia

Warnes CA (2009) Coarctation of the aorta and aortic disease. In: Michelle Z (ed) Gurvitz. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1405178205.html

WHO (2011) Global atlas on cardiovascular disease prevention and control. World Health Organization, Geneva

Wier WG, Pappano AJ (2013) Cardiovascular physiology, 10th edn, p 292. https://elsevier.ca/product.jsp?isbn=9780323086974

Willerson JT, Holmes DR (2015) Coronary artery disease. Springer, London, pp 3–4

Willerson JT, Wellens HJJ, Cohn JN, Holmes DR (eds) (2007) Cardiovascular medicine, 3rd edn. J Chem Inf Modeling, vol 53. Springer. http://www.springer.com/la/book/9781846281884

Windecker S, Kolh P, Alfonso F, Collet JP, Cremer J, Falk V, Filippatos G, Hamm C, Head SJ, Juni P, Kappetein AP, Kastrati A, Knuuti J, Landmesser U, Laufer G, Neumann FJ, Richter DJ, Schauerte P, Sousa Uva M, Stefanini GG, Taggart DP, Torracca L, Valgimigli M, Wijns W, Witkowski A, Authors/Task Force members (2014) 2014 ESC/EACTS guidelines on myocardial revascularization: the task force on myocardial revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS) developed with the special contribution of the European Association of Percutaneous Cardiovascular Interventions (EAPCI). Eur Heart J 35(37):2541–2619

Wyman RM, Safian RD, Portway V, Skillman JJ, McKay RG, Baim DS (1988) Current complications of diagnostic and therapeutic cardiac catheterization. J Am Coll Cardiol 12(6):1400–1406

Patient-Specific Modeling of the Coronary Circulation

3

Tiziano Passerini, Lucian Itu, and Puneet Sharma

Abstract

The instantaneous wave-Free Ratio (IFR) has been recently validated as a rest state pressure-derived index of coronary stenosis severity. We demonstrate that IFR and other patient-specific features of rest-state coronary hemodynamics can be quantified via a novel approach based on reduced-order computational modeling of blood flow. Blood flow is computed in image-based anatomical reconstructions of the coronary tree from Coronary Angiography (CA). A fully automatic two step parameter estimation framework ensures that the computations match the available patient-specific measurements. We evaluate a hybrid decision making strategy c-IFR–invasive FFR against an FFR-only strategy using a dataset comprising 125 lesions (64 patients). Lesions were considered functionally significant if c-IFR < 0.86, non-significant if c-IFR > 0.93, while lesions with intermediate c-IFR were classified based on FFR. Of the 125 lesions, 43 were hemodynamically significant (FFR ≤ 0.8). The hybrid c-IFR–FFR strategy resulted in a diagnostic accuracy of 96% when compared to the FFR-only strategy, while requiring invasive FFR assessment in only 34% (43) of the lesions.

T. Passerini (✉) • P. Sharma
Medical Imaging Technologies, Siemens Healthcare, 755 College Road,
Princeton, NJ 08540, USA
e-mail: tiziano.passerini@siemens.com

L. Itu
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

Automation and Information Technology, Transilvania University of Brasov,
Mihai Viteazu nr. 5, Brasov 5000174, Romania

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_3

3.1 Introduction

Previous investigations have shown that revascularization of blocked coronary arteries is preferred for severe lesions, whereas mild to moderate lesions are best treated using medical therapy alone (De Bruyne et al. 2012). The decision to revascularize blocked coronaries is commonly performed considering anatomical markers extracted from invasive coronary angiography, such as the percentage reduction in lumen diameter. Invasive coronary angiography is the gold standard in CAD imaging (Levin 1982; Ryan 2002). Subjective assessment of angiographically apparent CAD is inadequate due to high degrees of intra-observer and inter-observer variability. Hence, the significance of coronary stenoses is routinely assessed by computer-assisted quantitative coronary angiography (Ng and Lansky 2011).

There is strong evidence that this approach has a limited accuracy in evaluating the hemodynamic significance of lesions (Toth et al. 2014). Due to the tremendous improvement in medical imaging technologies, non-invasive imaging plays an increasingly important role in the diagnosis of CAD. Coronary Computed Tomography Angiography (CCTA) is a non-invasive imaging modality which is being increasingly used for the visualization and diagnosis of CAD, prior to invasive catheterization. While CCTA-based measurements of lesion anatomy correlate well with those from angiography, many severe lesions indicated by CCTA do not cause ischemia, leading to patients being unnecessarily referred to coronary angiography for invasive evaluation (Hachamovitch and Di Carli 2008; Meijboom et al. 2008). In view of the limitations of the pure anatomical evaluation of CAD, the functional index of Fractional Flow Reserve (FFR) has been recently introduced as an alternative. FFR is defined as the ratio of flow in the stenosed branch at hyperemia—a condition of stress, with maximum coronary blood flow—to the hypothetical hyperemic flow in the same branch under healthy conditions. This can be shown to be closely approximated by the ratio of hyperemic cycle-averaged pressure distal to the stenosis to the cycle-averaged aortic pressure (Pijls et al. 1996). Following multiple successful clinical trials which showed the superiority of FFR-guided decision making (Tonino et al. 2009), FFR is currently the gold standard for determining the functional severity of a lesion (Fihn et al. 2012; Windecker et al. 2014). Clinical evaluation of FFR is done under angiographic guidance, using a catheter-based pressure transducer. However, despite the advantages offered by FFR, studies have indicated that the use of FFR is still relatively uncommon due to additional costs, the need to administer drugs to induce hyperemia, and the invasive nature and the extended duration of the procedure, particularly during assessment of multivessel disease (Dattilo et al. 2012).

To address this issue, other hemodynamic indices have recently been introduced, which are based on measurements acquired during resting condition. These include rest state cycle average distal pressure divided by cycle average aortic pressure (basal Pd/Pa), basal stenosis resistance (BSR), and instantaneous wave-free ratio (IFR) (Sen et al. 2013). IFR is defined as the ratio of average distal pressure to the average aortic pressure measured during the diastolic wave-free period—a period of the heart cycle when resistance is constant and minimized (Sen et al. 2012). Several recent studies have reported the diagnostic accuracy of IFR in the range of 80–90% when compared against FFR (Petraco et al. 2013a, b; Jeremias et al. 2014; Escaned et al. 2015). When compared against an independent measure of ischemia (such as coronary flow reserve), however, IFR and FFR showed similar performance (Petraco et al. 2014a, b). To alleviate concerns around the moderate diagnostic accuracy of IFR, a hybrid IFR–FFR strategy was proposed (Petraco et al. 2013a, b, 2014a, b).

Analogous to FFR, the utility of IFR was also demonstrated for the assessment of serial lesions and of post-PCI hemodynamics (Nijjer et al. 2013, 2014).Two parallel large outcome-based clinical trials, with more than 4000 patients in total, are currently being conducted to investigate whether IFR is non-inferior to FFR based on a single cut-off point of IFR ≤ 0.89 (Götberg et al. 2015; Jeremias 2015).

Blood-flow computations, performed using computational fluid dynamics (CFD) algorithms, when used in conjunction with patient-specific anatomical models extracted from medical images, have been proposed for diagnosis, risk stratification, and surgical planning (Taylor and Steinman 2010). Model-based assessment of coronary stenoses has been previously performed using such techniques in several clinical studies, based on anatomical models reconstructed from Coronary Computed Tomography Angiography (CCTA) (Koo et al. 2011; Renker et al. 2014; Coenen et al. 2015) or X-ray coronary Angiography (XA) (Morris et al. 2013; Tu et al. 2014; Papafaklis et al. 2014; Tröbs et al. 2016). Computed FFR has been the main quantity of interest in these studies, all of which showed that computed FFR has good diagnostic accuracy compared to invasively measured FFR. Typically, such models are based on the simulation of the hyperemic condition, assuming that in this condition the blood-flow rate increases predictably when compared to the rest state.

In this chapter we introduce a method for computing IFR (c-IFR) from patient-specific computational blood flow modeling based on routine coronary angiography. The advantage compared to the above mentioned computational modeling approaches is that no assumption is required on the effect of hyperemia on the flow. In a retrospective analysis, we evaluate a novel hybrid strategy for the functional assessment of CAD based on c-IFR and invasive FFR measurements, and we evaluate it by comparison against an FFR-only strategy.

The concepts and information presented in this chapter are based on research and are not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

3.2 Instantaneous Wave-Free Ratio

In current practice, iFR is computed from invasively measured pressures obtained during the wave-free period (Sen et al. 2012). Instantaneous coronary resistance, computed as ratio of instantaneous pressure to instantaneous flow rate, varies during the heart cycle due to the interaction between the microvasculature and the myocardium, with higher fluctuations during systole. During the wave-free period however, coronary resistance is naturally constant and minimized as can be observed in Fig. 3.1 (Sen et al. 2012).

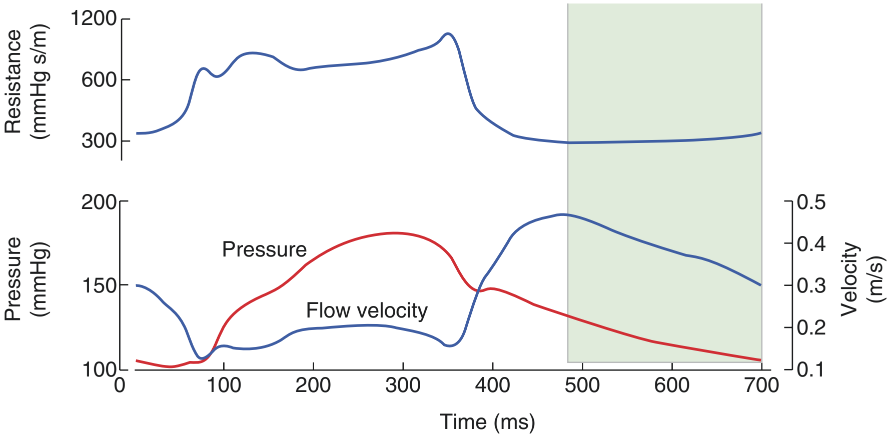


Fig. 3.1 Instantaneous pressure, velocity and resistance values during one heart cycle at rest (Sen et al. 2012)

Hence, to measure iFR the wave-free period has to be determined. In the arterial circulation waves can originate from both upstream locations (left ventricle)—forward travelling waves—or from downstream locations (bifurcations or the micro-circulation)—backward travelling waves. To establish the origin of the waves both pressure and flow rate profiles are required (Davies et al. 2006).

In the coronary circulation forward and backward travelling waves can occur at the same time. To separate the waves, a wave intensity analysis is performed (Parker et al. 1988).

Six different waves can be identified in the coronary circulation, which are displayed in Fig. 3.2 (Davies et al. 2006):

1. Early backward-travelling pushing wave: at the onset of ventricular contraction the myocardium compresses the coronary microcirculation.
2. Forward travelling pushing wave: contraction of the ventricle lumen generates ventricular ejection.
3. Late backward-travelling pushing wave: appears as a result of continued compression of the micro-circulation and wave reflection of the forward travelling pushing wave.
4. Forward travelling suction wave: the slowing of ventricular contraction creates a suction effect at the proximal end of the arteries.
5. Dominant backward travelling suction wave: appears as a result of the relief of myocardial compression of the coronary microcirculation.
6. Late forward-travelling pushing wave: is associated with the closure of the aortic valve.

The diastolic wave free-period runs from $WI-_{[\text{diastole}]} = 0$–5 ms before the end of diastole (end of diastole is defined as the moment when wave 1 from the above list is detected). $WI-_{[\text{diastole}]}$ is the wave intensity of distally originating waves.

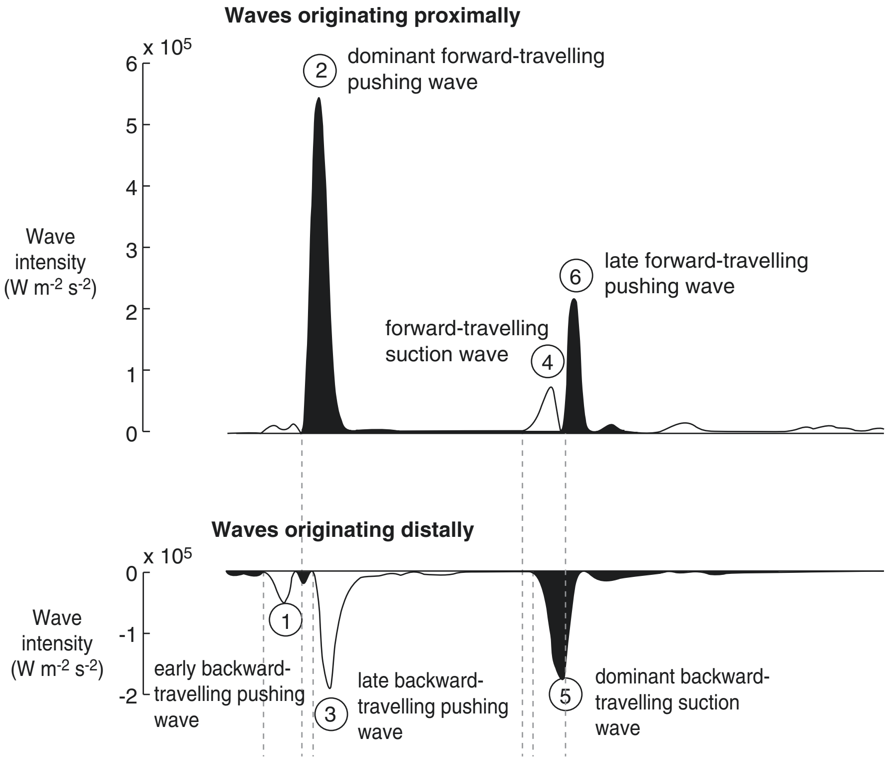


Fig. 3.2 Identification of waves in the human left coronary tree (Davies et al. 2006)

In clinical practice, only time-varying pressure is measured and hence it is not possible to perform wave separation. Consequently the wave free-period considered for the invasive measurement of iFR (Sen et al. 2012) runs from 25% of the way into diastole (the onset of diastole is identified from the dicrotic notch) to 5 ms before the end of diastole.

Depending on the availability of pressure and/or velocity profiles, other approaches for identifying the wave free period can be used (Davies et al. 2012):

- wave free period runs from 250 ms after dU_{max} is obtained, and lasts for 150 ms;
- wave free period runs from 150 ms after P_{max} is obtained until the end of the heart cycle minus 50 ms;
- wave free period is the period, after peak pressure, during which the standard deviation of the forward travelling wave is in the lowest 5% (or in the lowest 10% if no such period exists);
- wave free period is the mid window between the peak pressure time point and the end of the heart cycle (the 3/5 between these two time points);
- wave free period is the window during which dU is less than 10% (or 5%) of dU_{max}.

To separate the waves, a wave intensity analysis is performed. First, proximally and distally originating components of the change in pressure are determined:

$$dP_+ = \frac{1}{2}\left(dP + \rho v_w dU\right), \quad dP_- = \frac{1}{2}\left(dP - \rho v_w dU\right), \tag{3.1}$$

where dP and dU are the change in pressure and velocity, ρ is the blood density and v_w is the wave speed. Next, the wave intensity of proximally (WI_+) and distally originating waves (WI_-) is quantified:

$$WI_+ = \frac{1}{4\rho c}\left(\frac{dP}{dt} + \rho c \frac{dU}{dt}\right)^2, \quad WI_- = \frac{1}{4\rho c}\left(\frac{dP}{dt} - \rho c \frac{dU}{dt}\right)^2. \tag{3.2}$$

The diastolic wave free-period runs from the time when $WI_{-[\text{diastole}]} = 0$–5 ms before the end of diastole.

3.3 Methods

3.3.1 Rest State Flow Rate Estimation

The estimation of the rest state flow rate for a given geometry is one of the most important steps in the personalization of the CFD model. In the following we introduce an algorithm for computing the total rest flow for a patient-specific coronary tree composed of four main steps.

During the first step the reference radius and the rest flow rate for each branch are estimated.

Previous studies have established a power law relationship between flow rate and vessel reference radius in the arterial circulation (Murray 1926a, b) and in particular for the coronary circulation (Kassab et al. 1995):

$$q = k \cdot r_{ref}^n, \tag{3.3}$$

where q is the average flow rate in the vessel, r_{ref} is the reference radius of the vessel, k is a proportionality constant and n is the power coefficient, which takes values between 2 (for large arteries (Zamir et al. 1992)) and 3 (for small arteries) (Hutchins et al. 1976; Kamiya and Togawa 1980; Uylings 1977; van der Giessen et al. 2011; Kassab et al. 1995). Hence:

$$q = f_1\left(r_{ref}, n\right) \tag{3.4}$$

For a patient-specific geometry, the radius of a branch is continuously changing along its centerline. Furthermore, the branch may contain mild to severe stenoses, which in turn could be focal, diffuse, etc. Thus to compute the reference radius of a branch, an operator is applied to the longitudinally varying radius:

$$r_{ref} = f_2\left(r\left(x\right)\right) \tag{3.5}$$

This operator could perform one of the following operations (not limited to):

- average value of healthy radiuses of the entire branch or a part of the branch;
- average value of healthy radiuses obtained when excluding the largest x% and the smallest y% of the radius values of the entire branch or a part of the branch;
- maximum or minimum value of healthy radiuses of the entire branch or a part of the branch.

In a patient-specific geometry, if a branch is very short, or it displays diffuse disease along its entire length, the reference radius can be computed as a function of the reference radius values of the parent and/or daughter branches:

$$r_{ref} = f_3\left(r_{ref}^{par}, r_{ref}^{d1}, r_{ref}^{d2}, \ldots\right) \tag{3.6}$$

During the second step the total rest flow rate of a coronary tree is estimated based on the branches with a certain generation number.

Figure 3.3 displays a coronary tree where a generation number has been attached to each branch. The root branch has a generation number of 0, which then increases at each bifurcation by one. In the following we use the branches with generation number g. Before estimating the flow rate, a weight is attached to each branch, representing a confidence value for the correctness of the computed reference radius: wi. Very short branches (like the bottom branch with generation number equal to 1 in Fig. 3.3) or entirely diseased branches (like the diffusely diseased branch with generation number equal to 2 in Fig. 3.3) will receive a low confidence value, while long vessels without radius irregularities will receive large confidence values. Next, a total flow rate value for the entire coronary tree is estimated, based on the branches from generation g:

$$\left(q_{total}\right)_g = f_4\left(w_i, q_i\right), \tag{3.7}$$

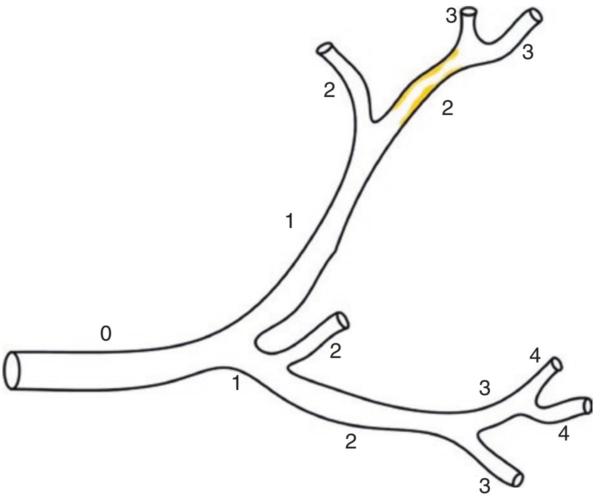


Fig. 3.3 Sample coronary tree with a generation number attached to each branch

where index i refers to all branches from generation g and all terminal branches with a generation number smaller than g. For example, the total flow rate could be computed as:

$$\left(q_{total}\right)_g = \frac{\sum_i w_i \cdot q_i}{\sum_i w_i}. \tag{3.8}$$

During the third step the total rest flow rate of a coronary tree is estimated based on all branches with generation number between gmin and gmax.

To improve the accuracy of the flow rate estimation different branch generations are used for computing one final total flow rate value. In the following, generation levels g_{min} to g_{max} are considered, whereas usually g_{min} is zero (it could be also larger than zero if the root branch is very short). Before estimating the total flow rate, a weight is attached to each generation number, representing a confidence value for the correctness of the computed total flow rate from the branches with the corresponding generation number: v_i. In a patient-specific setup, in general, low generation numbers will receive a large weight, while large generation number will receive low weights (typically small side branches are missed towards the large generation number, which would lead to an underestimation of the flow rate values). Hence the final flow rate value is estimated from:

$$q_{total} = f_5\left(v_j, \left(q_{total}\right)_j\right), \tag{3.9}$$

where index j refers to a generation between g_{min} and g_{max}. For example, the total flow rate could be computed as:

$$q_{total} = \frac{\sum_j v_j \cdot \left(q_{total}\right)_j}{\sum_j v_j}. \tag{3.10}$$

One advantage of this method is that you don't need both LCA and RCA to estimate the flow rates and boundary conditions for one of the coronary trees.

During the fourth and final step, starting from the total coronary flow determined in Eq. (3.10), the flow rate of each branch is determined. Hence, first we compute the flow rate of all terminal branches:

$$q_k = \frac{\left(r_{ref}\right)_k^n}{\sum_k \left(r_{ref}\right)_k^n} q_{total}, \tag{3.11}$$

where k refers to the coronary leaf segments. Finally, the ischemia weights of the branch and root segments are computed as a sum of the ischemia weights of all downstream leaf segments:

$$q_l = \sum_k q_k, \tag{3.12}$$

where k refers to all leaf segments lying downstream from the current segment l.

3.3.2 Computational Blood Flow Modeling

In the present study, we use a 1D-0D reduced-order geometrical multiscale model (Fig. 3.4) (Itu et al. 2012). Stenoses are automatically detected, and a semi-analytical pressure drop model (Huo et al. 2012), parameterized by the geometry, is employed. The patient-specific coronary geometry is coupled to a population-average aortic geometry. A time-varying elastance heart model is used for providing the inlet boundary condition of the aorta, and a specialized lumped model of coronary microvasculature is employed (Itu et al. 2012). This model simulates the effect of the

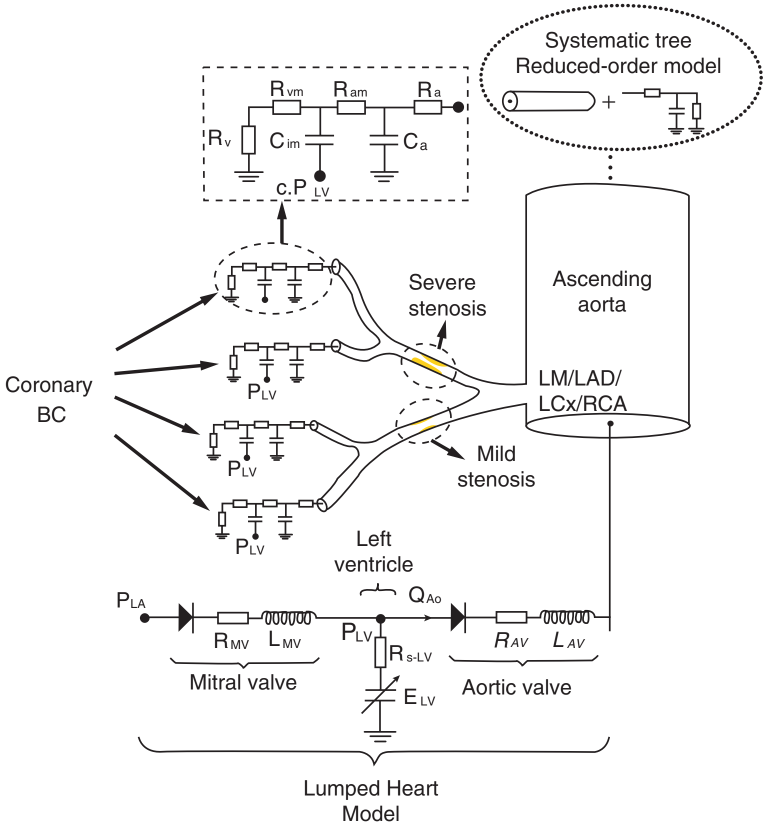


Fig. 3.4 Model of the systemic and coronary arterial circulation

myocardial contraction on the circulation by incorporating the intramyocardial pressure, approximated as the left ventricular pressure multiplied by a proportionality constant *c*.

To compute patient-specific hemodynamics, the parameters of the model are personalized through a parameter estimation framework consisting of two sequential steps. First, a series of parameters are computed directly, and, next, a fully automatic optimization-based calibration method is employed to estimate the values of the remaining parameters, ensuring that the personalized computations match the measurements performed on the patient. For the flow rate estimation we use the following simplified version of the algorithm described in the previous section. Hence, first we estimate the rest flow rate of each vessel in the coronary geometry as follows:

$$\left(Q_r\right)_i = k \cdot r_i^n, \tag{3.13}$$

where r_i is the radius of the healthy part of the vessel, k is a proportionality constant, and n is a power coefficient (Kassab et al. 1995). To estimate the total flow rate which enters at rest the coronary tree, first the average flow rate is computed for each generation of vessels. The final flow rate value Q_r is computed by averaging the estimations obtained for each generation. Finally, we determine the flow rate for each terminal vessel by distributing the previously computed total flow rate:

$$Q_j = Q_r \cdot r_j^3 \,/ \sum_{j=1}^{m} r_j^3, \tag{3.14}$$

where j refers to the current terminal branch and m is the total number of terminal branches. Next, the rest mean arterial pressure (MAP_r) is computed from cuff-based systolic (*SBP)* and diastolic blood pressures (*DBP*), and the heart rate (*HR*):

$$MAP_r = DBP + \left[\frac{1}{3} + \left(HR \cdot 0.0012\right)\right] \cdot \left(SBP - DBP\right). \tag{3.15}$$

We consider a venous pressure P_v, of 5 mmHg, and determine the total rest microvascular resistance for each terminal branch j:

$$\left(R_{t-r}\right)_j = \frac{MAP_r - P_v}{\left(Q_r\right)_j}. \tag{3.16}$$

To compute the total resistance of the systemic circulation we assume that coronary flow represents around 4.5% of the total cardiac output at rest. This flow is then distributed approximately evenly between the three main coronary arteries (LAD, LCx and RCA) (Wieneke et al. 2005), and the total systemic flow rate is estimated as:

$$Q_{syst} = \frac{100 - x}{x} Q_r, \tag{3.17}$$

where x is the percentage of the total cardiac output directed towards the patient-specific coronary model (1.5% for LAD, LCx and RCA, and 3% for Left Main), and Q_r is the rest flow rate computed for the root branch of the coronary tree. Hence, the total resistance of the arterial system is computed as:

$$R_{t-syst} = \frac{MAP_r - P_v}{Q_{syst}}. \tag{3.18}$$

During the second step of the parameter estimation framework, an optimization-based calibration method is employed to estimate the values of the remaining parameters. The following parameters are estimated: V_{ED} (the end diastolic volume of the heart), $\rho = R_p/R_t$ (the ratio of proximal to total resistance for the systemic windkessel model) and $\tau = R_d \cdot C$ (the time constant of the systemic windkessel's exponential pressure). To model coronary autoregulation, i.e. to ensure that the computed coronary flow rate values match the ones estimated in Eq. (3.14), the total microvascular resistance at each coronary outlet, $(R_{t-r})_j$, is also determined through the parameter estimation framework. The values of $(R_{t-r})_j$ computed in Eq. (3.16) are used as initial values: the final values are lower due to the resistance introduced by the stenoses.

Furthermore, iFR is computed during the wave-free window, and consequently it depends on the trans-stenotic pressure drop during diastole. The pressure drop in turn depends on the flow rate, and hence it is important to control the amount of coronary flow at systole and at diastole. Previous studies have examined the amount of coronary flow at systole and diastole (Spiller et al. 1983; Heller et al. 1994), indicating that systolic flow is proportionally lower in the LCA compared to the RCA. In the following we assume that systolic flow represents 20% of the total coronary flow for LCA branches, and 31% of the total coronary flow for RCA branches. To ensure that the computed percentage of systolic flow matches these values, the proportionality constant c is adapted separately for each outlet.

The calibration method is formulated as the solution of a system of nonlinear equations with a root where the computed objective values and the reference values match (Itu et al. 2015):

$$\mathbf{f}\begin{pmatrix} V_{ED} \\ \rho \\ \tau \\ (R_{t-r})_1 \\ c_1 \\ \dots \\ (R_{t-r})_n \\ c_n \end{pmatrix} = \begin{Bmatrix} (P_{avg})_{comp} - (P_{avg})_{ref} \\ (P_{\max})_{comp} - (P_{\max})_{ref} \\ (P_{\min})_{comp} - (P_{\min})_{ref} \\ \left[(Q_r)_1\right]_{comp} - \left[(Q_r)_1\right]_{ref} \\ \left[(\%SystFlow)_1\right]_{comp} - \left[(\%SystFlow)_1\right]_{ref} \\ \dots \\ \left[(Q_r)_n\right]_{comp} - \left[(Q_r)_n\right]_{ref} \\ \left[(\%SystFlow)_n\right]_{comp} - \left[(\%SystFlow)_n\right]_{ref} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{Bmatrix}, \tag{3.19}$$

where $(\cdot)comp$ refers to a value computed using the multiscale model, and $(\cdot)ref$ refers to the reference value (the patient-specific value). To provide a good initial estimate of the parameters in Eq. (3.19) and to limit the number of iterations required with the model displayed in Fig. 3.3, the nonlinear system of equations is first solved for a lumped parameter model, for which the aorta and the large coronary arteries are modeled as resistances.

3.3.3 Clinical Data

The above described methods were evaluated using a patient cohort consisting of 67 patients. All patients had an intermediate or high likelihood of obstructed coronary artery disease, were considered for PCI, were stable and at least 18 years old. Patients were excluded if they were unable to provide informed consent, had significant arrhythmia (heart rate over 120 bpm), atrial fibrillation, low systolic pressure (below 90 mmHg), contraindication to beta blockers, nitroglycerin or adenosine, a non-cardiac illness with a life expectancy of less than 2 years, pathological aortic valve, rest state angina or myocardial infarct during the last 6 months.

Coronary angiography (Siemens Artis Zee, Forchheim, Germany) was performed after iso-centering in posterior-anterior and lateral planes, via a transradial (preferred) or transfemoral approach. In all cases a six French diagnostic catheter was used after intracoronary injection of glyceryl trinitrate according to routine practice in our hospital, with manual contrast injection and cine acquisition at a frame rate of 15 frames/s. At least four acquisitions in different angulations were obtained for the left coronary artery and at least two acquisitions were obtained for the right coronary artery. The patient and the table were not moved between acquisitions, and overlapping and foreshortening were minimized for the target vessel segment.

Acquisition of physiological data for FFR calculation was performed according to conventional practice (Kern et al. 2006) with a commercially available FFR measurement system (PressureWire Aeris; St. Jude Medical, Minneapolis, MN, USA). The 0.014″ coronary wire with a pressure tip was advanced until the pressure sensor passed the orifice of the guiding catheter. Transcatheter aortic and intracoronary pressure tracings were equalized. Subsequently, the guidewire was advanced into the respective coronary artery until the pressure sensor passed the index lesion. Hyperemia was induced by administration of adenosine either intravenously at a constant rate of 140 μg/kg/min, or as an intracoronary bolus; the pressure recording was started, and FFR was determined.

A total of 132 coronary lesions in 67 patients underwent FFR analysis. After excluding 4 lesions in 4 patients due to lack of adequate visualization in invasive angiography, and 3 lesions in 2 patients due to incomplete pressure recordings, 125 coronary lesions in 64 patients remained that had at least two views separated by at least 30°, with sufficient contrast and without superposition of other branches. These lesions were included in the subsequent analyses.

Routine coronary angiograms from two views at least 30° apart were selected and two corresponding end-diastolic frames were chosen. On each frame, the centerline of the vessel(s) was traced manually by selecting at least three points. Next, the lumen boundary was automatically determined and manually corrected if deemed required. Subsequently, a 3D anatomic model was automatically generated based on the centerlines and the segmented lumen boundaries.

3.3.4 Evaluation of a Hybrid c-IFR–FFR strategy

This study evaluated retrospectively a hybrid strategy for the functional assessment of CAD, based on computed IFR (c-IFR) derived from image based computational blood flow modelling, and invasive FFR measurements. The same cut-off values were used as in the hybrid purely invasive IFR-FFR decision making strategy, which was recently proposed for increasing the adoption of functional assessment of coronary artery stenoses (Petraco et al. 2013a, b). Hence, if the c-IFR value was larger than 0.93 the lesion was considered functionally non-significant, if c-IFR was smaller than 0.86 the lesion was considered functionally significant, and lesions with intermediate c-IFR values were classified based on the invasively measured FFR value (Fig. 3.5). A stenosis was considered functionally significant if FFR ≤ 0.8.

The percentage/number of stenoses which were outside of the intermediate c-IFR window were determined, and the percentage/number of patients for which none of the stenoses were inside of the intermediate c-IFR window were determined (patients for which the hybrid c-IFR/FFR strategy could potentially obviate the need for invasive pressure measurements). Moreover, the effect of the hybrid cut-off values (the size of the hybrid window) on the diagnostic parameters and on the number of lesions no longer requiring invasive measurements was analyzed.

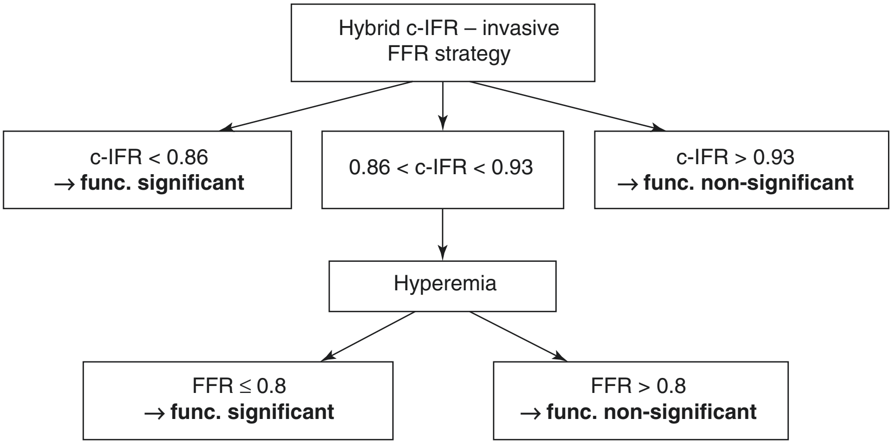


Fig. 3.5 Hybrid decision making strategy

3.4 Results

Baseline patient and lesion characteristics are summarized in Tables 3.1 and 3.2: 64 patients (41 male, 23 female) with 125 lesions were included in this study. Mean patient age was 63.3 ± 8.7 years and BMI was 29.3 ± 3.4 kg/m^2. Mean FFR was 0.81 ± 0.14; 66% of stenoses had FFR between 0.60 and 0.90, and 34% were hemodynamically significant according to the criterion FFR $\leq$ 0.80.

Using a c-IFR value of >0.93 for indicating functionally non-significant CAD, a c-IFR value of <0.86 for indicating functionally significant CAD, and a classification based on invasive FFR for lesions with intermediate c-IFR values, resulted in an accuracy of 96.0%, sensitivity of 97.7%, specificity of 95.1%, positive predictive value of 91.3%, and negative predictive value of 98.7% when compared to an FFR-only strategy. Overall, 82 lesions were classified based solely on c-IFR, 29 had c-IFR < 0.86, and 53 had c-IFR > 0.93, and only 43 lesions were classified based on invasive FFR (Fig. 3.6). The average computation time required for determining model-based IFR for one lesion was 42.2 ± 9.1 s on a workstation equipped with

Table 3.1 Baseline patient characteristics and risk factors (n = 64)

Male	41 (64%)
Female	23 (36%)
Age (years)	63.3 ± 8.7 years
Race	All Caucasian
BMI	29.3 ± 3.4 kg/m^2
Diabetes	29 (45.3%)
Hypertension	56 (87.5%)
Hypercholesterolemia	44 (68.8%)
Smoking history	32 (50.0%)
Family history of CAD	9 (14.1%)
Previous myocardial infarction	11 (17.2%)
Previous PCI	42 (65.6%)
Previous CABG	2 (3.1%)
Ejection fraction	51.7 ± 10.2%

Table 3.2 Baseline lesion characteristics (n = 125)

Index artery	
Left Anterior Descending artery (LAD)	66
Left Circumflex artery (LCx)	30
Right Coronary Artery (RCA)	29
Fractional Flow Reserve	
Mean ± SD	0.81 ± 0.14
Median (IQR)	0.84 (0.77–0.90)
FFR $\leq$ 0.80	43
FFR between 0.60 and 0.90	83

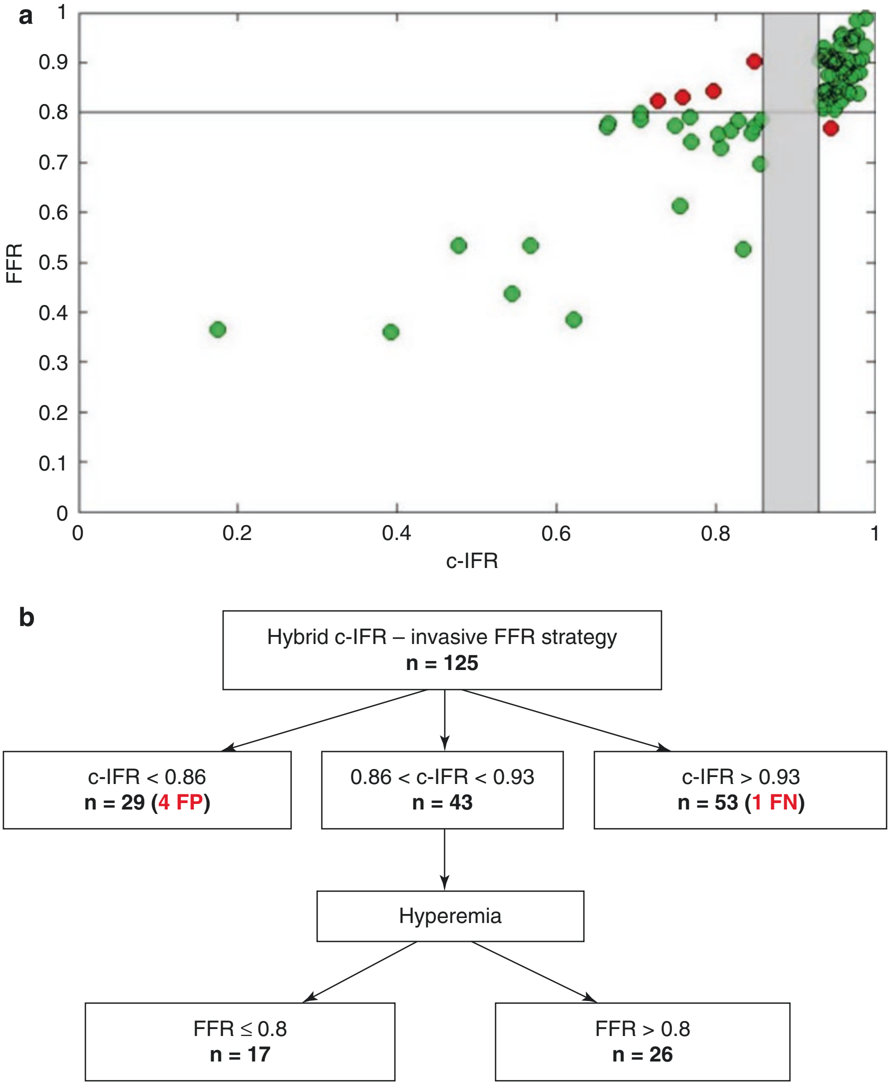


Fig. 3.6 c-IFR–FFR hybrid decision-making strategy: if the c-IFR value was larger than 0.93 the lesion was considered functionally non-significant, if c-IFR was smaller than 0.86 the lesion was considered functionally significant, and lesions with intermediate c-IFR values were classified based on the invasively measured FFR value; (**a**) scatter plot of FFR vs. c-IFR: *green dots* represent the agreement between c-IFR and FFR and *red dots* show disagreement points; (**b**) summary of the predicted results of the hybrid decision-making strategy

3.4 GHz Intel i7 8-core CPU. In Figs. 3.7 and 3.8 representative case examples are shown, illustrating the angiographic images that were used and the corresponding anatomical models colour coded based on the resulting c-IFR values.

The hybrid c-IFR–FFR strategy would significantly reduce the number of lesions and patients for whom invasive measurement is required. Thus, in this study population, 66% of the lesions and 45% of the patients would potentially not require invasive measurements, with functional assessment of CAD maintaining a high diagnostic accuracy with respect to an FFR-only strategy.

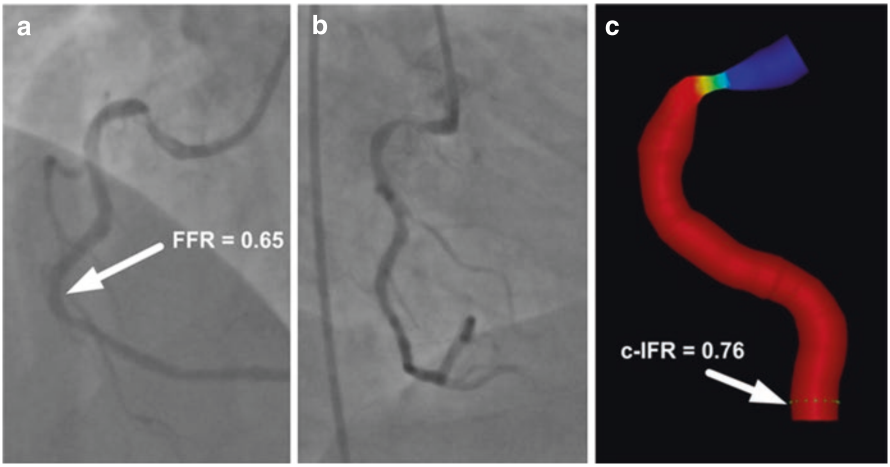


Fig. 3.7 Case example of a lesion in the proximal RCA. Invasively measured FFR was 0.65 and c-IFR was 0.76. (**a**, **b**) Frames from two different views that were used to reconstruct the anatomical model, (**c**) reconstructed anatomical model colour coded by the resulting c-IFR values

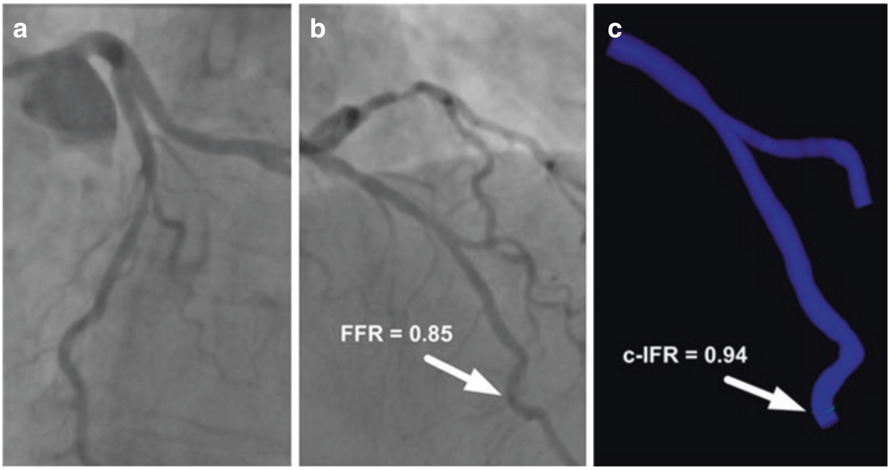


Fig. 3.8 Case example of a lesion in the mid LAD. Invasively measured FFR was 0.85 and c-IFR was 0.94. (**a**, **b**) Frames from two different views that were used to reconstruct the anatomical model, (**c**) reconstructed anatomical model colour coded by the resulting c-IFR values

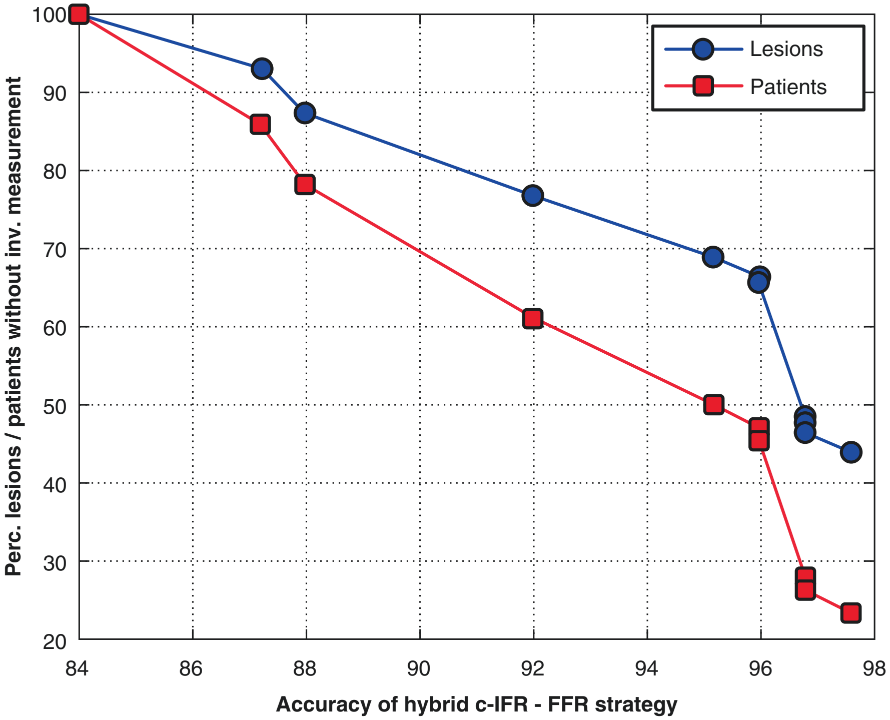


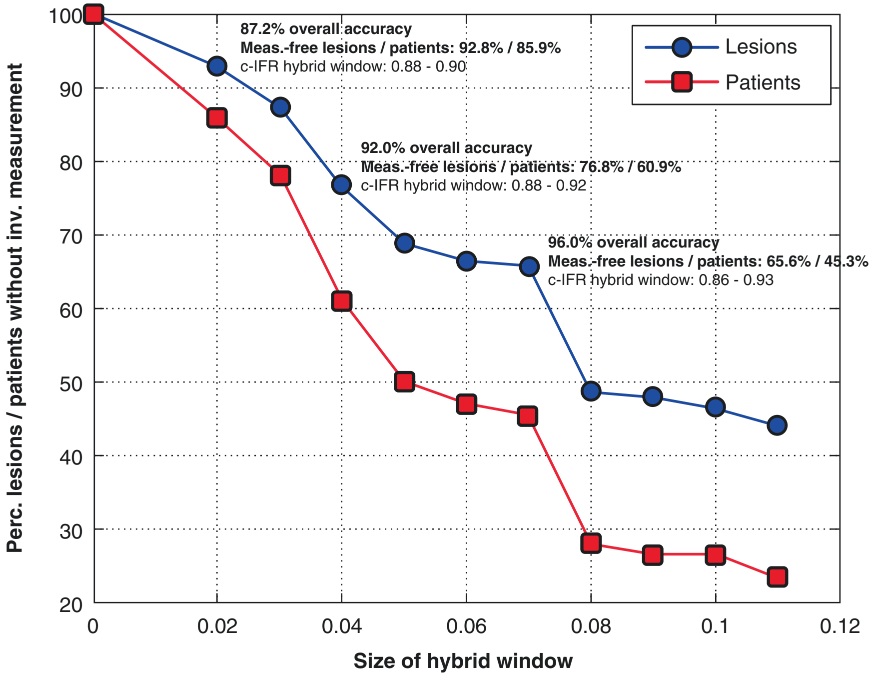
Fig. 3.9 The hybrid c-IFR–FFR strategy reduces the number of lesions and patients requiring invasive pressure measurements for a desired agreement with an FFR-only strategy, e.g. the invasive measurement requirement can be reduced by 77% for a 92% agreement with an FFR-only strategy

By reducing the width of the hybrid window, fewer invasive measurements are required, but the diagnostic accuracy of the hybrid strategy decreases. For a classification agreement between the c-IFR–FFR hybrid strategy and the FFR-only strategy of 87%, 92%, and 96% respectively, the percentage of lesions (patients) no longer requiring invasive measurements would be 92.8% (85.9%), 76.8% (60.9%) and respectively 65.6% (45.3%) (Figs. 3.9 and 3.10). The corresponding cut-off values for the hybrid window would be 0.88 and 0.9, 0.88 and 0.92, and 0.86 and 0.93 respectively. If invasive measurements are performed for a larger window of c-IFR values, the diagnostic accuracy with respect to the FFR-only strategy increases, but at the expense of more lesions requiring an invasive assessment.

3.4.1 Detailed Results for a Subset of Cases

We present in the following detailed results for a subset of three cases:

- Case 1: anatomically severe stenosis;
- Case 2: anatomically intermediate stenosis;
- Case 3: anatomically mild stenosis.



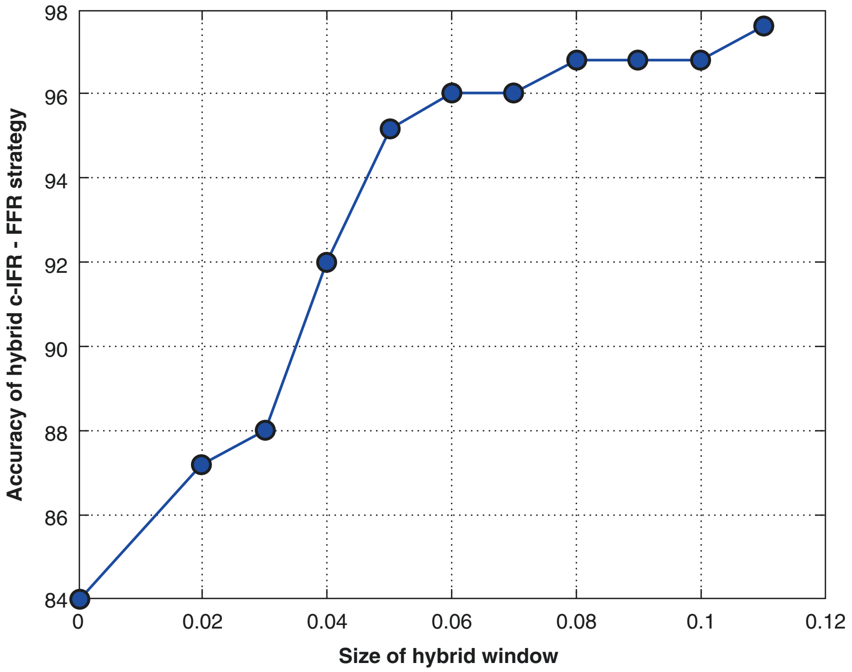


Fig. 3.10 (**a**) Percentage of lesions and patients requiring no invasive measurements according to the hybrid c-IFR–FFR strategy, and (**b**) overall accuracy with respect to the FFR-only strategy, both as a function of the size of the intermediate c-IFR window requiring invasive FFR measurements (for each hybrid window size, the pair of cut-off values leading to the largest accuracy was chosen)

Computations were run until convergence and results are displayed for each case only for the last heart cycle.

Figure 3.11 displays for Case 1 the computed time-varying proximal and distal pressures, time-varying flow rate, time-varying resistance, instantaneous pressure ratio, and wave intensity of forward and backward travelling waves. The diastolic period runs from 6.828 s (dicrotic notch) to 7.5 s. The wave

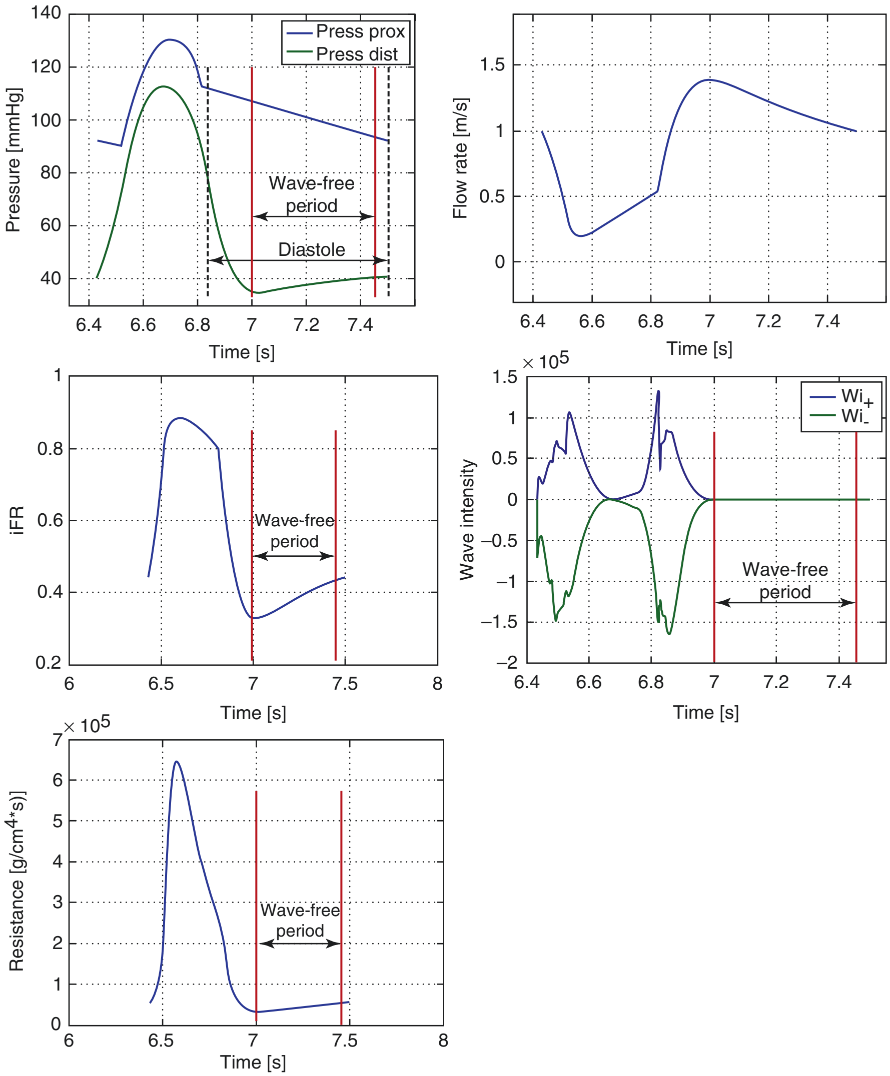


Fig. 3.11 Case 1: Computed time-varying proximal and distal pressures, time-varying flow rate, time-varying resistance, instantaneous pressure ratio, wave intensity of forward and backward travelling waves

Table 3.3 Average proximal and distal computed pressures and pressure ratios (Basal P_d/P_a, iFR) for three cases at rest and hyperemia

Configuration	Quantity	Case 1	Case 2	Case 3
Entire heart cycle—rest	P_a [mmHg]	106.084	100.172	102.803
	P_d [mmHg]	59.609	95.07	100.312
	Basal P_d/P_a	0.562	0.949	0.976
Wave-free period—rest	P_a [mmHg]	100.051	91.895	97.233
	P_d [mmHg]	37.736	85.154	93.499
	iFR	0.377	0.927	0.962

free-period, considered for the computation of iFR, runs from 25% of the way into diastole to 5 ms before the end of diastole (from 6.99 to 7.45 s). The resistance is constant during the wave free-period and also the instantaneous distal to proximal pressure ratio is relatively constant. The wave intensity of forward and backward travelling waves was computed using the method described in (Parker et al. 1988) and when representing the wave-free window on the plot of the two wave intensities one can observe that indeed there are no waves in the interval from 6.99 to 7.45 s.

Table 3.3 displays the average pressures and pressure ratios (Basal P_d/P_a, iFR) for the entire heart cycle and for the wave-free period for the three cases. Both basal P_d/P_a and iFR are determined from the proposed method using imaging data.

The computed iFR value is 0.377 for case 1, which indicates that this is a hemodynamically significant lesion (this finding is confirmed by its anatomical significance).

Figure 3.12 displays for case 2 the computed quantities of interest. The diastolic period runs from 6.65 to 7.24 s. The wave free-period runs from 6.797 to 7.19 s. Again, the resistance is constant during the wave free-period and also the instantaneous distal to proximal pressure ratio is relatively constant. When plotting the wave intensity of forward and backward travelling waves one can observe that there are no waves during the wave-free period.

The computed iFR value is 0.927, which falls into the grey zone identified in previous studies (Petraco et al. 2013a, b) (this finding is confirmed by its intermediate anatomical nature).

Finally, Fig. 3.13 displays for case 3 the computed quantities of interest. The diastolic period runs from 4.235 to 4.565 s. The wave free-period runs from 4.317 to 4.515 s. Once again, the resistance is constant during the wave free-period and also the instantaneous distal to proximal pressure ratio is relatively constant. When plotting the wave intensity of forward and backward travelling waves one can observe that there are no waves during the wave-free period.

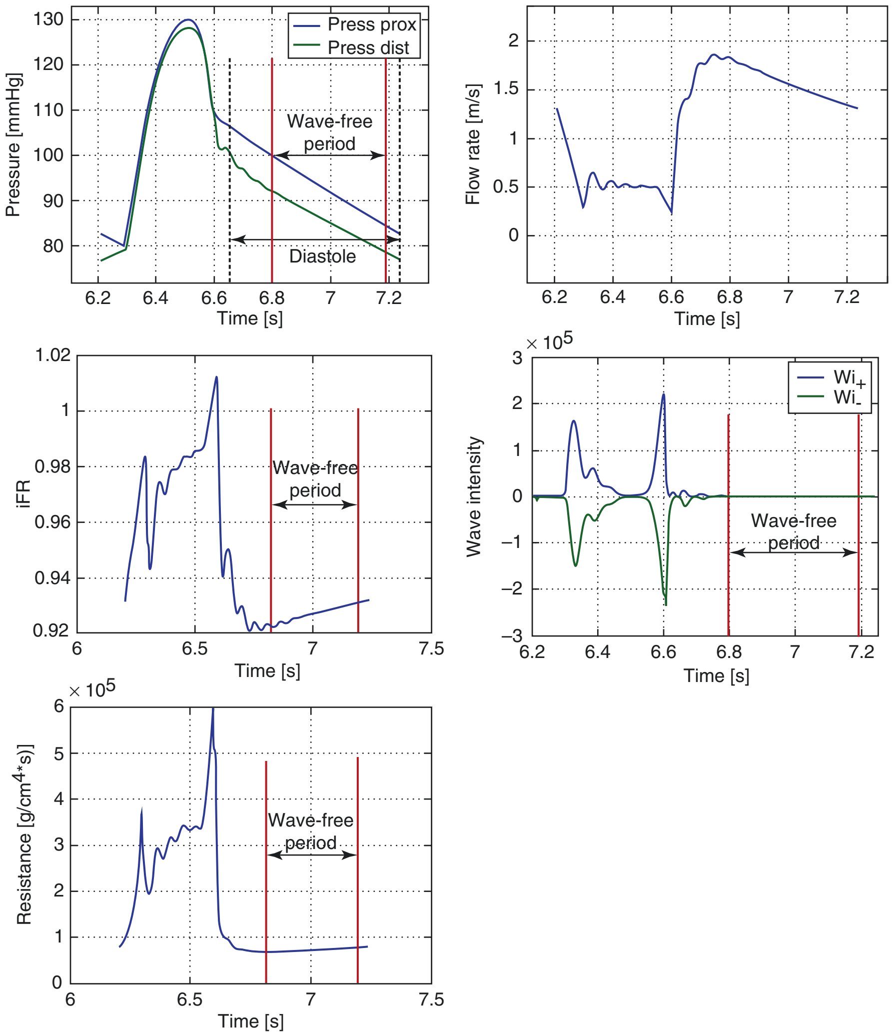


Fig. 3.12 Case 2: Computed time-varying proximal and distal pressures, time-varying flow rate, time-varying resistance, instantaneous pressure ratio, wave intensity of forward and backward travelling waves

The computed iFR value is 0.962 for case 3, which indicates that this is a hemodynamically non-significant lesion (this finding is confirmed by its mild anatomical nature).

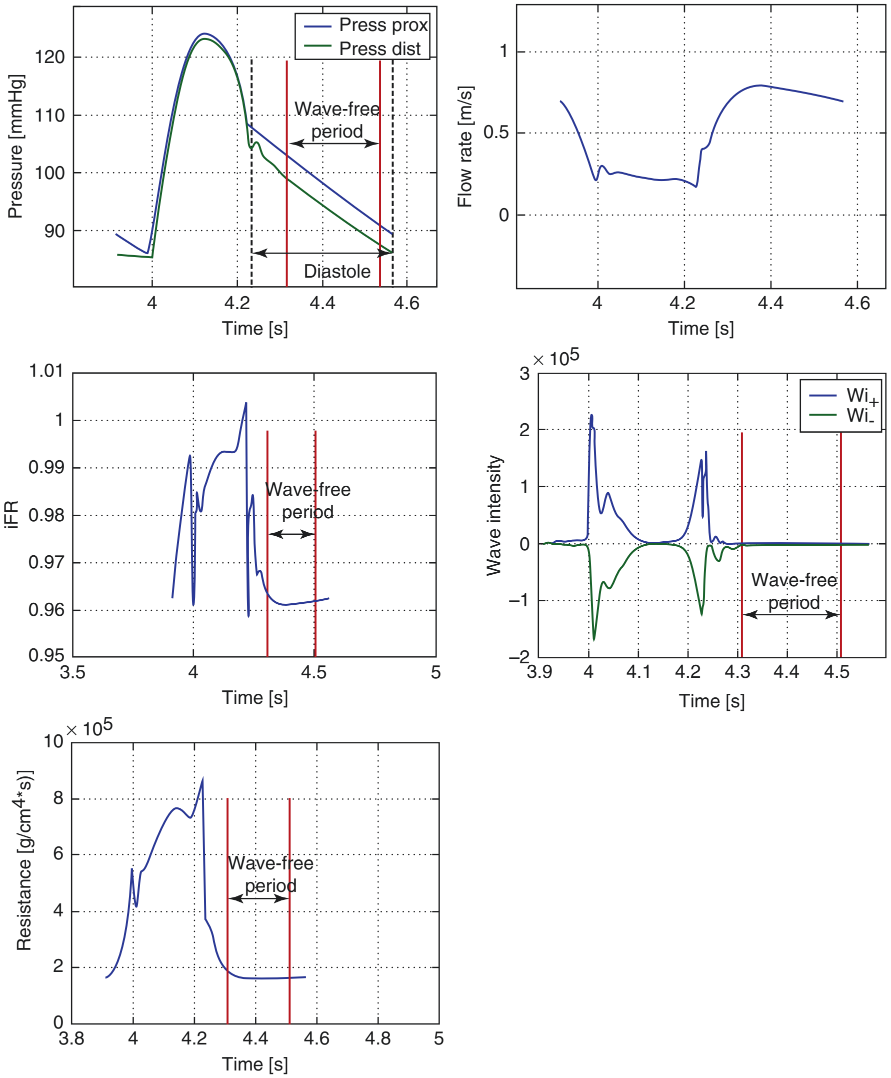


Fig. 3.13 Case 3: Computed time-varying proximal and distal pressures, time-varying flow rate, time-varying resistance, instantaneous pressure ratio, wave intensity of forward and backward travelling waves.

3.5 Discussion and Conclusions

As more data is emerging from studies based on computational modelling, the incremental diagnostic value of computed coronary functional diagnostic indices over the traditional XA based visual or quantitative lesion grading is becoming more evident.

We have introduced a method for image-based computation of Instantaneous Wave-free Ratio (c-IFR) from routine coronary angiography, and have evaluated a novel hybrid strategy for functional assessment of CAD based on c-IFR and invasive FFR measurements. This hybrid strategy has a high diagnostic accuracy when compared to an FFR-only strategy. Similarly to previously published approaches based on computational modelling, our approach provides a non-invasive assessment of a coronary diagnostic index, but, to the best of our knowledge, this is the first published work that addresses computation of IFR (c-IFR) from medical images.

One of the key ingredients for the design of our computational method has been the inclusion of a heart model and a systemic circulation model to provide proper proximal and distal boundary conditions for the coronary circulation. Since c-IFR is determined from the time-varying pressures during the diastolic wave-free period, it is particularly important to correctly model time-varying coronary hemodynamics: low flow during systole and large flow during diastole. The low systolic flow is caused by the ventricular contractions which generate an intramyocardial pressure. With the inclusion of a heart model in the computational modelling framework this aspect is properly taken into account for the determination of c-IFR.

Another key aspect of our method is the fact that no assumption about the effect of hyperemia on the flow is required, which has been one of the limitations of CFD-based studies reported in the past. Therein, modelling was based on a normal microvascular function, e.g. four- to fivefold decrease of microvascular resistance during hyperaemia. The microvascular function is however likely to be hampered in patients with diabetes, hypertension, and/or severe lesions, which in turn is difficult to assess non-invasively.

Despite the overwhelming clinical evidence that a FFR-guided revascularization strategy improves outcome, still, less than 10% of the coronary interventions are preceded by FFR measurements due to the limitations of invasive pressure measurements mentioned above (Dattilo et al. 2012). Hence a hybrid c-IFR–FFR strategy would significantly increase the adoption of physiology guided coronary interventions, while drastically reducing the requirement for invasive pressure measurements.

The proposed method is potentially well suited for a clinical setting given the relatively short execution time on a standard hardware configuration. The diagnostic accuracy of the proposed hybrid approach is high (96%), and in the same range as that of previously published purely invasive hybrid IFR–FFR decision making strategies, which varied in the range of 94–95% (Petraco et al. 2013a, b, 2014a, b).

The approach only requires knowledge of the coronary luminal geometry, which can be determined from routine coronary angiography. For around two thirds of the lesions and almost half of the patients in this study, CAD could have been assessed by the proposed hybrid strategy without the need for invasive pressure measurements, which are linked to additional procedure time, equipment, cost and risk. The target accuracy of the hybrid decision-making strategy is closely linked to the percentage of lesions and patients no longer requiring invasive measurements. The previously proposed invasive IFR hybrid window of 0.86–0.93 leads to an agreement of 96% with the FFR-only based assessment. However, if an agreement of 92% was considered sufficient, the number of required invasive measurements could decrease significantly (only 24% of the lesions would have required invasive measurements in this study).

The vast majority of patients in this study had multi-vessel disease: the average number of lesions per patient was 1.95. This is much higher than the average number encountered in other similar studies, which varied between 1.13 and 1.37 (Morris et al. 2013; Tu et al. 2014; Papafaklis et al. 2014; Tröbs et al. 2016; Petraco et al. 2014a, b). A larger average number of lesions per patient increases the likelihood that at least one lesion has a c-IFR value in the hybrid window, and thus reduces the number of patients for which no invasive pressure measurements are required. Nevertheless, even a reduction in the number of invasive pressure measurements performed for a patient has positive effects on costs, duration and risks associated with the procedure. On the other hand, as noted for the hybrid invasive IFR–FFR strategy (Petraco et al. 2013a, b), if the FFR grey zone 0.75–0.80 was accounted for, the proportion of lesions and patients no longer requiring invasive measurements would further increase, whilst maintaining the same accuracy with the FFR-only strategy (as a result of a smaller hybrid window in the c-IFR–FFR strategy).

As displayed in Fig. 3.6, disagreement between the hybrid c-IFR–FFR strategy and the FFR-only strategy was found in only five lesions. In the range of negative c-IFR values (>0.93), disagreement was found in 1 lesion, with an FFR value in the FFR grey zone (0.77). In the range of positive c-IFR values (<0.86), disagreement was found in four lesions, whereas three of them had FFR values between 0.80 and 0.85.

Since c-IFR is computed at all locations in the reconstructed anatomical model, a natural extension of this work is to generate virtual pull-back curves, where the variation of c-IFR along a path from inlet to outlet is depicted. Such an analysis could be useful for determining the most significant lesion in case of serial stenoses. Furthermore, once the most significant stenosis is determined, a virtual PCI may be performed (i.e. virtual stenting), and virtual post-PCI hemodynamics may be evaluated. Such an approach may help the clinician in selecting the optimal treatment plan for the patient.

Recently, a machine learning model for FFR computation from CCTA data, as an alternative to CFD-based modelling, was proposed and it was shown that the performance of the machine learning model is statistically not discernible from that of the CFD approach (Itu et al. 2016). Similarly, a machine learning model may be trained to predict c-IFR. Once the model is trained, the computational time for predicting the c-IFR value would be significantly lower than that of the CFD model, providing results in near real-time.

The motivation to perform invasive FFR was mostly clinical, which resulted in a large proportion of angiographically borderline lesions in a population with extensive atherosclerotic disease. Only few cases were excluded because of the inability to find two proper views of the lesions of interest. The results should be interpreted with the consideration that this was a retrospective single-center study that included a relatively small patient cohort.

The definition of the parameters of the blood flow model is based on physiological assumptions, which would require validation on larger data sets. Such assumptions include for example the allometric scaling laws applied for estimating the flow rate distribution, and the use of population-averaged rheological properties of the blood (density, viscosity, etc.). Contrast-propagation information could be additionally used to improve the estimation of total flow rate and its distribution towards the outlet branches (Tu et al. 2014). Moreover, the collateral circulation has not been taken into account. Collaterals can have a significant impact on the hemodynamics, especially for very severe lesions.

The anatomical model that was reconstructed from angiography and used for CFD based modelling may not accurately reflect the true luminal geometry due to patient motion, e.g. in case of slight respiratory changes. Furthermore, the manual correction of automatically detected vessel contours introduces intra- and inter-observer variability.

The performance of the model needs to be further evaluated in patients that were excluded from the current study, e.g. low systolic pressure, pathological aortic valves, recent myocardial infarction, etc.

Finally, to validate our findings and to provide more representative results the proposed method and hybrid decision making strategy require further validation in larger, prospective studies, conducted at multiple clinical sites.

References

Coenen A, Lubbers MM, Kurata A, Kono A, Dedic A, Chelu RG, Dijkshoorn ML, Gijsen FJ, Ouhlous M, van Geuns RJM, Nieman K (2015) Fractional flow reserve computed from noninvasive CT angiography data: diagnostic performance of an on-site clinician-operated computational fluid dynamics algorithm. Radiology 274:674–683

Dattilo PB, Prasad A, Honeycutt E, Wang TY, Messenger JC (2012) Contemporary patterns of fractional flow reserve and intravascular ultrasound use among patients undergoing percutaneous coronary intervention in the United States: insights from the National Cardiovascular Data Registry. J Am Coll Cardiol 60:2337–2339

Davies JE, Whinnett ZI, Francis DP et al (2006) Evidence of a dominant backward-propagating “suction” wave responsible for diastolic coronary filling in humans, attenuated in left ventricular hypertrophy. Circulation 113:1768–1778

Davies HCS et al (2012) Apparatus and method of assessing a narrowing in a fluid filled tube, WO patent, Medsolve Limited, WO2012093266

De Bruyne B, Pijls NH, Kalesan B, Barbato E, Tonino PA, Piroth Z, Jagic N, Möbius-Winkler S, Rioufol G, Witt N, Kala P, MacCarthy P, Engström T, Oldroyd KG, Mavromatis K, Manoharan G, Verlee P, Frobert O, Curzen N, Johnson JB, Jüni P, Fearon WF (2012) Fractional flow reserve–guided PCI versus medical therapy in stable coronary disease. N Engl J Med 367:991–1001

Escaned J, Echavarría-Pinto M, Garcia-Garcia HM, van de Hoef TP, de Vries T, Kaul P, Raveendran G, Altman JD, Kurz HI, Brechtken J, Tulli M, Von Birgelen C, Schneider JE, Khashaba AA, Jeremias A, Baucum J, Moreno R, Meuwissen M, Mishkel G, van Geuns RJ, Levite H, Lopez-Palop R, Mayhew M, Serruys PW, Samady H, Piek JJ, Lerman A, ADVISE II Study Group (2015) Prospective assessment of the diagnostic accuracy of instantaneous wave-free ratio to assess coronary stenosis relevance: results of ADVISE II international, multicenter study (ADenosine Vasodilator Independent Stenosis Evaluation II). JACC Cardiovasc Interv 8:824–833

Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR, Smith SC, Spertus JA, Williams SV (2012) 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines, and the American College of Physicians, American Association for

Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 60:e44–e164

Götberg M, Christiansen EH, Gudmundsdottir I, Sandhall L, Omerovic E, James SK, Erlinge D, Fröbert O (2015) Instantaneous Wave-Free Ratio versus Fractional Flow Reserve guided intervention (iFR-SWEDEHEART): rationale and design of a multicenter, prospective, registry-based randomized clinical trial. Am Heart J 170:945–950

Hachamovitch R, Di Carli MF (2008) Methods and limitations of assessing new noninvasive tests part II: outcomes-based validation and reliability assessment of noninvasive testing. Circulation 117:2793–2801

Heller L et al (1994) Blood flow velocity in the right coronary artery: assessment before and after angioplasty. J Am Coll Cardiol 24:1012–1017

Huo Y et al (2012) A validated predictive model of coronary fractional flow reserve. J R Soc Interface 9:1325–1338

Hutchins GM, Miner MM, Boitnott JK (1976) Vessel caliber and branch-angle of human coronary artery branch-points. Circ Res 38:572–576

L. Itu et al. (2012) A patient-specific reduced-order model for coronary circulation. In: Intern Symp Biomed Imag, Barcelona, Spain, pp 832–835

Itu L et al (2015) A parameter estimation framework for patient-specific hemodynamic computations. J Comput Phys 281:316–333

Itu L, Rapaka A, Passerini T, Georgescu B, Schwemmer C, Schoebinger M, Flohr T, Sharma P, Comaniciu D (2016) A machine learning approach for computation of fractional flow reserve from coronary computed tomography. J Appl Physiol 121(1):42–52

Jeremias A (2015) Functional lesion assessment of intermediate stenosis to guide revascularization (DEFINE-FLAIR); NCT02053038. http://clinicaltrials.gov/show/NCT02053038

Jeremias A, Maehara A, Généreux P, Asrress KN, Berry C, De Bruyne B, Davies JE, Escaned J, Fearon WF, Gould KL, Johnson NP, Kirtane AJ, Koo BK, Marques KM, Nijjer S, Oldroyd KG, Petraco R, Piek JJ, Pijls NH, Redwood S, Siebes M, Spaan JA, van't Veer M, Mintz GS, Stone GW (2014) Multicenter core laboratory comparison of the instantaneous wave-free ratio and resting Pd/Pa with fractional flow reserve: the RESOLVE study. J Am Coll Cardiol 63:1253–1261

Kamiya A, Togawa T (1980) Adaptive regulation of wall shear stress to flow change in the canine carotid artery. Am J Physiol 239:14–21

Kassab G et al (1995) The pattern of coronary arteriolar bifurcations and the uniform shear hypothesis. Ann Biomed Eng 23:13–20

Kern MJ, Lerman A, Bech JW, De Bruyne B, Eeckhout E, Fearon WF, Higano ST, Lim MJ, Meuwissen M, Piek JJ, Pijls NH, Siebes M, Spaan JA (2006) Physiological assessment of coronary artery disease in the cardiac catheterization laboratory: a scientific statement from the American Heart Association Committee on Diagnostic and Interventional Cardiac Catheterization, Council on Clinical Cardiology. Circulation 114:1321–1341

Koo BK, Erglis A, Doh JH, Daniels DV, Jegere S, Kim HS, Dunning A, DeFrance T, Lansky A, Leipsic J, Min JK (2011) Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms: results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study. J Am Coll Cardiol 58:1989–1997

Levin DC (1982) Invasive evaluation (coronary arteriography) of the coronary artery disease patient: clinical, economic and social issues. Circulation 66:III71–III79

Meijboom WB, Van Mieghem CA, van Pelt N, Weustink A, Pugliese F, Mollet NR, Boersma E, Regar E, van Geuns RJ, de Jaegere PJ, Serruys PW, Krestin GP, de Feyter PJ (2008) Comprehensive assessment of coronary artery stenoses: computed tomography coronary angiography versus conventional coronary angiography and correlation with fractional flow reserve in patients with stable angina. J Am Coll Cardiol 52:636–643

Morris PD, Ryan D, Morton AC, Lycett R, Lawford PV, Hose DR, Gunn JP (2013) Virtual Fractional Flow Reserve from coronary angiography: modeling the significance of coronary lesions. J Am Coll Cardiol Interv 6:149–157

Murray CD (1926a) The physiological principle of minimum work: I. The vascular system and the cost of blood volume. Proc Natl Acad Sci U S A 12:207–214

Murray CD (1926b) The physiological principle of minimum work: II. Oxygen exchange in capillaries. Proc Natl Acad Sci U S A 12:299–304

Ng VG, Lansky AJ (2011) Novel QCA methodologies and angiographic scores. Int J Cardiovasc Imaging 27:157–165

Nijjer SS, Sen S, Petraco R, Sachdeva R, Cuculi F, Escaned J, Broyd C, Foin N, Hadjiloizou N, Foale RA, Malik I, Mikhail GW, Sethi AS, Al-Bustami M, Kaprielian RR, Khan MA, Baker CS, Bellamy MF, Hughes AD, Mayet J, Kharbanda RK, Di Mario C, Davies JE (2013) Improvement in coronary haemodynamics after percutaneous coronary intervention: assessment using instantaneous wave-free ratio. Heart 99:1740–1748

Nijjer SS, Sen S, Petraco R, Escaned J, Echavarria-Pinto M, Broyd C, Al-Lamee R, Foin N, Foale RA, Malik IS, Mikhail GW, Sethi AS, Al-Bustami M, Kaprielian RR, Khan MA, Baker CS, Bellamy MF, Hughes AD, Mayet J, Francis DP, Di Mario C, Davies JE (2014) Pre-angioplasty instantaneous wave-free ratio pullback provides virtual intervention and predicts hemodynamic outcome for serial lesions and diffuse coronary artery disease. JACC Cardiovasc Interv 7:1386–1396

Papafaklis MI, Muramatsu T, Ishibashi Y, Lakkas LS, Nakatani S, Bourantas CV, Ligthart J, Onuma Y, Echavarria-Pinto M, Tsirka G, Kotsia A, Nikas DN, Mogabgab O, van Geuns RJ, Naka KK, Fotiadis DI, Brilakis ES, Garcia-Garcia HM, Escaned J, Zijlstra F, Michalis LK, Serruys PW (2014) Fast virtual functional assessment of intermediate coronary lesions using routine angiographic data and blood flow simulation in humans: comparison with pressure wire - fractional flow reserve. EuroIntervention 10:574–583

Parker KH, Jones CJ, Dawson JR, Gibson DG (1988) What stops the flow of blood from the heart. Heart Vessel 4(4):2415

Petraco R, Escaned J, Sen S, Nijjer S, Asrress KN, Echavarria-Pinto M, Lockie T, Khawaja MZ, Cuevas C, Foin N, Broyd C, Foale RA, Hadjiloizou N, Malik IS, Mikhail GW, Sethi A, Kaprielian R, Baker CS, Lefroy D, Bellamy M, Al-Bustami M, Khan MA, Hughes AD, Francis DP, Mayet J, Di Mario C, Redwood S, Davies JE (2013a) Classification performance of instantaneous wave-free ratio (iFR) and fractional flow reserve in a clinical population of intermediate coronary stenoses: results of the ADVISE registry. EuroIntervention 9: 91–101

Petraco R, Park JJ, Sen S, Nijjer SS, Malik IS, Echavarría-Pinto M, Asrress KN, Nam CW, Macías E, Foale RA, Sethi A, Mikhail GW, Kaprielian R, Baker CS, Lefroy D, Bellamy M, Al-Bustami M, Khan MA, Gonzalo N, Hughes AD, Francis DP, Mayet J, Di Mario C, Redwood S, Escaned J, Koo BK, Davies JE (2013b) Hybrid iFR-FFR decision-making strategy: implications for enhancing universal adoption of physiology-guided coronary revascularization. EuroIntervention 8:1157–1165

Petraco R, Al-Lamee R, Gotberg M, Sharp A, Hellig F, Nijjer SS, Echavarria-Pinto M, van de Hoef TP, Sen S, Tanaka N, Van Belle E, Bojara W, Sakoda K, Mates M, Indolfi C, De Rosa S, Vrints CJ, Haine S, Yokoi H, Ribichini FL, Meuwissen M, Matsuo H, Janssens L, Katsumi U, Di Mario C, Escaned J, Piek J, Davies JE (2014a) Real-time use of instantaneous wave-free ratio: results of the ADVISE in-practice: an international, multicenter evaluation of instantaneous wave-free ratio in clinical practice. Am Heart J 168:739–748

Petraco R, van de Hoef TP, Nijjer S, Sen S, van Lavieren MA, Foale RA, Meuwissen M, Broyd C, Echavarria-Pinto M, Foin N, Malik IS, Mikhail GW, Hughes AD, Francis DP, Mayet J, Di Mario C, Escaned J, Piek JJ, Davies JE (2014b) Baseline instantaneous wave-free ratio as a pressure-only estimation of underlying coronary flow reserve: results of the JUSTIFY-CFR study (Joined Coronary Pressure and Flow Analysis to Determine Diagnostic Characteristics of Basal and Hyperemic Indices of Functional Lesion Severity-Coronary Flow Reserve). Circ Cardiovasc Interv 7:492–502

Pijls NH, de Bruyne B, Peels K, van der Voort PH, Bonnier HJ, Bartunek J, Koolen JJ (1996) Measurement of fractional flow reserve to assess the functional severity of coronary-artery stenoses. N Engl J Med 334:1703–1708

Renker M, Schoepf UJ, Wang R, Meinel FG, Rier JD, Bayer RR, Möllmann H, Hamm CW, Steinberg DH, Baumann S (2014) Comparison of diagnostic value of a novel noninvasive coronary computed tomography angiography method versus standard coronary angiography for assessing fractional flow reserve. Am J Cardiol 114:1303–1308

Ryan TJ (2002) The coronary angiogram and its seminal contributions to cardiovascular medicine over five decades. Circulation 106:752–756

Sen S, Escaned J, Malik IS, Mikhail GW, Foale RA, Mila R, Tarkin J, Petraco R, Broyd C, Jabbour R, Sethi A, Baker CS, Bellamy M, Al-Bustami M, Hackett D, Khan M, Lefroy D, Parker KH, Hughes AD, Francis DP, Di Mario C, Mayet J, Davies JE (2012) Development and validation of a new adenosine-independent index of stenosis severity from coronary wave-intensity analysis results of the ADVISE (ADenosine Vasodilator Independent Stenosis Evaluation) study. J Am Coll Cardiol 59:1392–1402

Sen S, Asrress KN, Nijjer S, Petraco R, Malik IS, Foale RA, Mikhail GW, Foin N, Broyd C, Hadjiloizou N, Sethi A, Al-Bustami M, Hackett D, Khan MA, Khawaja MZ, Baker CS, Bellamy M, Parker KH, Hughes AD, Francis DP, Mayet J, Di Mario C, Escaned J, Redwood S, Davies JE (2013) Diagnostic classification of the instantaneous wave-free ratio is equivalent to fractional flow reserve and is not improved with adenosine administration. Results of CLARIFY (Classification Accuracy of Pressure-Only Ratios Against Indices Using Flow Study). J Am Coll Cardiol 61:1409–1420

Spiller P et al (1983) Measurement of systolic and diastolic flow rates in the coronary artery system by X-ray densitometry. Circulation 68:337–347

Taylor CA, Steinman DA (2010) Image-based modeling of blood flow and vessel wall dynamics: applications, methods and future directions. Ann Biomed Eng 38:1188–1203

Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, vant Veer M, Klauss V, Manoharan G, Engstrøm T, Oldroyd KG, Ver Lee PN, PA MC, Fearon WF (2009) Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med 360:213–224

Toth G, Hamilos M, Pyxaras S, Mangiacapra F, Nelis O, De Vroey F, Di Serafino L, Muller O, Van Mieghem C, Wyffels E, Heyndrickx GR, Bartunek J, Vanderheyden M, Barbato E, Wijns W, De Bruyne B (2014) Evolving concepts of angiogram: fractional flow reserve discordances in 4000 coronary stenoses. Eur Heart J 35:2831–2838

Tröbs M, Achenbach S, Röther J, Redel T, Scheuering M, Winneberger D, Klingenbeck K, Itu L, Passerini T, Kamen A, Sharma P, Comaniciu D, Schlundt C (2016) Comparison of Fractional Flow Reserve based on computational fluid dynamics modeling using coronary angiographic vessel morphology versus invasively measured Fractional Flow Reserve. Am J Cardiol 117:29–35

Tu S, Barbato E, Köszegi Z, Yang J, Sun Z, Holm NR, Tar B, Li Y, Rusinaru D, Wijns W, Reiber JH (2014) Fractional Flow Reserve calculation from 3-dimensional quantitative coronary angiography and TIMI frame count: a fast computer model to quantify the functional significance of moderately obstructed coronary arteries. JACC Cardiovasc Interv 7:768–777

Uylings H (1977) Optimization of diameters and bifurcation angles in lung and vascular tree structures. Bull Math Biol 39:509–520

van der Giessen A, Groen H, Doriot PA, de Feyter P, van der Steen A, van de Vosse F, Wentzel JJ, Gijsen FJ (2011) The influence of boundary conditions on wall shear stress distribution in patients specific coronary trees. J Biomech 44:1089–1095

Wieneke H et al (2005) Determinants of coronary blood flow in humans: quantification by intracoronary Doppler and ultrasound. J Appl Physiol 98:1076–1082

Windecker S, Kolh P, Alfonso F, Collet JP, Cremer J, Falk V, Filippatos G, Hamm C, Head SJ, Jüni P, Kappetein AP, Kastrati A, Knuuti J, Landmesser U, Laufer G, Neumann FJ, Richter DJ, Schauerte P, Uva MS, Stefanini GG, Taggart DP, Torracca L, Valgimigli M, Wijns W, Witkowski A (2014) 2014 ESC/EACTS Guidelines on myocardial revascularization. Eur Heart J 46:517–592

Zamir M, Sinclair P, Wonnacott TH (1992) Relation between diameter and flow in major branches of the arch of the aorta. J Biomech 25:1303–1310

A Parameter Estimation Framework for Patient-Specific Assessment of Aortic Coarctation

4

Lucian Itu, Puneet Sharma, Tiziano Passerini, Ali Kamen,
and Constantin Suciu

Abstract

In this chapter we introduce a method based on computational fluid dynamics for non-invasively assessing patients with aortic coarctation. While in practice the pressure gradient across the coarctation is typically measured invasively with a catheter, the proposed method determines the pressure gradient using a computational modeling approach, which relies on medical imaging data, routine non-invasive clinical measurements and physiological principles. The main components of the method are a reduced-order model coupled with a comprehensive pressure-drop formulation, and a parameter estimation method for personalizing the boundary conditions and the vessel wall parameters. The parameter estimation method is fully automated, and is based on an iterative tuning procedure to obtain a close match between the computed and the non-invasively determined quantities. A key feature is a warm-start to the optimization procedure, with better initial solution for the nonlinear system of equations, to reduce the number of iterations needed for the calibration of the geometrical multiscale models. To achieve these goals, the initial solution, computed with a lumped

---

Large parts of Sects. 4.2, 4.3 and 4.4 have been published before in the paper ‘A Parameter Estimation Framework for Patient-specific Hemodynamic Computations’, Journal of Computational Physics, Volume 281, 15 January 2015, Pages 316–333.

L. Itu (✉) • C. Suciu
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

Automation and Information Technology, Transilvania University of Brasov,
Mihai Viteazu nr. 5, Brasov 5000174, Romania
e-mail: lucian.itu@siemens.com

P. Sharma • T. Passerini • A. Kamen
Medical Imaging Technologies, Siemens Healthcare, 755 College Road,
Princeton, NJ 08540, USA

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_4

parameter model, is adapted before solving the parameter estimation problem for the geometrical multiscale circulation model: the resistance and the compliance of the circulation model are estimated and compensated. This feature is based on research, and is not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

4.1 Introduction

Aortic coarctation (CoA) represents a congenital cardiac pathology which is encountered in 5–8% of the patients with congenital heart pathologies (Ringel and Jenkins 2007). Its luminal appearance is similar to that of an atherosclerotic stenosis: a local narrowing of the aortic wall. Different methods are used for assessing the CoA severity: the anatomy is typically assessed using Magnetic Resonance Imaging (MRI) or Computed Tomography (CT). The functional assessment of aortic coarctations is based on the pressure drop along the CoA. The gold standard for determining the pressure drop is invasive catheterization. Due to the costs and the risks associated with this invasive procedure, other less accurate methods have also been proposed. The velocities measured using Doppler echocardiography are used in the simplified or modified Bernoulli equations to determine the pressure drop, but, on one hand, these velocities are sometimes difficult to measure due to the placement of the descending aorta, and, on the other hand, the computed values often overestimate the actual gradient (Seifert et al. 1999). The difference between cuff-based blood pressure measurements performed at the arms and legs represents another simple, non-invasive and cost-effective method, but is also unreliable (Hom et al. 2008).

To address this issue, Computational Fluid Dynamics (CFD) based blood-flow models have been proposed for analyzing the hemodynamics in idealized and/or patient-specific healthy and diseased aortic geometries. These models provide important insights into the structure and function of the cardiovascular system, and have been proposed in recent years, for diagnosis, risk stratification, and surgical planning (Cebral et al. 2011; Haggerty et al. 2013; Quarteroni et al. 2000; Taylor and Steinman 2010).

Different approaches have been proposed for specifying the outlet boundary conditions, ranging from pressure or flow rate profiles to lumped parameter models (0D models). For an accurate patient-specific computation, the role of physiologically sound boundary conditions is well appreciated in the literature. The effect of distal vasculature is modeled by outlet boundary conditions coupled with the computational domain (region of interest), resulting in a geometrical multiscale model. The boundary conditions are represented by lumped parameter models, which are designed to capture one (or more) of the (1) total resistance, (2) total compliance, and (3) the wave propagation and reflection effects in the distal vasculature. The most widely used lumped parameter model is the three-element windkessel model (Westerhof et al. 2009), which is characterized by its simplicity (only three parameters), and ability to capture two important characteristics of the distal circulation (compliance and resistance).

In a clinical scenario, the values of the windkessel model parameters are not available on a per-patient basis. Instead, multiple pressure or flow measurements are usually available for each patient. A clinically feasible and accurate flow computation should not only be in agreement with these measurements, but should also have means to model other hemodynamic states for the same patient. To achieve this, one has to estimate a set of personalized windkessel model parameters, while ensuring that the computations match the measured data.

Different calibration procedures for the outlet boundary conditions have been proposed. Olufsen et al. proposed a calibration method for determining the dynamic cerebral blood flow response to sudden hypotension during posture change (Olufsen et al. 2002). Their experimental calibration procedure depends strongly on the patient-specific state, i.e. on the position of the patient.

A fully automatic optimization-based calibration method for the windkessel models was suggested (Spilker and Taylor 2010), where the input was specified by non-invasively acquired systolic/diastolic pressures and, in some cases, additional flow data. The windkessel parameters were obtained by solving a system of nonlinear equations, formulated based on a set of objectives for the pressure and flow rate waveforms at various locations. A Broyden method was employed for solving the nonlinear system of equations. The initial parameter values were determined by a reduced-order model, composed of only the windkessel models of the geometrical multiscale model.

An adjoint based method for calibrating the windkessel parameters was proposed (Ismail et al. 2013a, b), where the Jacobian was computed without the use of finite-differences, and eliminated the risk of that the Broyden's method would not converge for initial guesses far away from the solution. Furthermore, a reduced-order model with resistance outlet boundary conditions was introduced (Blanco et al. 2012), under which the terminal resistance values of the arterial model of the arm were adapted to obtain desired flow rate distribution between vascular territories.

A competitive alternative to the above mentioned optimization based methods is represented by filtering based methods. These methods were successfully used to estimate different aspects of fluid-structure interaction applications, like arterial wall stiffness (Pant et al. 2014), the surrounding tissue support (Moireau et al. 2013), or windkessel parameters (Bertoglio et al. 2012).

Herein we propose a parameter estimation method for personalizing hemodynamic computations (Itu et al. 2015). The proposed method is inspired by the approach introduced previously (Spilker and Taylor 2010), and has been developed based on the following strategies: a warm-start of the optimization procedure with better initial solution, and a reduction in the number of iterations performed for the calibration of the geometrical multiscale models (for simplifying phrasing, we will refer to geometrical multiscale models simply as multiscale models). The first strategy aims at reducing the risk of a possible failure of the optimization approach due to a bad initial guess (Ismail et al. 2013a, b)—a fairly common occurrence for pathologic cases. Moreover, we can achieve a faster computation time for the overall estimation method by reducing the number of repetitive iterations on the same geometry.

The parameter estimation method automatically determines the windkessel parameters in the multiscale circulation model, by solving a system of nonlinear equations. To obtain an initial guess, the equations are first solved for a lumped parameter model. The main characteristic of the proposed method is that, when switching from the lumped model to the multiscale model, the windkessel parameters are appropriately adapted to take into account the hemodynamic properties (resistance and compliance) of the multiscale model.

The concepts and information presented in this chapter are based on research and are not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

4.2 Methods

In this section, we describe an efficient optimization-based algorithm, which ensures that the personalized flow computations are in close agreement with the physiological measurements. In the present study, we use a 1D–0D reduced-order geometrical multiscale model as proof of concept. One-dimensional models have been shown to accurately predict time-varying flow rate and pressure wave forms (Reymond et al. 2011). Most of the one-dimensional models introduced in literature use elastic wall laws (Stergiopulos et al. 1992; Olufsen et al. 2000; Formaggia et al. 2003; Mynard and Nithiarasu 2008), but viscoelastic arterial wall models have also been applied (Alastruey et al. 2011; Malossi et al. 2012). Further, recent research activities have shown the growing interest in the one-dimensional blood flow model not only for the computation of a full body arterial model, but also for specific parts of the circulation in pathologic situations: the coronary circulation (Itu et al. 2012), the abdominal aorta (Raghu et al. 2011; Low et al. 2012), proximal part of the aorta (Itu et al. 2013), and the aortic valve (Mynard et al. 2012). The 1D model used in this study has been previously introduced in (Itu et al. 2012, 2013). Time-varying flow rate profiles are used as inlet boundary condition, while three-element windkessel models were coupled at the outlets of the highest order model. The 3-element windkessel model is represented by the following relationship between instantaneous flow and pressure:

$$\frac{dp}{dt} = R_p \frac{dq}{dt} - \frac{p}{R_d \cdot C} + \frac{q\left(R_p + R_d\right)}{R_d \cdot C}, \tag{4.1}$$

where p is the instantaneous pressure at the inlet of the windkessel model, q is the instantaneous flow rate, R_p and R_d are the proximal and distal resistance respectively, and C is the compliance.

Next, we briefly review the calibration method proposed previously and introduce the proposed parameter estimation method.

4.2.1 Calibration Method for Windkessel Parameters

The calibration method automatically estimates the free parameters (windkessel parameters) to ensure that the computed pressure and flow-rate values minimize the objective function (Spilker and Taylor 2010). More specifically, the algorithm iteratively estimates the total resistance of the windkessel model, R_t, the proximal resistance ratio, ρ, and the time constant of the exponential pressure decay in the windkessel model at zero flow, τ:

$$R_t = R_p + R_d, \tag{4.2}$$

$$\rho = R_p / R_t, \tag{4.3}$$

$$\tau = R_d \cdot C. \tag{4.4}$$

Measured pressure and/or flow rate values are used as objectives of the calibration method. The parameter estimation problem is formulated as a solution to a system of nonlinear equations, with each equation representing the residual error between the computed and measured quantity of interest. To determine the values of all the residuals ($\boldsymbol{f}(\boldsymbol{x}_i)$), a computation with the parameter values $\boldsymbol{x}_i$ is required. Since the absolute values of the adapted parameters and of the residuals generally differ by orders of magnitude, for the calibration method both the parameter and the objective residuals have been scaled using typical values, as is described below.

The nonlinear system of equations is first solved for a 0D model, composed of the windkessel models used in the multiscale model. To find an initial solution for the 0D model, a grid of physiological parameter value sets is considered. The parameter value set leading to the smallest L_2 norm for the objective residual, is used as initial solution for a dogleg trust region method, which determines the solution $\boldsymbol{x}_0$ used during the subsequent steps.

Next, a fixed-point approach is used to compute a finite-difference Jacobian which is determined using the typical step sizes, and is consistent with the chosen typical values of the objective residuals. The components of the Jacobian approximations are computed as follows:

$$J_{ij} = \frac{1}{s_j^{typ}}\left[\mathbf{f}\left(\mathbf{x}_0 + \frac{1}{2}s_j^{typ}\mathbf{e}_j\right) - \mathbf{f}\left(\mathbf{x}_0 - \frac{1}{2}s_j^{typ}\mathbf{e}_j\right)\right]\cdot \mathbf{e}_i, \tag{4.5}$$

where $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ represent the unit vectors in the *i*th and *j*th direction respectively, and s_j^{typ} is the typical step size for parameter *j*, given by:

$$s_j^{typ} = 1 \Big/ \sqrt{\sum_{i=1}^{n_{eq}} \left(J_{ij} / f_i^{typ}\right)}, \tag{4.6}$$

where f_i^{typ} is the typical value of the objective residual f_i.

The process composed of Eqs. (4.5) and (4.6) is an iterative procedure which is terminated once the Euclidian norm of the difference of two consecutive Jacobians, normalized by the corresponding f_i^{typ} and s_j^{typ} values, is less than 10^{-6}.

Next, the multiscale model is set up and run, and the objective residuals are evaluated. For the first run with the multiscale model, the solution variables are initialized based on the results obtained with the 0D model. Each computation, with a given set of parameter values, is run until the L_2 norms of the normalized differences between the pressure and flow rate profiles at the current and the previous cardiac cycle are smaller than 10^{-5}.

If all objective residuals are smaller than the tolerance limit (taken here equal to $f_i^{typ}/10$), the calibration method is terminated. Otherwise, a quasi-Newton method is employed to update the parameter values. First the Jacobian matrix is updated: since the computation of a finite-difference Jacobian matrix with the multiscale model is expensive, the Jacobian determined with the fixed-point approach is updated at each iteration, using:

$$\mathbf{J}_{i+1} = \mathbf{J}_i + \frac{\left[\mathbf{f}\left(\mathbf{x}_{i+1}\right) - \mathbf{f}\left(\mathbf{x}_i\right) - \mathbf{J}_i\mathbf{s}_i\right]\left(\mathbf{D}_s^2\mathbf{s}_i\right)^T}{\mathbf{s}_i^T\mathbf{D}_s^2\mathbf{s}_i}, \tag{4.7}$$

where $\mathbf{s}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ is the current step and $\boldsymbol{D}_s$ is a diagonal scaling matrix:

$$\left(D_s\right)_{ij} = \begin{cases} 1/s_j^{typ}, & i = j \\ 0, & i \neq j \end{cases} \tag{4.8}$$

Finally, the new parameter values are estimated:

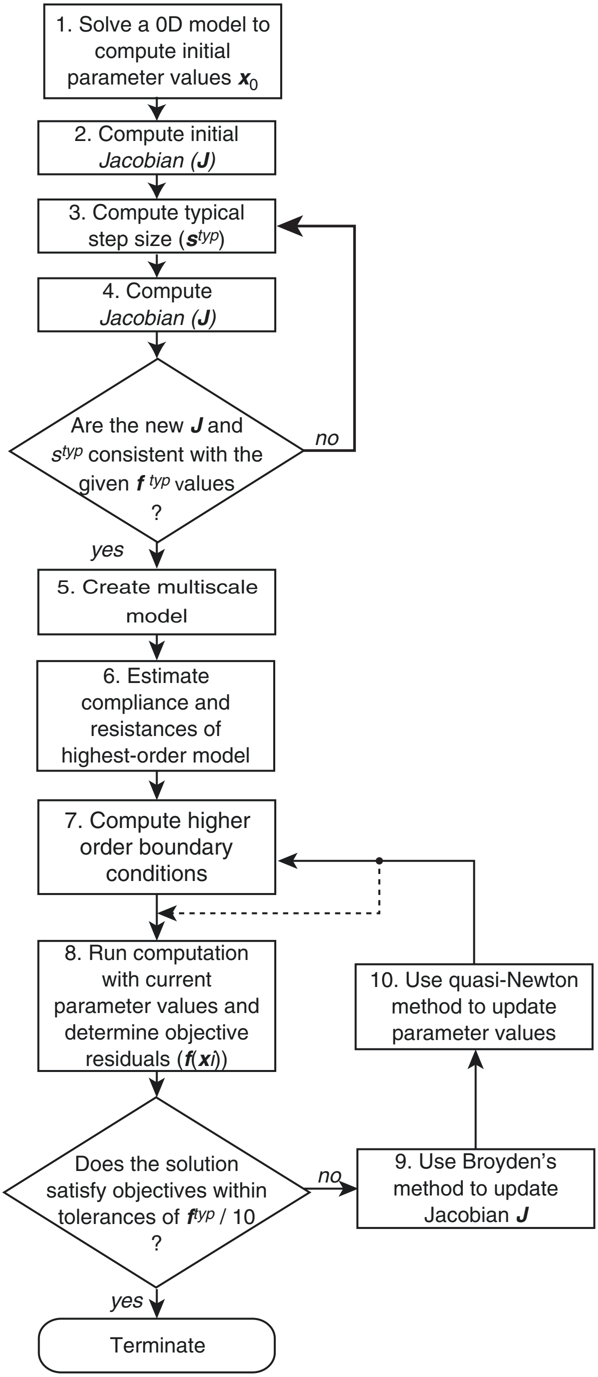
$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{J}_i^{-1}\mathbf{f}\left(\mathbf{x}_i\right). \tag{4.9}$$

The main drawback of the above described calibration method is that the hemodynamic properties of the highest order model of the multiscale model are not taken into account when switching from the 0D model to the multiscale model. This aspect has several consequences. First of all, the solution obtained for the 0D model, which is used as initial solution for the nonlinear system of equations solved for the multiscale model, is considerably different from the final solution of the parameter values. This can lead to a failure of the minimization approach employed for the nonlinear system. Furthermore, this risk is enhanced when the hemodynamic computation is performed for pathologic cases, i.e. when the hemodynamic properties of the highest order model become more important compared to the properties of the 0D model. Secondly, the number of calibration iterations required to obtain the final solution for the parameter values increases. These repetitive computations increase the total execution time, an aspect which is a significant disadvantage when the calibration method is applied for patient specific computations in a clinical setting.

4.2.2 Proposed Parameter Estimation Framework

We introduce a parameter estimation method which eliminates the drawbacks of the calibration method described in the previous section. An outline of the algorithm is illustrated in Fig. 4.1, which is composed of the steps described previously and two critical steps (displayed in bold: steps 6 and 7). One of the main advantages of the proposed algorithm is that it can be efficiently applied not only for healthy cases, but also for pathologic cases. The main idea is to efficiently account for the hemodynamic

Fig. 4.1 The proposed parameter estimation method (Itu et al. 2015)



properties of the highest order model inside the multiscale model, when switching from the 0D model to the multiscale model. Particularly, we focus on two properties of the multiscale model: the compliance and the resistance.

At step 6, the compliance and resistances of the highest order model are estimated. These two quantities have not been taken into account when solving the nonlinear system of equations with the 0D model. Next, at step 7, the parameter values of the windkessel models are recomputed to account for the hemodynamic properties of the highest order model. Note that the values of the parameters of the nonlinear system are not modified at this stage and only the windkessel parameters are changed. The algorithms used for the estimation and compensation of the resistances and compliances are described in the next two subsections.

As a result of the two additional steps, the parameter estimation method adapts the overall properties of the multiscale models. For the multiscale computation, the windkessel parameters are the only quantities updated from one iteration to the next.

The parameter estimation methods do not always use the total resistances or the compliances as calibrated parameters. Nevertheless, regardless of the choice of parameters, the resistances and the compliances of the terminal windkessel models are always adapted when switching from the 0D model to the multiscale model. Additionally, if the resistances or the compliance of the multiscale model are adapted directly or indirectly, the parameters of the windkessel models are recomputed at the end of each iteration, based on the new parameter values determined through Eq. (4.9) and the estimated quantities of the highest order model. If the quantities are not adapted at all, then step 7 is applied only once (an aspect indicated by the dashed line in Fig. 4.1). The target objectives are: maximum pressure, P_{max}, minimum pressure, P_{min}, and flow rate split between outlets of the highest order model, Φ. The typical value of the residuals for pressure and flow-rate split objectives was set to 1 mmHg and 0.005 respectively.

4.2.3 Resistance Estimation and Compensation

To estimate the resistance of a healthy vascular segment for the highest order model, $(R_{HO})_i$, we assume a velocity profile:

$$\left(R_{HO}\right)_i = \frac{2(\gamma+2)\mu}{\pi}\int_0^{L_i}\frac{1}{r_i^4(l)}dl \tag{4.10}$$

where γ is the power coefficient in the velocity profile $u=\bar{U}\frac{\gamma+2}{\gamma}\left[1-\left(\frac{r}{R}\right)^{\gamma}\right]$, μ is the viscosity and r_i is the radius. For this work, we have chosen a parabolic profile (γ = 2) without any loss of generality. Alternatively a power law profile (γ = 9) or a Womersley profile can be considered (Womersley 1955).

For a pathologic vascular segment (e.g. coarctation, stenosis) we use a comprehensive pressure-drop model to estimate the resistance of the pathological segment.

This approach was validated against catheter-based ground-truth measurements in our previous work (Itu et al. 2013). Since the time-varying flow rate through the descending aorta is not known apriori, the resistance of the CoA segment is computed from the average flow rate (which represents an objective of the parameter estimation method, i.e. it is known apriori):

$$\left(R_{HO}\right)_i = R_{CoA}\left(Q\right) = \Delta P\left(\bar{Q}_{DAo}\right) / \bar{Q}_{DAo}, \tag{4.11}$$

where $\Delta P(\bullet)$ refers to the pressure drop model (Itu et al. 2013), and $\bar{Q}_{DAo}$ is the average flow rate through the descending aorta.

A similar approach can also be used if the multiscale model contains stenosed arteries. Depending on the region of interest, different pressure drop models have been proposed in the literature, e.g. coronary artery stenosis (Huo et al. 2012), femoral artery stenosis (Young et al. 1975), renal artery stenosis (Steele 2007). For aneurysms, a different pressure drop model (Bessems 2007) can be applied to estimate the resistance.

Equations (4.10)/(4.11) are used at step 6 in Fig. 4.1 to determine the resistance of each branch of the multiscale model. Next, we introduce a recursive algorithm, employed at step 7, for adapting the total terminal resistances of the windkessel models, which is used during the switch from the 0D model to the multiscale model. Let n be the number of outlets in the geometric model, $(R_{t-0D})_j$ represents the total (sum of proximal and distal) resistance of the windkessel model used at each outlet ($j = 1 \ldots n$). The objective is to estimate the total resistances at each outlet of the geometric multiscale model $(R_{t-MS})_j$. Algorithm 1 illustrates the recursive function used for the resistance adaptation.

Algorithm 1: *adaptResistance* (totalRes, vesselNr)

```
    $R_t \leftarrow$ totalRes
    $i \leftarrow$ vesselNr
    $(R_{t-MS})_i \leftarrow R_t - (R_{HO})_i$
    if vessel $i$ is a terminal vessel
    return
    else
    $\left(R_{t-0D}\right)_i \leftarrow 1 \Big/ \sum_k \frac{1}{\left(R_{t-0D}\right)_k}$,   $k \rightarrow$ terminal vessels downstream from vessel $i$
    for each daughter vessel $j$ of vessel $i$
    $\Phi_j \leftarrow (R_{t-0D})_i/(R_{t-0D})_j$
    $(R_{t-MS})_j \leftarrow (R_{t-MS})_i/\Phi_j$
    adaptResistance $((R_{t-MS})_j, j)$
    end (for)
    end (if)
    return
```

The function *adaptResistance* is called exactly once for the root segment of the arterial tree, which recursively computes all terminal resistances of the multiscale model. If the current segment is a terminal segment, the total terminal resistance is determined by subtracting the multiscale resistance of the current segment from the total resistance at the root of the vessel. If the current segment has daughter vessels, the goal is to distribute the new total resistance, $(R_{t-MS})_i$, to the outlets of segment i, in a manner that maintains the flow rate ratio in each daughter vessel. In the first step, the total resistance of the daughter vessels inside the 0D model is determined. Next, for each daughter vessel, using the ratio of total resistance of current vessel to terminal resistance of the current daughter vessel (Φ_j), the new terminal resistance of the current daughter vessel is determined ($(R_{t-MS})_j$). If the daughter vessel j is not a terminal vessel, $(R_{t-0D})_j$ is computed using the resistances of the downstream terminal vessels. Finally, function *adaptResistance* is called to further distribute the resistance of the current daughter vessel.

Equations (4.10)/(4.11) are used to estimate only once the resistance of each vascular segment. If the resistances are adapted directly or indirectly, Algorithm 1 is also applied, in a slightly modified version, at the end of each calibration iteration performed for the multiscale model (after computing the new parameter values through Eq. (4.9)). Hence, the total resistances of the 0D model used in Algorithm 1, $(R_{t-0D})_j$, are substituted by the total equivalent resistances of the multiscale model at each outlet of the highest order model, $(R_t)_j$, which is determined before applying the modified version of Algorithm 1. If the directly or indirectly adapted resistance refers to a single outlet, $(R_t)_j$ is set equal to this parameter value. Otherwise the parameter value is distributed to all outlets covered by the parameter. For the current study, we use a power law:

$$\left(R_t\right)_j = R_t \cdot \sum_k r_k^m \Big/ r_j^m, \tag{4.12}$$

where R_t refers to the adapted parameter, index k is used to iterate over all outlets covered by the adapted parameter and m is a power coefficient, whose physiological range of values is between two for large arteries (Zamir et al. 1992) and three for small arteries (Murray 1926).

4.2.4 Compliance Estimation and Compensation

To estimate the compliance of a vascular segment i of the highest order model in the multiscale model, $(C_{HO})_i$, we use the material properties of the arterial wall:

$$\left(C_{HO}\right)_i = \int_0^{L_i} \frac{3\pi r_i^2(l)}{2} \cdot \frac{r_i(l)}{E_i \cdot h_i} dl, \tag{4.13}$$

where E_i is the Young's modulus and h_i is the wall thickness. Most of the compliance in an arterial/venous tree resides in the large vessels, hence the compliance of the highest order model may become significant compared to the compliance of the windkessel elements, especially when using domains with large arteries. Next, the total compliance of the highest order model is determined $C_{HO} = \sum_i (C_{HO})_i$, which is then distributed to the outlets of the geometric model:

$$(C_{MS-a})_j = \left(C_{HO} \cdot r_j^2\right) \Big/ \sum_{j=1}^{n} r_j^2 , \tag{4.14}$$

where $(C_{MS-a})_j$ represents the assumed additional compliance introduced at each outlet of the highest order model inside the multiscale model. The goal is to determine the compliance of the windkessel models, $(C_{MS})_j$, so as to obtain for the multiscale model the same total compliance as the one used for the 0D model, $(C_0D)_j$. Figure 4.2 introduces different representations for the compliance of the multiscale model, alongside the 0D model.

Figure 4.2a displays the classic representation of the multiscale model, composed of the highest order model and the terminal windkessel element. In Fig. 4.2b, the compliance of the highest order model is lumped into a 0D compliance at outlet j, $(C_{MS-a})_j$, placed before the windkessel model. To determine the compliances of the windkessel elements of the multiscale model, $(C_{MS})_j$, $(C_{MS-a})_j$ is ported inside the windkessel model, as displayed in Fig. 4.2c. For the models in Fig. 4.2c, d to be equivalent, the following relationship must hold:

$$(C_{MS})_j = (C_{0D})_j - (C_{MS-a-WK})_j \tag{4.15}$$

Hence, to determine $(C_{MS})_j$, $(C_{MS-a-WK})_j$ must be computed. $(C_{MS-a-WK})_j$ represents the assumed additional compliance introduced by the multiscale model at outlet j and ported inside the three-element windkessel model. Three different approaches were tested for computing $(C_{MS-a-WK})_j$.

Approach 1—Direct compliance compensation. The influence of the proximal resistance of the windkessel model is neglected:

$$(C_{MS-a-WK})_j = (C_{MS-a})_j \tag{4.16}$$

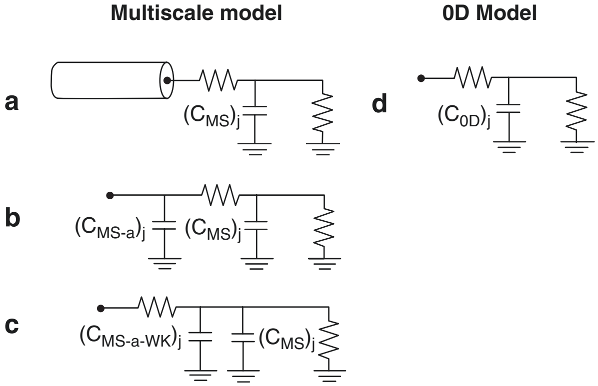


Fig. 4.2 (**a**) Classic representation of the multiscale model; (**b**) the compliance of the highest order model is lumped into a 0D compliance at outlet j placed before the windkessel model; (**c**) the compliance of the highest order model is ported inside the windkessel element; (**d**) classic representation of the windkessel model

Approach 2—Analytical compliance compensation. Approach 1 does not take into account the fact that the proximal resistance of the windkessel element diminishes the influence of the compliance. Thus, following a method proposed previously (Grinberg and Karniadakis 2008):

$$\left(C_{MS-a-WK}\right)_j = \left(C_{MS-a}\right)_j \Big/ \left(1-\left(R_{p-MS}\right)_j / \left(R_{t-MS}\right)_j\right), \tag{4.17}$$

where $(R_{p-MS})_j$ is the proximal resistance of the windkessel model at outlet j of the highest order model.

Approach 3—Numerical compliance compensation. Thirdly, we propose a numerical approach based on the pulse pressure method (PPM) (Stergiopulos et al. 1994), which is described in Algorithm 2. PPM has been introduced to estimate the compliance downstream of a location in an arterial tree, at which the time-varying flow rate values and the average and pulse pressure are known. In the following we use a modified PPM, based on a three-element windkessel model (PPM-WK3) which replaces the two-element windkessel model in the original PPM.

Algorithm 2: PPM-Based Compliance Compensation

```
for each terminal vessel $j$
        Run computation with two element WK model
            $((Q_{0D}(t))_j, (R_{t-MS})_j, (C_{MS-a})_j) \rightarrow (PP_{ref})j$
        Initialize PPM-WK3 $\rightarrow (Q_{0D}(t))_j, (R_{p-MS})_j, (R_{d-MS})_j, (C_{MS-a})_j, (PP_{ref})_j$
        Run PPM-WK3 $\rightarrow (C_{MS-a-PPM})_j$
        $(C_{MS-a-WK})_j \leftarrow (C_{MS-a-PPM})_j$
    end (for)
```

First, a two-element windkessel model is used to determine a reference pulse pressure, $(PP_{ref})_j$, using the flow rate profile obtained at outlet j of the highest order model during the last computation performed with the 0D model, $(Q_{0D}(t))_j$, the total resistance at outlet j after resistance compensation, $(R_{t-MS})_j$, and the assumed additional compliance introduced at outlet j, $(C_{MS-a})_j$. Next, PPM-WK3 is initialized and run, and as a result the compliance value, to be subtracted in Eq. (4.15) from the compliance used inside the 0D model, is determined.

Approaches 1–3 are used to estimate only once the assumed additional compliance introduced by the multiscale model at each outlet j of the highest order model. If the compliance of the multiscale model is adapted directly or indirectly, the compliance values of the windkessel elements are recomputed at each calibration iteration performed for the multiscale model. In this case, considering C the total compliance of the multiscale model, this value is first distributed to the outlets of the highest order model, using a relationship similar to Eq. (4.14):

$$\left(C_{MS-t}\right)_j = \left(C \cdot r_j^2\right) \Big/ \sum_{j=1}^{n} r_j^2 , \tag{4.18}$$

where $(C_{MS-t})_j$ represents the total equivalent compliance of the multiscale model at each outlet of the highest order model.

For the recomputation of the windkessel compliance, a relationship similar to Eq. (4.15) is employed:

$$\left(C_{MS}\right)_j = \left(C_{MS-t}\right)_j - \left(C_{MS-a-WK}\right)_j. \tag{4.19}$$

Compliance compensation is always performed after resistance compensation.

4.3 Results

To evaluate the performance of the proposed parameter estimation method, next we present results for a patient-specific aortic coarctation model extracted from MRI images. In this case, the parameter estimation method ensures that the computational setup is personalized, and, consequently, computed pressure and flow values are in close agreement with the clinical measurements.

Results are reported for four different parameter estimation methods: the calibration method for windkessel parameters (WKC) (Spilker and Taylor 2010), the parameter estimation method with resistance compensation and Direct Compliance compensation (DC), the parameter estimation method with resistance compensation and Analytical Compliance compensation (AC), and the parameter estimation method with resistance compensation and Numerical Compliance compensation (NC).

Blood is modeled as an incompressible Newtonian fluid with a density of 1.050 g/cm^3 and a dynamic viscosity of 0.040 dynes/(cm^2 s). The grid size is 0.05 cm, while the time-step (limited by the CFL-condition) is set equal to $2.5e-5$ s.

Several CFD-based methods for non-invasive assessment of trans-stenotic pressure gradient have been proposed recently (Itu et al. 2013; Keshavarz-Motamed et al. 2011; LaDisa et al. 2011; Ismail et al. 2013a, b). For an accurate patient-specific estimation, the CFD-based solution should be in close agreement with the measured pressure and flow-data, a task that requires automatic calibration of the boundary conditions.

The patient-specific anatomical model (CFD Challenge 2013) is composed of the ascending aorta, three supra-aortic branches, aortic arch, coarctation, and the descending aorta (Fig. 4.3a). Figure 4.3b displays the multiscale model corresponding to the CoA patient-specific geometry. A measured flow rate profile is provided at the inlet, while time-averaged flow-splits are provided for each of the outlets of the geometric model. The objective is to compute the pressure drop across the coarctation, under the constraint that the CFD-based solutions should (1) maintain the same flow-split at each outlet as with the measured data, and (2) replicate the measured systolic and diastolic pressure in the aorta. Figure 4.3c displays the corresponding 0D model.

To build the discretized geometric mesh required for the blood flow computation, we first used the vascular modeling toolkit (VMTK 2014) to extract the centerline

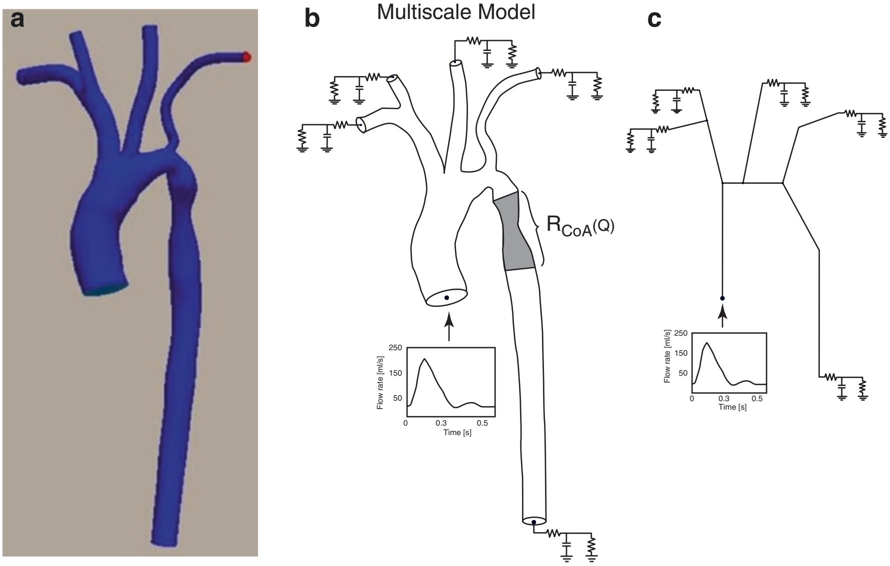


Fig. 4.3 (**a**) Proximal aorta geometry with coarctation, (**b**) Multiscale model used for determining the windkessel parameter values of the patient-specific model, (**c**) 0D model used during the first steps of the model personalization algorithm for finding an initial solution of the parameter values (Itu et al. 2015)

and the cross-sectional areas along the centerline of each arterial segment. Next we used an approach similar to previously introduced ones (Steele et al. 2003), wherein for each vessel of the arterial model, we used several distinct 1D segments with spatially varying cross-sectional area values in order to obtain a geometry close to the 3D geometry acquired through MRI.

Given the high compliance of the ascending aortic wall, this example underlines the advantages of a compliance compensation. At the same time, it also demonstrates the advantages of the resistance compensation for high resistance segments, such as the coarctation.

The parameters to be estimated are the total resistances of the three supra-aortic vessels and of the descending aorta, and the total compliance. The following system of nonlinear equations is numerically solved to obtain the optimum value of each parameter:

$$\mathbf{f}\begin{pmatrix} R_{t-BC} \\ R_{t-LCC} \\ R_{t-LS} \\ R_{t-DAo} \\ C \end{pmatrix} = \begin{Bmatrix} (P_{\max})_{comp} - (P_{\max})_{ref} \\ (P_{\min})_{comp} - (P_{\min})_{ref} \\ (\Phi_{BC})_{comp} - (\Phi_{BC})_{ref} \\ (\Phi_{LCC})_{comp} - (\Phi_{LCC})_{ref} \\ (\Phi_{DAo})_{comp} - (\Phi_{DAo})_{ref} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{4.20}$$

where P_{max} is the maximum (systolic) pressure, P_{min} is the minimum (diastolic) pressure, $(\Phi)_\cdot$ represents a flow rate split, while $(\bullet)_{comp}$ refers to a value computed using the 0D/Multiscale model, and $(\bullet)_{ref}$ refers to the reference value. Index BC refers to the brachiocephalic artery, LCC to the left common carotid artery and DAo to the descending aorta. The reference systolic and diastolic pressures (115 mmHg and 65 mmHg respectively), and the reference flow-rate splits are taken from literature data (CFD Challenge 2013). Only three of the four flow rate splits are used as objectives in Eq. (4.20) since the fourth one is obtained as difference. The characteristics of the pressure waveform are determined at the outlet of the left subclavian artery (Ismail et al. 2013a, b).

One of the estimated total resistances corresponds to several outlets ($R_{t\text{–}BC}$ represents the equivalent total resistance of the right subclavian artery and the right common carotid artery). Hence, before applying Algorithm 1, at the end of the calibration iterations performed for the multiscale model, the total equivalent resistance of the multiscale model at each of the two outlets of the brachiocephalic artery is determined using Eq. (4.12), with a power coefficient equal to two. The proximal resistance of each windkessel model is set equal to the characteristic resistance and was maintained constant throughout the parameter estimation method.

Since total resistance is adapted directly, Algorithm 1 is applied both when switching from the 0D model to the multiscale model, and during each additional calibration iteration performed for the multiscale model. Since compliance is also adapted directly, Eqs. (4.14) and (4.15) are used for the initial compliance compensation, and Eqs. (4.18) and (4.19) are applied at each further calibration iteration.

The objective and parameter values obtained with the four different parameter estimation methods are displayed in Fig. 4.4. Whereas three iterations are required with the WKC method, two iterations are required with the DC and AC methods, and only one iteration is required with the NC method.

With any of the proposed parameter estimation methods, the objective and parameter values are significantly closer to their reference/final values. When the WKC method is used, the pulse pressure at iteration 0 is significantly smaller than the reference pulse pressure, caused by the additional compliance of the highest order model (especially the initial diastolic pressure is considerably closer to the final value). This aspect is also reflected by the fact that the compliance value at iteration 0 is significantly closer to its final value for the proposed methods. Furthermore all initial computed flow rate split values are closer to their final values. This is mainly caused by the significantly improved initial estimate of the total resistance on the descending aorta $R_{t\text{–}DAo}$.

As opposed to the objectives, where the final values are identical for all four parameter estimation methods, the final parameter values are different for the two methods. This is caused by the fact that the herein introduced methods adapt the overall properties of the multiscale model, whereas the WKC method adapts directly the parameters of the three-element windkessel models. Hence, although the parameter values of the two methods at iteration 0 are identical, the values of the windkessel parameters are in fact different. The initial estimate of $R_{t\text{–}DAo}$ for the proposed

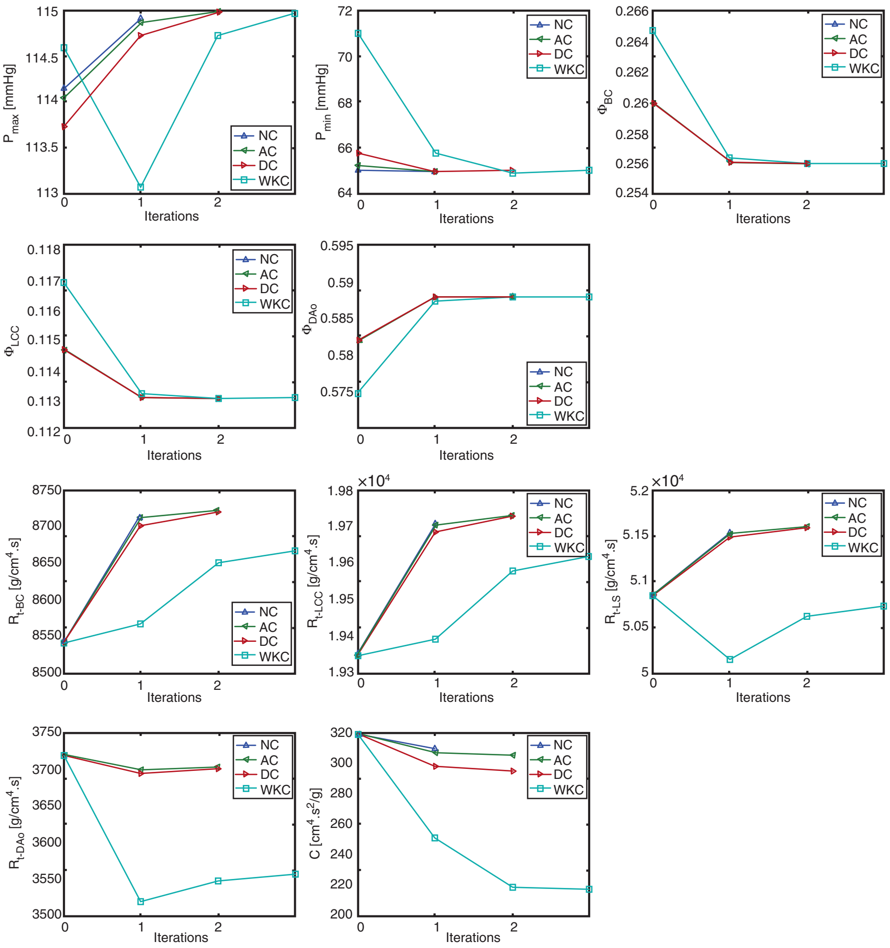


Fig. 4.4 Parameter estimation progression for the proximal aorta model. The total resistance of each of the three supra-aortic branches, the total resistance of the descending aorta windkessel model and the sum of all compliances were the adapted parameters. The desired mean fractions of flow through supra-aortic branches and through the descending aorta were used as objectives, besides systolic and diastolic pressure (Itu et al. 2015)

parameter estimation methods is superior since Eq. (4.11) is employed for computing the resistance of the CoA segment.

Figure 4.5 displays a comparison of the pressure profiles at the outlet of the left subclavian artery, corresponding to the initial solution of the WKC method, the initial solution of the NC method and the final solution (identical for both parameter estimation methods). The initial solution refers to the results obtained at iteration zero with the multiscale model. The initial solution provided by the NC method is clearly superior to the initial solution provided by the WKC method and almost identical to the final solution.

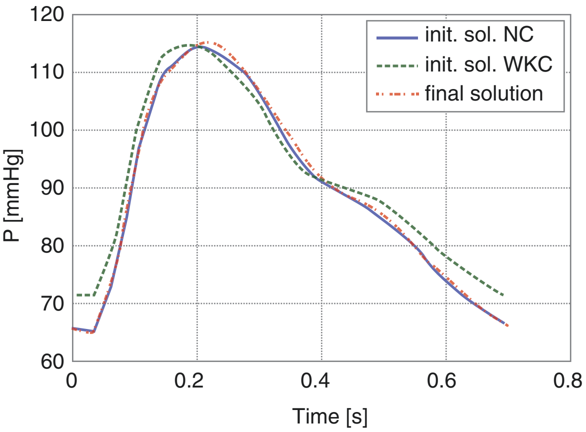


Fig. 4.5 Comparison of the pressure profiles at the outlet of the left subclavian artery of the proximal aorta model, corresponding to the initial solution of the NC method, the initial solution of the WKC method and the final solution (Itu et al. 2015)

Table 4.1 Comparison of reference and computed flow rate split values

Artery	Reference flow split [%]	Computed flow split (NC) [%]
Brachiocephalic artery	25.6	25.6044
Left common carotid artery	11.3	11.3022
Left subclavian artery	4.26	4.3009
Descending aorta	58.8	58.7925

Table 4.2 Final values of the adapted parameters for four different parameter estimation methods applied for the nonlinear system in Eq. (4.20)

Parameter estimation method	$(R_{\text{t-BC}})_{\text{final}}$ [g/(cm^4 s)]	$(R_{\text{t-LCC}})_{\text{final}}$ [g/(cm^4 s)]	$(R_{\text{t-LS}})_{\text{final}}$ [g/(cm^4 s)]	$(R_{\text{t-DAo}})_{\text{final}}$ [g/(cm^4 s)]	$(C)_{\text{final}}$ [10^{-6} cm^4 s^2/g]
NC	8711.3	19,710.6	51,541.6	3698.0	308.71
AC	8720.0	19,730.8	51,595.4	3700.7	304.63
DC	8718,8	19,728.2	51,587.6	3700.0	294.18
WKC	8665.7	19,615.9	50,715.9	3555.4	216.14

Table 4.1 compares the reference and the computed flow rate split values (obtained with the NC method): the values match almost perfectly. Tables 4.2 and 4.3 display for each parameter estimation method the final values of the adapted parameters and the compensated multiscale compliance and CoA resistance values respectively. The compensated compliance is highest for the NC method. Importantly, the sum of the five compensated compliance values reflects the difference between the final total compliance values obtained with the proposed parameter estimation methods and with the WKC method. Furthermore, the compensated CoA resistance (R_{CoA}) roughly reflects the difference between the final total resistance value on the descending aorta ($R_{t\text{-}DAo}$) obtained with the proposed methods and with the WKC method.

Table 4.3 Compensated CoA resistance and compliance values for four different parameter estimation methods applied for the nonlinear system in Eq. (4.20)

Parameter estimation method	R_{CoA} [g/(cm^4 s)]	$C_{MS\text{-}a\text{-}WK}$ [10^{-6} cm^4 s^2/g]				
		RS	RCC	LCC	LS	DAo
NC	125.17	18.15	9.679	13.23	5.830	45.68
AC	125.17	17.45	0.245	12.77	5.442	44.37
DC	125.17	15.44	8.097	11.48	4.700	38.85
WKC	0.0	0.0	0.0	0.0	0.0	0.0

RS right subclavian artery, *RCC* right common carotid artery, *LCC* left common carotid artery, *LS* left subclavian artery

The computed average pressure drop across the coarctation is 4.12 mmHg, while the peak pressure difference is of 16.88 mmHg, both in the range of the values reported for the challenge (Ismail et al. 2013a, b).

4.4 Discussion

The results presented in the previous section demonstrate the potential advantages of the model personalization algorithm, which leads to superior calibration results for the considered patient-specific CoA geometry. The DC and AC methods are easier to implement and computationally faster due to their analytical nature. Given the high computational cost associated with multiscale models, the additional execution time for the NC method, compared to the DC and AC methods, is negligible. Since the NC method has the best performance, we consider this variant of the proposed parameter estimation method to be the most efficient. Note that the configurations of the 0D model and the multiscale model are not affected by the additional steps in the proposed parameter estimation method, and the framework developed for the WKC method can be completely reused.

The geometrical multiscale framework for the aortic model is useful for modeling multiple pathologies. Recently, we validated (with in-vivo measurements) a reduced-order multiscale model for the non-invasive assessment of aortic coarctation (Itu et al. 2013). A fully automated parameter estimation method was proposed for personalizing the hemodynamic computations. The goals of this estimation method were to match the patient-specific average aortic pressure, the flow distribution between supra-aortic arteries and the descending aorta, and to determine the aortic wall properties. Promising validation results were obtained for the non-invasive computation of the trans-coarctation peak-to-peak pressure-gradient.

The configuration used in the previous section for the CoA geometry provides the possibility to enhance the personalization of the hemodynamic computations by matching not necessarily the mean arterial pressure, computed from the systolic and diastolic pressure, but directly the two cuff-based pressures. Matching the flow rate split for the descending aorta is crucial since the trans-coarctation pressure-drop is highly dependent on the flow rate through the descending aorta. Previously only the average flow rate was considered, but, using an automated parameter estimation method, the maximum and/or minimum flow rate can also be constrained.

The proposed parameter estimation method could potentially provide significant advantages when the multiscale model is an FSI model (when rigid wall models are used, only the resistance compensation is required). For the non-invasive assessment of aortic coarctation FSI modeling is required: since the trans-coarctation pressure gradient is computed as a peak-to-peak pressure difference between the ascending aorta and the descending aorta, the pressure drop is not mainly determined by the maximum flow rate and the geometry, but by the complex interaction between these two aspects, the phase lag introduced by the compliance, the wave propagation speed, and the backward travelling pressure and flow rate waves. Furthermore, the wall properties can also be used as adapted parameters in order to match a specific feature of the wave propagation aspects.

Parameter estimation methods for model personalization are not only useful for performing patient-specific computations using measured data, but also to perform computations for other conditions, like exercise, drug-induced hyperemia, or post surgical intervention. Patient-specific computations for CoA patients corresponding to the exercise state were reported previously (LaDisa et al. 2011), where time-varying and average trans-coarctation pressure gradient, wall shear stress and oscillatory shear index parameters were analyzed. Parameter estimation methods may provide the possibility to enhance the value of such (predictive) computations by imposing different changes for the reference values, compared to the rest state. For example, different elevations of mean or systolic aortic pressure, or of flow rate through the descending aorta could be imposed. After further evaluation and clinical research, such an approach could have the potential to eliminate the risks and the costs involved in measuring these quantities in-vivo. Regarding treatment planning evaluation, parameter estimation methods could be useful for imposing a certain post-interventional state for the hemodynamic computation, which could then enable an accurate assessment of the treatment options.

References

Alastruey J, Khir A, Matthys K, Segers P, Sherwin S, Verdonck P, Parker K, Peiro J (2011) Pulse wave propagation in a model human arterial network: assessment of 1-D visco-elastic simulations against in vitro measurements. J Biomech 44:2250–2258. doi:10.1016/j.jbiomech.2011.05.041

Bertoglio C, Moireau P, Gerbeau J-F (2012) Sequential parameter estimation for fluid–structure problems: application to hemodynamics. Int J Numer Method Biomed Eng 28:434–455. doi:10.1002/cnm.1476

Bessems D (2007) On the propagation of pressure and flow waves through the patient-specific arterial system. PhD Thesis, TU Eindhoven, Netherlands

Blanco PJ, Watanabe SM, Feijo RA (2012) Identification of vascular territory resistances in one-dimensional hemodynamics simulations. J Biomech 45:2066–2073. doi:10.1016/j.jbiomech.2012.06.002

Cebral JR, Mut F, Weir J, Putman CM (2011) Association of hemodynamic characteristics and cerebral aneurysm rupture. Am J Neuroradiol 32:264–270. doi:10.3174/ajnr.A2274

CFD Challenge (2013) Simulation of hemodynamics in a patient-specific aortic coarctation model. http://www.vascularmodel.org/miccai2012/. Accessed 15 June 2013

Formaggia L, Lamponi D, Quarteroni A (2003) One dimensional models for blood flow in arteries. J Eng Math 47:251–276. doi:10.1023/B:ENGI.0000007980.01347.29

Grinberg L, Karniadakis GE (2008) Outflow boundary conditions for arterial networks with multiple outlets. Ann Biomed Eng 36:1496–1514. doi:10.1007/s10439-008-9527-7

Haggerty CM, Kanter KR, Restrepo M, de Zelicourt DA, Parks WJ, Rossignac J, Fogel MA, Yoganathan AP (2013) Simulating hemodynamics of the Fontan y-graft based on patient-specific in vivo connections. J Thorac Cardiovasc Surg 145:663–670. doi:10.1016/j.jtcvs.2012.03.076

Hom JJ, Ordovas K, Reddy GP (2008) Velocity-encoded cine MR imaging in aortic coarctation: functional assessment of hemodynamic events. Radiographics 28:407–416

Huo Y, Svendsen M, Choy JS, Zhang ZD, Kassab GS (2012) A validated predictive model of coronary fractional flow reserve. J R Soc Interface 9:1325–1338. doi:10.1098/rsif.2011.0605

Ismail M, Gee MW, Wall WA (2013a) CFD challenge: hemodynamic simulation of a patient-specific aortic coarctation model with adjoint-based calibrated windkessel elements. Lect Notes Comput Sci 7746:44–52. doi:10.1007/978-3-642-36961-2_6

Ismail M, Wall WA, Gee MW (2013b) Adjoint-based inverse analysis of windkessel parameters for patient-specific vascular models. J Comput Phys 244:113–130. doi:10.1016/j.jcp.2012.10.028

Itu L, Sharma P, Mihalef V, Kamen A, Suciu C, Comaniciu D (2012) A patient-specific reduced-order model for coronary circulation. In: Proceedings of the international symposium on biomedical imaging, Barcelona, Spain, pp 832–835. doi: 10.1109/ISBI.2012.6235677

Itu L, Sharma P, Ralovich K, Mihalef V, Ionasec R, Everett A, Ringel R, Kamen A, Comaniciu D (2013) Non-invasive hemodynamic assessment of aortic coarctation: validation with in vivo measurements. Ann Biomed Eng 41:669–681. doi:10.1007/s10439-012-0715-0

Itu L, Sharma P, Passerini T, Kamen A, Suciu C, Comaniciu D (2015) A parameter estimation framework for patient-specific hemodynamic computations. J Comput Phys 281:316–333

Keshavarz-Motamed Z, Garcia J, Pibarot P, Larose E, Kadem L (2011) Modeling the impact of concomitant aortic stenosis and coarctation of the aorta on left ventricular workload. J Biomech 44:2817–2825. doi:10.1016/j.jbiomech.2011.08.00

LaDisa JFJ, Figueroa CA, Vignon-Clementel IE, Kim HJ, Xiao N, Ellwein LM, Chan FP, Feinstein JA, Taylor CA (2011) Computational simulations for aortic coarctation: representative results from a sampling of patients. J Biomech Eng 133:091008. doi:10.1115/1.4004996

Low K, van Loon R, Sazonov I, Bevan RLT, Nithiarasu P (2012) An improved baseline model for a human arterial network to study the impact of aneurysms on pressure-flow waveforms. Int J Numer Method Biomed Eng 28:1224–1246. doi:10.1002/cnm.2533

Malossi C, Blanco P, Deparis S (2012) A two-level time step technique for the partitioned solution of one-dimensional arterial networks. Comput Methods Appl Mech Eng 237:212–226. doi:10.1016/j.cma.2012.05.017

Moireau P, Bertoglio C, Xiao N, Figueroa CA, Taylor CA, Chapelle D, Gerbeau JF (2013) Sequential identification of boundary support parameters in a fluid-structure vascular model using patient image data. Biomech Model Mechanobiol 12:475–496. doi:10.1007/s10237-012-0418-3

Murray CD (1926) The physiological principle of minimum work: I. The vascular system and the cost of blood volume. Proc Natl Acad Sci U S A 12:207–214

Mynard JP, Nithiarasu P (2008) A 1D arterial blood flow model incorporating ventricular pressure, aortic valve and regional coronary flow using the locally conservative Galerkin (LCG) method. Int J Numer Method Biomed Eng 24:367–417. doi:10.1002/cnm.1117

Mynard JP, Davidson MR, Penny DJ, Smolich JJ (2012) A simple, versatile valve model for use in lumped parameter and one-dimensional cardiovascular models. Int J Numer Method Biomed Eng 28:626–641. doi:10.1002/cnm.1466

Olufsen M, Peskin C, Kim WY, Pedersen EM, Nadim A, Larsen J (2000) Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions. Ann Biomed Eng 28:1281–1299. doi:10.1114/1.1326031

Olufsen MS, Nadim A, Lipsitz LA (2002) Dynamics of cerebral blood flow regulation explained using a lumped parameter model. Am J Physiol Regul Integr Comp Physiol 282:611–622. doi:10.1152/ajpregu.00285.2001

Pant S, Fabrèges B, Gerbeau J-F, Vignon-Clementel IE (2014) A multiscale filtering-based parameter estimation method for patient-specific coarctation simulations in rest and exercise, Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. Lect Notes Comput Sci 8330:102–109. doi:10.1007/978-3-642-54268-8_12

Quarteroni A, Tuveri M, Veneziani A (2000) Computational vascular fluid dynamics: problems, models and methods. Comput Vis Sci 2:163–197. doi:10.1007/s007910050039

Raghu R, Vignon-Clementel I, Figueroa CA, Taylor CA (2011) Comparative study of viscoelastic arterial wall models in nonlinear one-dimensional finite element simulations of blood flow. J Biomech Eng 133:081003. doi:10.1115/1.4004532

Reymond P, Bohraus Y, Perren F, Lazeyras F, Stergiopulos N (2011) Validation of a patient-specific one-dimensional model of the systemic arterial tree. Am J Physiol Heart Circ Physiol 301:1173–1182. doi:10.1152/ajpheart.00821.2010

Ringel RE, Jenkins K 2007 Coarctation of the aorta stent trial (coast). http://clinicaltrials.gov/ct2/show/NCT00552812. Accessed 10 Mar 2012

Seifert BL, DesRochers K, Ta M, Giraud G, Zarandi M, Gharib M, Sahn DJ (1999) Accuracy of Doppler methods for estimating peak-to-peak and peak instantaneous gradients across coarctation of the aorta: an in vitro study. J Am Soc Echocardiogr 12:744–753

Spilker R, Taylor CA (2010) Tuning multidomain hemodynamic simulations to match physiological measurements. Ann Biomed Eng 38:2635–2648. doi:10.1007/s10439-010-0011-9

Steele BN (2007) Using one-dimensional finite element analysis to estimate differential pressure of renal artery stenoses. In: Proceedings of computer in cardiology, Durham, NC, USA, pp 381–384. doi: 10.1109/CIC.2007.4745504

Steele BN, Wan J, Ku JP, Hughes TJR, Taylor CA (2003) In vivo validation of a one-dimensional finite-element method for predicting blood flow in cardiovascular bypass grafts. IEEE Trans Biomed Eng 50:649–656. doi:10.1109/TBME.2003.812201

Stergiopulos N, Young DF, Rogge TR (1992) Computer simulation of arterial flow with applications to arterial and aortic stenosis. J Biomech 25:1477–1488. doi:10.1016/0021-9290(92)90060-E.

Stergiopulos N, Meister JJ, Westerhof N (1994) Simple and accurate way for estimating total and segmental arterial compliance: the pulse pressure method. Ann Biomed Eng 22:392–397. doi:10.1007/BF02368245

Taylor CA, Steinman DA (2010) Image-based modeling of blood flow and vessel wall dynamics: applications, methods and future directions. Ann Biomed Eng 38:1188–1203. doi:10.1007/s10439-010-9901-0

The Vasular Modeling Toolkit (2014) http://www.vmtk.org/. Accessed 1 May 2014

Westerhof N, Lankhaar JW, Westerhof BE (2009) The arterial windkessel. Med Biol Eng Comput 47:131–141. doi:10.1007/s11517-008-0359-2

Womersley JR (1955) Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. J Physiol 127:553–563

Young DF, Cholvin NR, Roth AC (1975) Pressure drop across artificially induced stenoses in the femoral arteries of dogs. Circ Res 36:735–743

Zamir M, Sinclair P, Wonnacott TH (1992) Relation between diameter and flow in major branches of the arch of the aorta. J Biomech 25:1303–1310. doi:10.1016/0021-9290(92)90285-9

Lumped Parameter Whole Body Circulation Modelling

5

Viorel Mihalef, Lucian Itu, Tommaso Mansi,
and Puneet Sharma

Abstract

In this chapter we introduce methodologies for modeling whole-body cardiovascular dynamics. Lumped parameter modeling techniques are employed to model both open-loop and closed-loop dynamics. The main constituents of the model are the pulmonary arterial and venous circulation, the systemic arterial and venous circulation, and the four chambers of the heart. A fully automated parameter estimation framework is introduced, which is based on two sequential steps: first, a series of parameters are computed directly, and, next, a fully automatic optimization-based calibration method is employed to iteratively estimate the values of the remaining parameters. A detailed sensitivity analysis has been performed for identifying the parameters which require calibration. Advanced objectives defined based on slopes and interval of times determined from the measured volume and pressure curves are formulated to improve the overall agreement between computed and measured quantities. Furthermore, methods for modeling subtle influences, e.g. from the KG diaphragm, and pathologic heart valves (stenosed, regurgitant) are introduced.

Parts of Sect. 5.2 have been published before in the paper ‘Model based non-invasive estimation of PV loop from echocardiography’, 36th Annual Inter. Conf. of the IEEE Engineering in Medicine & Biology Society—EMBC 2014, Chicago, USA, August 26–30, pp. 6774–6777, 2014.

V. Mihalef (✉) • T. Mansi • P. Sharma
Medical Imaging Technologies, Siemens Healthcare, 755 College Road,
Princeton, NJ 08540, USA
e-mail: viorel.mihalef@siemens.com

L. Itu
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

Automation and Information Technology, Transilvania University of Brasov,
Mihai Viteazu nr. 5, Brasov 5000174, Romania

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_5

The methodology has been validated both for healthy volunteers and for a patient with mild aortic valve regurgitation: a close-agreement between the computed and measured time-varying LV volumes, time-varying LV and aortic pressures, and PV loops has been obtained. This feature is based on research, and is not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

5.1 Introduction

The capacity of the heart to pump sufficient blood to match its own demands and the demands of the body depends on both intrinsic and extrinsic factors. The modeling of these factors can lead to better approaches for the evaluation and management of cardiac disease, as well as better patient stratification and therapy planning.

In this chapter we present comprehensive whole body circulation models which are able to simulate physiological and pathophysiological characteristics, and quantify the cardiac workload from those characteristics. These methods enable a better understanding of the complex relationship between heart disease and the extra workload on the heart due to various pathologies such as hypertrophy, cardiomyopathy (arrhythmogenic right ventricular cardiomyopathy, isolated ventricular non-compaction, mitochondrial myopathy, dilated cardiomyopathy, restrictive cardiomyopathy, peripartum cardiomyopathy, takotsubo cardiomyopathy, loeffler endocarditis, etc.), mitral regurgitation, aortic stenosis, aortic regurgitation, and hypertension.

Due to the prohibitive computational cost of spatially defined blood flow models (three-dimensional models in particular), whole body circulation models of the cardiovascular system rely heavily on lumped parameter models. These models are based on the analogy between hydraulics and electricity, in the form of RLC circuits (Table 5.1) (Quarteroni et al. 2001).

Closed loop lumped parameter or multiscale models of the cardiovascular system have been introduced in the past (Segers et al. 2003; Korakianitis et al. 2006; Lakin et al. 2007; Kim et al. 2010; Paeme et al. 2011). To model different patient states (steady-state and transient), the cardiovascular model was coupled to a series of models representing the cardiovascular regulatory systems (Fig. 5.1) (Ottesen et al. 2003). The bidirectional exchange of information between the systems leads to a continuous adaptation of the cardiovascular activity and operation.

Table 5.1 Analogy between hydraulics and electricity

Hydraulics	Electricity	
Pressure	Voltage/potential	P
Flow rate	Current	Q
Viscosity	Resistance	R
Inertia	Inductance	L
Compliance	Capacitance	C

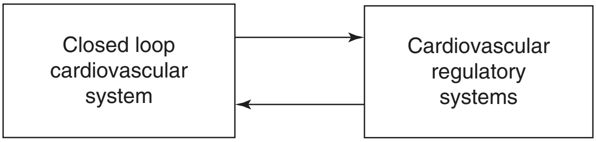


Fig. 5.1 Bidirectional exchange of information between the cardiovascular and the regulatory systems

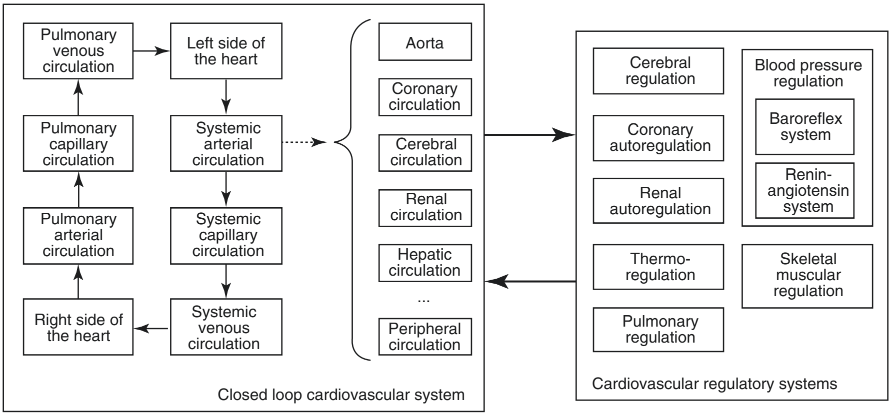


Fig. 5.2 Main components of the cardiovascular system and regulatory systems which act on the cardiovascular system

The closed loop models introduced in the past focus mainly on population average computations in order to quantify the interactions between the different components of the system. No systematic or automatic personalization procedures have been applied. Furthermore, at most one regulatory system was taken into consideration.

In Fig. 5.2 we display a comprehensive model, whereas we have detailed the main components of the systems in Fig. 5.1. The cardiovascular model contains a heart model (left and right side of the heart, each of them with atrium and ventricle), the systemic circulation (arteries, capillaries, veins) and the pulmonary circulation (arteries, capillaries, veins). Figure 5.2 presents a set of possible models for the systemic and pulmonary circulation. Furthermore, various regulatory systems which act on the cardiovascular system are presented in Fig. 5.2 right. The objectives of these systems are to maintain certain levels of blood pressure, flow rate to a certain organ, body temperature, filtration rate, oxygen level in the blood, etc. Specifically, most systems of the body show some degree of autoregulation, the heart and the brain are very sensitive to over- and underperfusion. Coronary autoregulation ensures that the coronary blood supply matches the oxygen demand of the myocardium, both at rest and at exercise (hyperemia), by adapting the resistance of the coronary microvasculature. Cerebral autoregulation also focuses on maintaining an appropriate blood flow to the subtended cerebral tissue.

Furthermore, blood pressure regulation at the systemic level is performed by the baroreflex system, which uses input data provided by the baroreceptors situated mainly in the aortic arch and at the carotid sinuses, and by the renin-angiotensin system which is triggered by pressure and flow receptors in the afferent arterioles of the renal arterial circulation. The renal autoregulation system adapts the resistance of the renal microvasculature in order to maintain the reference glomerular filtration rate.

In the following we first present an open-loop system focusing on the left side of the heart and the systemic circulation, followed by a closed-loop whole body-circulation model.

The concepts and information presented in this chapter are based on research and are not commercially available. Due to regulatory reasons their future availability cannot be guaranteed.

5.1.1 Short-Term Control Mechanisms of the Human Circulatory System

The short term control is mainly performed by the central nervous system (CNS) and includes baroreceptors, mechanoreceptors and chemoreceptors (Guyton 1991). Its goal is to distribute the total flow to the various parts of the body by controlling the heart (heart rate and cardiac contractility) and the vessel constrictions and relaxations. The chemoreceptors react to changes in the concentration of different gases (oxygen, etc.). The baroreceptors react to changes in pressure (the most important receptors of this type are located at the aortic arch and at the carotid sinus). Physiological and historical details regarding the baroreceptor model can be found in Guyton (1991) and Acierno (1994). Mechanoreceptors are mostly found in the atria and pulmonary veins, which are low pressure areas. Out of these three types of receptors, the baroreceptors are the most easily accessible receptors, while the mechanoreceptors are less studied. The baroreceptors play the main role regarding the short term pressure control and have no role in the long term regulation functions.

Two different baroreceptor models have been introduced in Ottesen et al. (2004), which are composed of three parts:

- The afferent part: takes as input the average arterial pressure and outputs the firing rates n;
- The CNS which takes as input the firing rates and generates the sympathetic and parasympathetic nervous activities (n_s and n_p);
- The efferent part which takes as input the two nervous activities and generates the new values for the controlled parameters.

Based on these three components, a simple and a unified model baroreceptor model are introduced. For the simple model, the sympathetic and parasympathetic activities are computed directly from the mean arterial pressure, by assuming a sigmoidal function:

$$n_s\left(\overline{p}_{cs}\right)=\frac{1}{1+\left(\frac{\overline{p}_{cs}}{\mu}\right)^{\nu}}, \tag{5.1}$$

$$n_p\left(\overline{p}_{cs}\right)=\frac{1}{1+\left(\frac{\overline{p}_{cs}}{\mu}\right)^{-\nu}}, \tag{5.2}$$

where $\overline{p}_{cs}$ is the measured average pressure, μ is the steady state arterial pressure at the baroreceptors and ν determines the steepness of the sigmoidal curves.

The efferent part is similar for both baroreceptor models and consists of a static and a dynamic component (first order ordinary differential equation). Thus, the generic model, which was adopted, is:

$$\frac{dx_i(t)}{dt}=\frac{1}{\tau_i}\left(-x_i(t)+\alpha_i n_s\left(\overline{p}_{cs}\right)-\beta_i n_p\left(\overline{p}_{cs}\right)+\gamma_i\right) \tag{5.3}$$

where $x_i(t)=\{HR, E_{\max}, R_{ps}, V_{un}, C_v\}$, τ_i represents the transition time until the efferent response i takes effect, α_i and β_i represent the strength of the sympathetic and parasympathetic activity on the response i, whereas γ_i is equal to x_i when no nervous activity is considered. The five controlled quantities are: HR—the heart rate, E_{max}—the maximum contractility of the heart, R_{ps}—the peripheral systemic resistance, V_{un}—the venous unstressed volume and C_v—the venous compliance.

For the unified baroreceptor model described in Ottesen et al. (2004), the firing rates n are computed separately and all three parts of the model are treated separately. Several models have been proposed, which model some or all of the nonlinear behavioral aspects of the firing rates (Leaning et al. 1983; Ursino 1999; Ursino 2000; Taher et al. 1988; Ottesen 1997).

When analyzing Eq. (5.3) from the point of view of control theory, it follows that it represents a controller of type PT1 (proportional controller with first order phase lag). Hence, when the hemodynamic conditions change and as a consequence, the average pressure changes, the controllers will counteract this change, but since no integrative component is present in the controller, an error between the reference pressure μ and the measured pressure will appear once the system is perturbed.

The baroreceptor model has been coupled to multiscale model of the human circulation in Kim et al. (2010) and the results have shown that indeed, once a perturbation appears, the average arterial pressure does not reach its previous value.

In a different study (Coogan et al. 2011) a patient-specific simulation, performed for the evaluation of aortic coarctation, was described. In order to match the patient-state as close as possible, the mean average pressure was tuned so as to be identical to the non-invasively measured one. A simple heart model, based on varying elastance, was coupled at the inlet of the aorta and the average aortic pressure was controlled through the left atrial pressure (only the arterial part was modeled).

5.1.2 Coronary Autoregulation

Autoregulation is control mechanism which acts locally and which is independent from the central nervous system. Thus, dilation and contraction of vessels may appear as a consequence of moderate changes in flow rate or pressure (Guyton 1991; Tortora and Anagnostakos 1990). Autoregulation is mainly performed through smooth muscles which surround the vessels (Ganong 1975). In the following we will refer specifically to the coronary autoregulation, since the focus of the current work lies on the simulation and control of the coronary circulation.

Coronary autoregulation plays an important role in stenosed coronary tree at rest state. Its goal is to maintain a certain level of blood flow, given by the myocardial oxygen demand, to the microvascular beds perfused by arteries which contain stenosis. Through autoregulation, the microvascular resistances are decreased so as to eliminate the effect of the stenosis which introduce additional resistances along the epicardial vessels. During hyperemia, the microvascular resistances are already minimal and cannot be decreased further through autoregulation.

One of the most comprehensive studies regarding coronary flow has been described in Miyashiro and Feigl (1995), where both feedforward and feedback mechanism have been used to model coronary autoregulation. Thus α-receptor mediated vasoconstriction and β-receptor mediated vasodilation are modeled through feedforward loops. The feedforward vasodilation generated by the β-receptors appears in order to meet the oxygen demand, but also a feedback loop is necessary in order to obtain a high level of accuracy for the control mechanism. The feedback loop uses as error signal the coronary venous partial pressure of CO_2 and the coronary venous partial pressure of O_2.

Different experiments have shown that when only the feedback loop is active, the delay of the regulation is higher and overshoot appears, while when the feedforward mechanisms are activated the response is much faster and no overshoot is present. Also, it has been shown the α-receptor mediated vasoconstriction increases the errors and decreases the accuracy of the flow response.

The detailed mechanism which determines a good match between myocardial oxygen consumption and coronary blood flow is not completely known. Several other models have been proposed which use the partial venous pressure of O_2 as the error signal for the metabolic feedback loop (Dankelmann et al. 1992; Drake-Holland et al. 1984). Additionally, other factors intervene like flow-dependent coronary vasodilation induced by endothelium-derived factor (which is a positive feedback flow mechanism) or the myogenic response.

5.2 Non-invasive Estimation of the Left Ventricular Pressure-Volume Loop

The left ventricular pressure-volume (PV) loop represents an efficient tool for understanding and characterizing cardiac function. It contains information regarding stroke volume, cardiac output, ejection fraction, myocardial contractility, cardiac oxygen consumption, and other important measures of the heart and the systemic circulation. For example, the extent of ventricular remodeling, the degree of ventricular-arterial mismatching (Burkhoff 2013), and the left ventricular end-diastolic pressure-volume relationship (Spevack et al. 2013) represent strong predictors of congestive heart failure. Pathologies such as left ventricular hypertrophy, dilated cardiomyopathy, aortic and mitral valve stenosis, and regurgitation (Hall 2011) are manifested in the PV-loop. Hence, a method for an efficient estimation of the PV loop would represent a powerful diagnostic tool for clinicians. Medical imaging modalities such as MRI, CT, and echocardiography can be used to estimate the time-varying LV volume through the heart cycle in a non-invasive manner, which can then be combined with an invasive measurement of LV pressure to obtain the PV loop (van Slochteren et al. 2012).

In the following, we introduce a model-based approach for the non-invasive estimation of left ventricular, patient-specific PV loops: a lumped parameter circulation model is personalized using a two step parameter estimation framework (Itu et al. 2014). The input data required for the model personalization are acquired through routine non-invasive clinical measurements and echocardiography.

5.2.1 Lumped Parameter Model

The lumped parameter circulation model employed herein is displayed in Fig. 5.3. It comprises three main components: venous pulmonary circulation, the left heart and the systemic circulation. For the venous part of the pulmonary circulation, we use a model composed of a resistance (R_{pulVen}) and compliance (C_{pulVen}):

$$C_{pulVen}\frac{dP_{LA}}{\mathrm{dt}} = Q_{C-pulVen}, \tag{5.4}$$

$$Q_{pulVen} = Q_{C-pulVen} + Q_{LA-in}, \tag{5.5}$$

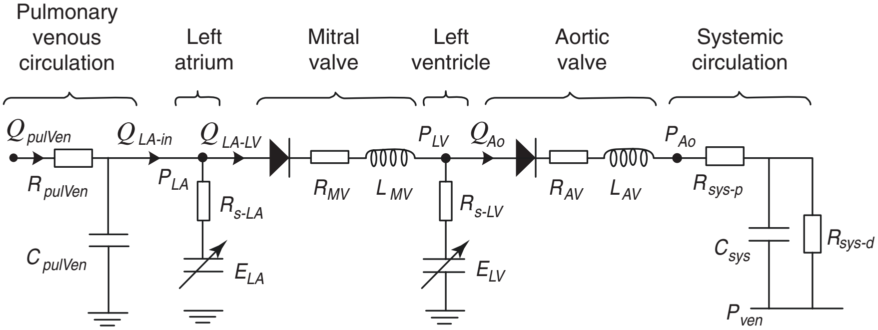


Fig. 5.3 Lumped parameter model representing the venous pulmonary circulation, the left heart and the systemic circulation (Itu et al. 2014)

where the venous pulmonary flow rate (Q_{pulVen}) is considered to be constant in time.

The left heart model has four components: left atrium (LA), mitral valve, left ventricle (LV) and aortic valve. We use a time-varying elastance model for the LA and the LV (Suga 1971):

$$P(t) = E(t) \cdot (V(t) - V_0) - R_s Q(t) \tag{5.6}$$

where E is the time-varying elastance, V is the cavity volume, V_0 is the dead volume of the cavity, and Rs is a source resistance which accounts for the dependence between the flow and the cavity pressure (Shroff et al. 1985) ($R_s = K_s E(t)$ $(V(t) - V_0(t))$, K_s—constant). The cavity volume is equal to:

$$dV / \mathrm{dt} = Q_{in} - Q_{out}, \tag{5.7}$$

where Q_{in} represents the inlet flow rate into the cavity and Q_{out} represents the outlet flow rate from the cavity. The mitral valve and the aortic valve are modeled using a resistance, an inertance and a diode to simulate the closure and the opening of the valve (Mynard et al. 2012). When the valve is open, the following relationship holds:

$$P_{in} - P_{out} = R_{valve} \cdot Q + L_{valve} \cdot dQ / dt, \tag{5.8}$$

where P_{in} and P_{out} represent the pressures at the inlet and respectively the outlet of the valve. When the valve is closed, the flow rate through the valve is set to zero. Each valve opens when P_{in} becomes greater than P_{out}, and closes when the flow rate becomes negative. A three-element Windkessel model is used for the systemic circulation, represented by the following relationship between instantaneous flow and pressure.

$$\frac{dP_{Ao}}{\mathrm{dt}} = R_{sys-p} \frac{dQ_{Ao}}{\mathrm{dt}} - \frac{P_{Ao} - P_{ven}}{R_{sys-d} \cdot C_{sys}} + \frac{Q_{Ao}\left(R_{sys-p} + R_{sys-d}\right)}{R_{sys-d} \cdot C_{sys}}, \tag{5.9}$$

where $R_{sys\text{-}p}$ and $R_{sys\text{-}d}$ are the proximal and distal resistances respectively, C_{sys} is the compliance, and P_{ven} is the venous pressure. A total of nine equations are obtained, which are solved implicitly using the forward Euler time integration scheme.

5.2.2 Parameter Estimation Framework

To compute patient-specific left ventricular PV loops using the lumped parameter model, the parameters of the model are personalized. The model personalization framework consists of two sequential steps. First, a series of parameters are computed directly, and next, a fully automatic optimization-based calibration method is employed to estimate the values of the remaining parameters, ensuring that the

Table 5.2 List of patient-specific input parameters

Input	Source
Systolic blood pressure (SBP)	Cuff measurement (arms)
Diastolic blood pressure (DBP)	Cuff measurement (arms)
Heart rate (HR)	Routine measurement
Ejection fraction (EF)	Echocardiography
End-diastolic volume (EDV)	Echocardiography

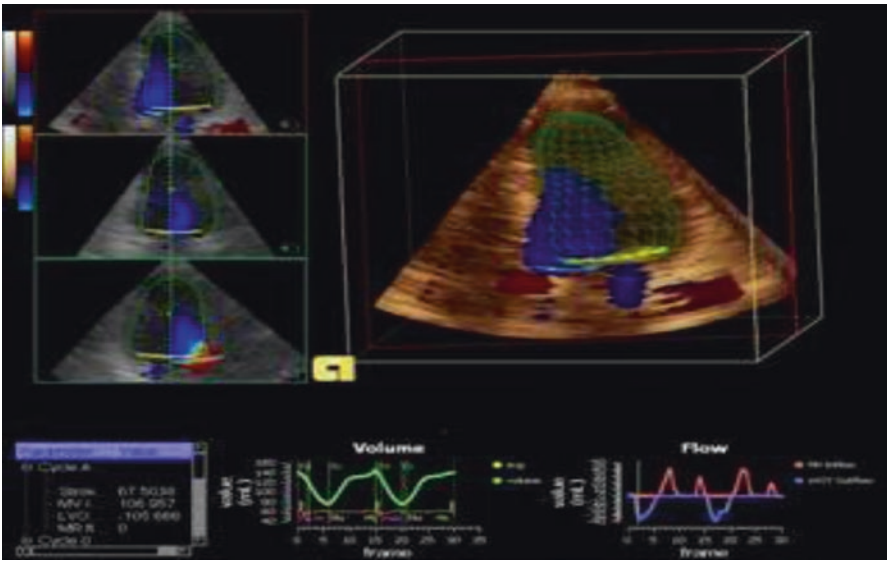


Fig. 5.4 Image acquired through echocardiography illustrating the steps required to extract the end-diastolic volume and the ejection fraction (Itu et al. 2014)

personalized computations match the measurements. Table 5.2 lists the patient-specific input parameters used in the current study, together with their source. Figure 5.4 displays an image acquired through echocardiography, illustrating the steps required for extracting the last two quantities from Table 5.2.

During the first step of the parameter estimation framework, the mean arterial pressure (MAP) is determined:

$$MAP = DBP + \left[1/3 + \left(HR \cdot 0.0012\right)\right] \cdot \left(SBP - DBP\right). \tag{5.10}$$

Then, the end-systolic volume is computed:

$$ESV = EDV \cdot \left(1 - EF\right)/100. \tag{5.11}$$

Next, the stroke volume is determined:

$$SV = EDV - ESV, \tag{5.12}$$

and the average aortic flow rate is computed:

$$\overline{Q_{Ao}} = SV \cdot 60 / HR. \tag{5.13}$$

Finally, the total systemic resistance, as well as the proximal and distal components, are determined:

$$\begin{aligned} R_{sys-t} &= \left(MAP - P_v \right) / \overline{Q_{Ao}}, \\ R_{sys-p} &= \rho \cdot R_{sys-t}; \quad R_{sys-d} = \left(1-\rho\right) \cdot R_{sys-t}, \end{aligned} \tag{5.14}$$

where ρ is the proximal resistance fraction. Since the lumped model is used for a pulsatile steady-state computation, the average inlet flow rate (Q_{pulVen}) is equal to the average outlet flow rate, given by Eq. (5.13). Hence:

$$Q_{pulVen} = \overline{Q_{Ao}}. \tag{5.15}$$

The normalized elastance curve is used for the left ventricle model (Suga 1971), which is denormalized using the minimum and maximum elastance values, and the time at which the maximum elastance is reached. The minimum elastance value is set to 0.08 mmHg/ml, and the time at which the maximum elastance of the left ventricle is reached is computed using $t_{max} = 0.16\,T + 0.17$, where T is the period. The maximum elastance value is estimated as described further down. A two-hill function is used to determine the elastance curve for the left atrium, whereas the minimum elastance is set to 0.08 mmHg/ml, the maximum elastance is set to 0.17 mmHg/ml, and the onset of the contraction is set at 0.85 T (Mynard et al. 2012).

During the second step of the parameter estimation framework, an optimization-based calibration method is employed to estimate the maximum elastance of the left ventricle model, $E_{max\text{-}LV}$, the dead volume of the left ventricle, $V_{0\text{-}LV}$, and the compliance of the systemic Windkessel model, C_{sys}.

The parameter estimation problem is formulated as a numerical optimization problem, the goal of which is to find a set of parameter values for which a set of objectives are met. Since the number of parameters to be estimated is set equal to the number of objectives, the parameter estimation problem becomes a problem of finding the root for a system of nonlinear equations. To solve the system of equations, we use the dogleg trust region method (Nocedal and Wright 2006). The objectives of the parameter estimation method are formulated based on the systolic and diastolic pressures, and the ejection fraction, leading to the system of nonlinear equations:

$$\mathbf{r}\begin{pmatrix} E_{\max-LV} \\ V_{0-LV} \\ C_{sys} \end{pmatrix} = \begin{Bmatrix} \left(SBP\right)_{comp} - \left(SBP\right)_{ref} \\ \left(DBP\right)_{comp} - \left(DBP\right)_{ref} \\ \left(EF\right)_{comp} - \left(EF\right)_{ref} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}, \tag{5.16}$$

where, $\boldsymbol{r}(\boldsymbol{x})$ is a vector function, called in the following objective function, and $\boldsymbol{x}$ is the vector of the unknowns, i.e. the parameters to be estimated. Each component of the objective function is formulated as the difference between the computed value of a quantity—$(\bullet)_{comp}$ (determined using the lumped parameter model) and its reference value—$(\bullet)_{ref}$ (determined through measurement). To evaluate the objective function for a given set of parameter values, the lumped parameter model is run exactly once. The details of the parameter estimation framework have been introduced in Chap. 4.

5.2.3 Results

To evaluate the performance of the proposed methodology for the non-invasive estimation of left ventricular PV loops, next we present results for three healthy volunteers. Systolic and diastolic pressure values were acquired using cuff-based measurements and the ejection fraction and end diastolic volumes were estimated from the echocardiography performed at rest state in a horizontal position using the Siemens ACUSON SC 2000 ultrasound system. The computed time-varying pressure profiles and PV loops are displayed in Fig. 5.5.

To perform an initial validation of the methodology, we computed the PV loop for a patient with mild aortic valve regurgitation and compared the results against the invasively determined quantities. Figure 5.6 displays a comparison between model-based computed results and invasively performed measurements. The input data used for the parameter estimation framework were extracted from the invasive measurements as follows: *SBP* was the maximum aortic pressure (181.5 mmHg), *DBP* was the minimum aortic pressure (89.7 mmHg), *EDV* was the maximum LV volume (196.68 ml), *EF* (53.1%) was computed from EDV and ESV, determined as minimum LV volume (92.26 ml), and *HR* was determined from the period of the time-varying pressure (47 bpm). All these values are matched exactly for the output parameter values: $E_{max\text{-}LV} = 0.968$ mmHg/ml, $V_{0-LV} = -88.71$ ml, $C_{sys} = 1.065 \times 10^{-3}$ cm^4 s^2/g. There is a close agreement between the time-varying LV and aortic pressures, time-varying LV volumes, and PV loops. Moreover, the four phases of the cardiac cycle can be clearly identified in the computed results (Fig. 5.6a, b): (1) isovolumetric contraction phase, (2) ventricular ejection phase, (3) isovolumetric relaxation phase, and (4) ventricular filling phase. The mild aortic valve regurgitation can be observed in the PV loop in Fig. 5.6c, where the line corresponding to the isovolumetric relaxation has a slight curvature, and in Fig. 5.6b, during the second part of phase 3, where the LV volume increases slightly. The average execution time for the four volunteers/patients was of 28.9 s on a standard Intel i7 CPU core with 3.4 GHz.

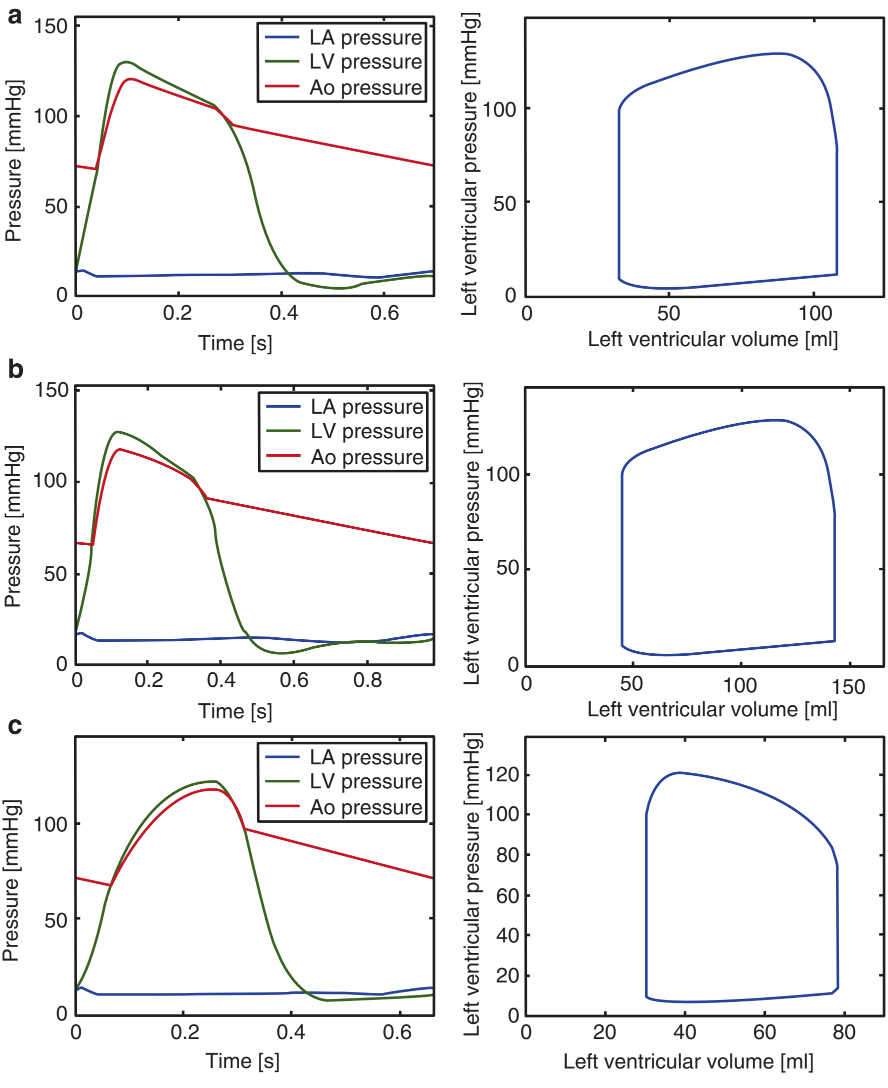


Fig. 5.5 Computed time-varying LA pressure, LV pressure and aortic pressure (*left side*) and PV loops for (**a**) volunteer 1, (**b**) volunteer 2, and (**c**) volunteer 3 (Itu et al. 2014)

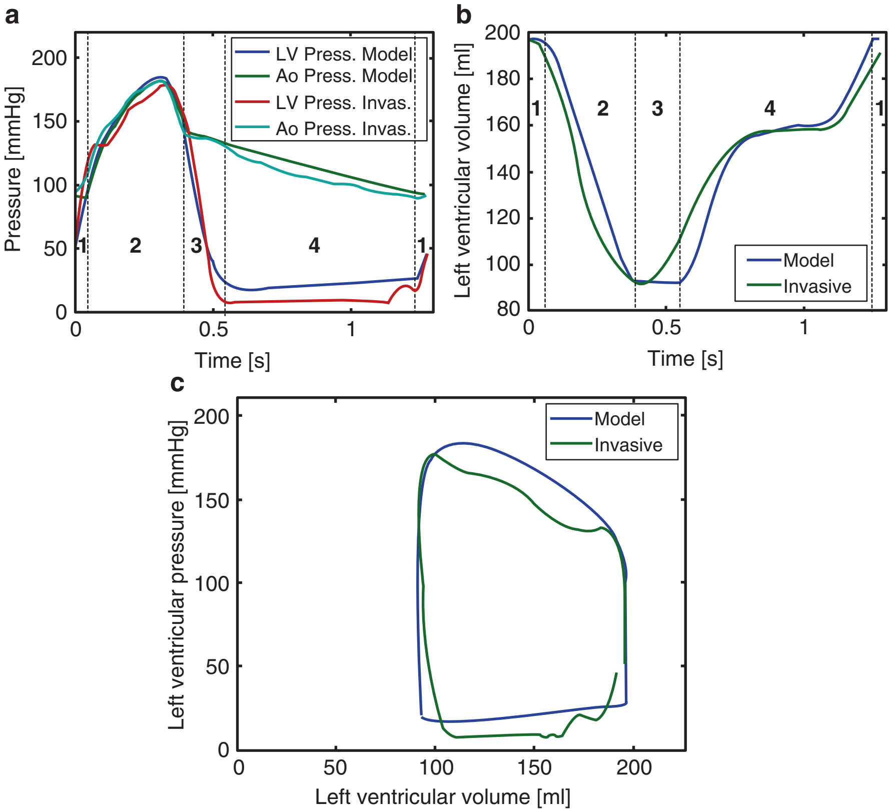


Fig. 5.6 Comparison of model-based computation against invasive measurements, for (**a**) time-varying left ventricular and aortic pressures, (**b**) time-varying left ventricular volume, and (**c**) PV loop (Itu et al. 2014)

5.3 Closed-Loop Whole Body-Circulation Model

The proposed closed-loop whole body-circulation model is displayed in Fig. 5.7.

Time-varying elastance models are used for all four chambers of the heart (Eq. (5.6)):

The models of all four valves (mitral, aortic, tricuspid and pulmonary) of the heart include a resistance, an inertance and a diode (for simulating the opening and the closure of the valve based on the pressure gradient between the two sides of the valve). When the valve is closed the flow across the valve is set to 0. When the valve is open Eq. (5.8) holds:

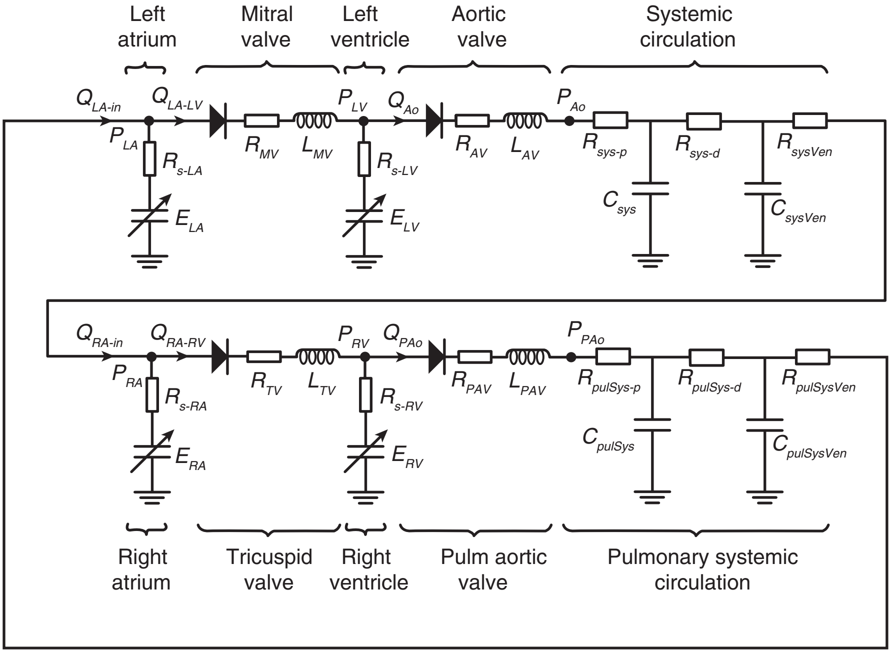


Fig. 5.7 Lumped parameter closed loop model of the cardiovascular system

A three-element Windkessel model is used for the systemic circulation, represented by the following relationship between instantaneous flow and pressure (Eq. (5.9)):

A two-element Windkessel model is used for the systemic venous circulation:

$$\frac{dP_{ven}}{\mathrm{dt}} = \frac{Q_{ven}}{C_{sysVen}} - \frac{P_{ven} - P_{RA}}{R_{sysVen} \cdot C_{sysVen}}. \tag{5.17}$$

Similar models are used in the pulmonary circulation.

The following parameters may be used as input or output parameters for the closed loop (lumped) models of the cardiovascular system:

- Systolic aortic pressure [mmHg]
- Diastolic aortic pressure [mmHg]
- Heart rate [bpm]
- Ejection fraction [%]
- End-diastolic volume [ml]
- Stroke volume [ml]
- Left ventricular end-systolic pressure [mmHg]
- Left ventricular end-systolic elastance [mmHg/ml]
- Arterial compliance
- $V_{0,*}$ [ml]—dead volume of the * chamber of the heart

- V_{100} [ml]—left ventricular volume corresponding to a left ventricular pressure of 100 mmHg
- Proximal arterial resistance [g/(cm^4 s)]
- Distal arterial resistance [g/(cm^4 s)]
- Total arterial resistance [g/(cm^4 s)]
- Stroke work PV [J]: stroke work determined from computed PV loop
- Normalized stroke work PV [J/ml]: stroke work PV divided by stroke volume
- Stroke work PQt [J]: stroke work determined from computed left ventricular pressure and aortic flow rate
- Normalized stroke work PQt [J/ml]: stroke work PQt divided by stroke volume
- Arterial elastance [mmHg/ml]: computed as end systolic pressure divided by stroke volume
- Arterial ventricular coupling: arterial elastance divided by left ventricular end-systolic elastance

To compute patient-specific hemodynamics, the model in Fig. 5.7 requires personalization. Figure 5.8 displays the overall workflow used for fully automatic model personalization. The first step is to define the metrics of interest: the main output(s) of the model. Herein these are:

- Arterial compliance
- $V_{0,*}$ and V_{100}

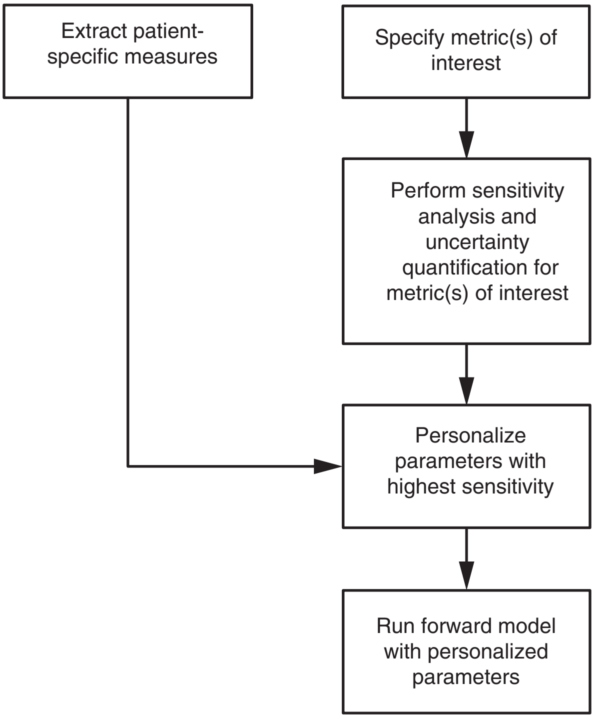


Fig. 5.8 Overall workflow used for personalization

- Proximal, distal and total arterial resistance
- Stroke work
- Arterial elastance
- Arterial ventricular coupling

The next step is to perform a sensitivity analysis to identify the main parameters of the model which affect the outputs of interest. To achieve this, both global sensitivity analysis and uncertainty quantification are performed (based on the stochastic collocation method and polynomial chaos expansion). Once the parameters with the highest influence on the metrics of interest have been identified they are personalized based on patient-specific measures. Herein, these measures are:

- Non-invasive measurements (cuff-based pressures systemic systolic and diastolic pressures, heart rate)
- left and right ventricle volume

Population-average values are used for the parameters which are not personalized. The personalization consists in running a forward model multiple times until the objectives in the model outputs are met.

5.3.1 Local Sensitivity Analysis

The results of a local sensitivity analysis performed with the whole body circulation lumped parameter model are described below. In each test one parameter is modified, while the others are maintained constant. Effects have been analyzed for the LV pressure, LV volume and the systemic arterial pressure.

Figure 5.9 displays the effects of a variation in the initial volume in the LA and the LV on the computed quantities (LV pressure, aortic pressure and LV volume). In open loop models the inlet boundary condition controls the volume of blood in the model, and this value may change from one heart cycle to the next, as the computation progresses from a transient to a steady state. In closed loop models however, the initial volume in the system is maintained from one heart cycle to the next, and hence the initialization of the model with the ‘correct’ amount of blood is crucial.

Figure 5.10 displays the effects of a variation in the maximum elastance of the LV. A lower maximum elastance mainly affects early systole in the LV and aortic pressures, it leads to a slight decrease of the diastolic aortic pressure and to an increase of the LV volume. The lower the maximum elastance, the smaller becomes the stroke volume and, as a result, the levels of the LV volume increases.

Figure 5.11 displays the effects of a variation in the systemic compliance. A lower compliance leads to a higher afterload, and hence to higher LV and aortic pressures, the diastolic aortic pressure decay becomes more prominent (the time constant of the systemic circulation decreases), and the LV volume increases (due to the higher afterload less blood is pumped out of the LV).

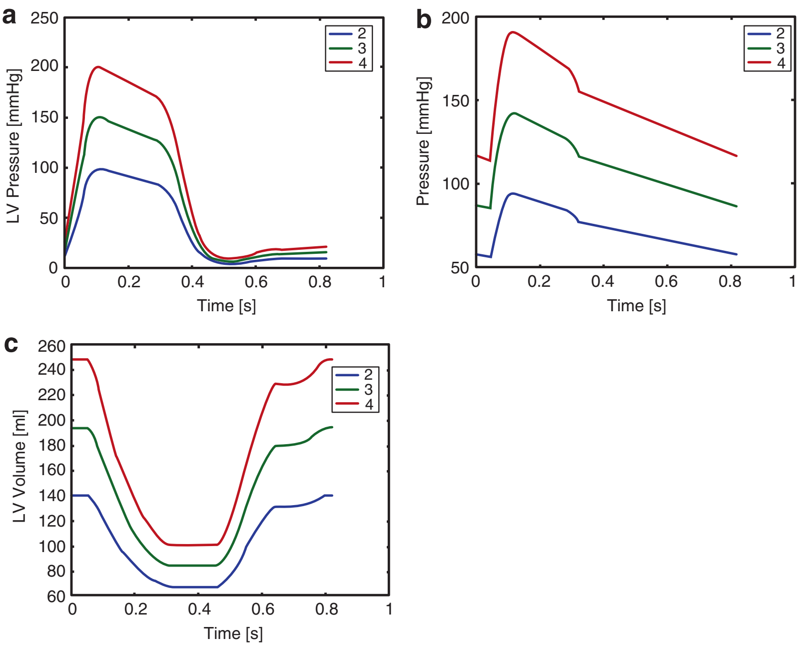


Fig. 5.9 Effect of initial LA and LV volume on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. A manifold of the patient-specific LV end-diastolic volume is used (2, 3 or 4)

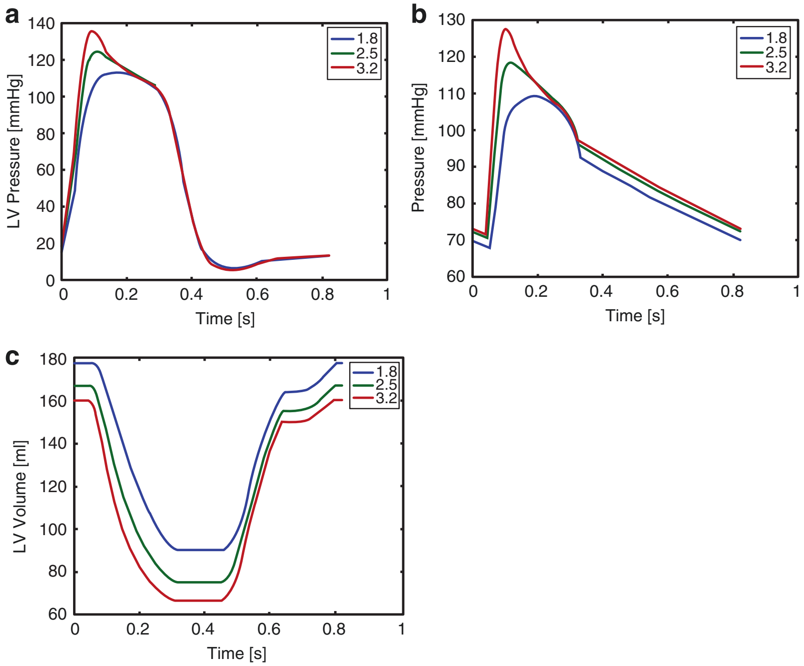


Fig. 5.10 Effect of LV maximum elastance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Elastance is expressed in [mmHg/ml]

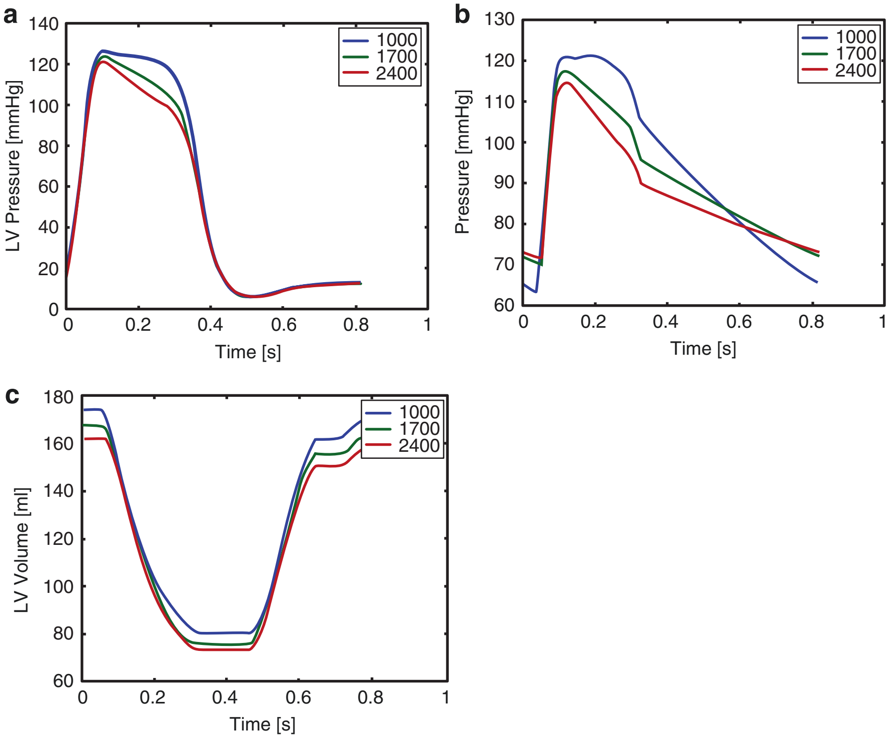


Fig. 5.11 Effect of systemic compliance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Compliance is expressed in [10^{-6} cm^4 s^2/g]

Figure 5.12 displays the effects of a variation in the dead volume of the LV. LV and aortic pressures are largely unaffected, while the LV volume levels are shifted by a value approximately equal to the variation in the dead volume.

Figure 5.13 displays the effects of a variation of the time when LV elastance is maximum: the systolic period (the ejection time) is increased.

Figure 5.14 displays the effects of a variation in the maximum LA elastance: effects are negligible.

Figure 5.15 displays the effects of a variation in the minimum LA elastance: the filling phase of the LV after the atrial kick is slightly affected.

Figure 5.16 displays the effects of a variation in the LA onset of contraction: the filling phase of the LV after atrial the kick is slightly affected: with later onset, the atrial kick becomes less pronounced.

Figure 5.17 displays the effects of a variation in the LA dead volume. With increasing dead volume LV pressure, aortic pressure, and LV volume are slightly shifted downwards.

Figure 5.18 displays the effects of a variation in the systemic venous compliance. Since the systemic venous compliance is very large and stores a large amount of blood, a decrease of this compliance will lead to higher overall levels of pressure and LV volume.

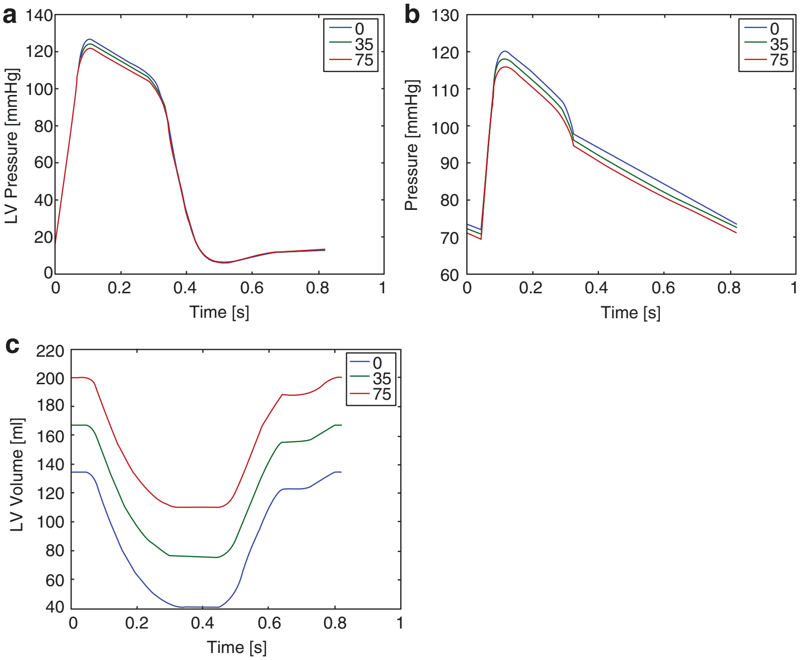


Fig. 5.12 Effect of LV dead volume on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Volume is expressed in [ml]

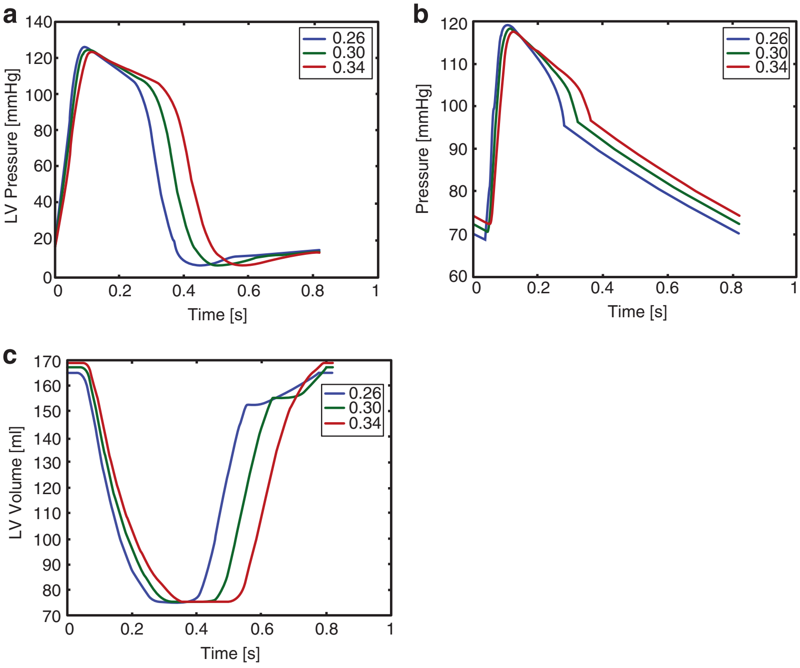


Fig. 5.13 Effect of time when LV elastance is maximum on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Time is expressed in [s]

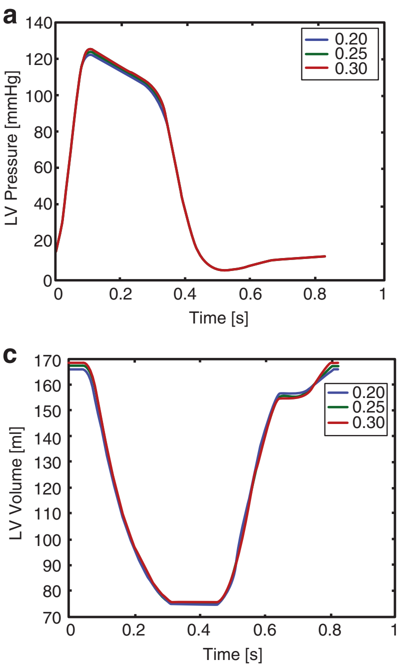


Fig. 5.14 Effect of LA maximum elastance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Elastance is expressed in [mmHg/ml]

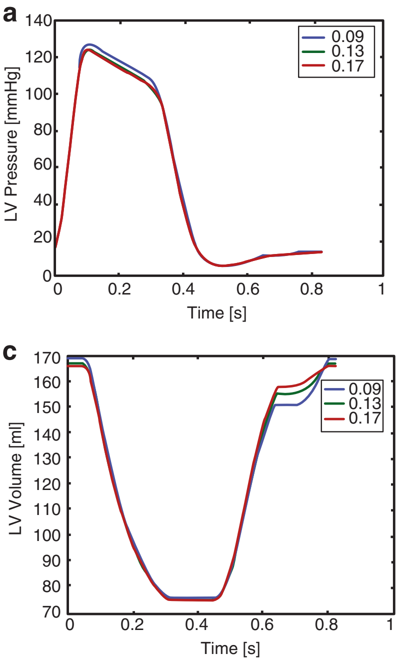


Fig. 5.15 Effect of LA minimum elastance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Elastance is expressed in [mmHg/ml]

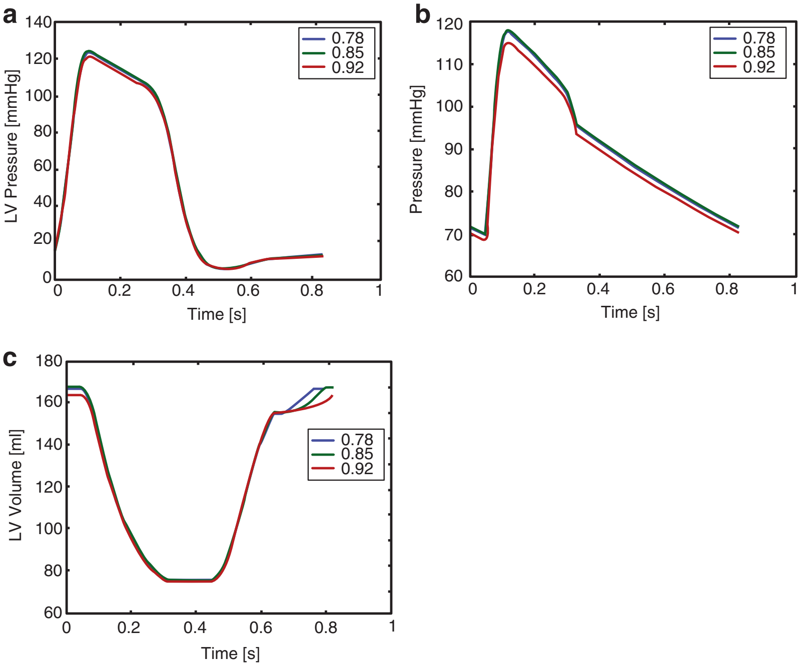


Fig. 5.16 Effect of LA onset of contraction on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Moment of onset is expressed relative to the cycle period

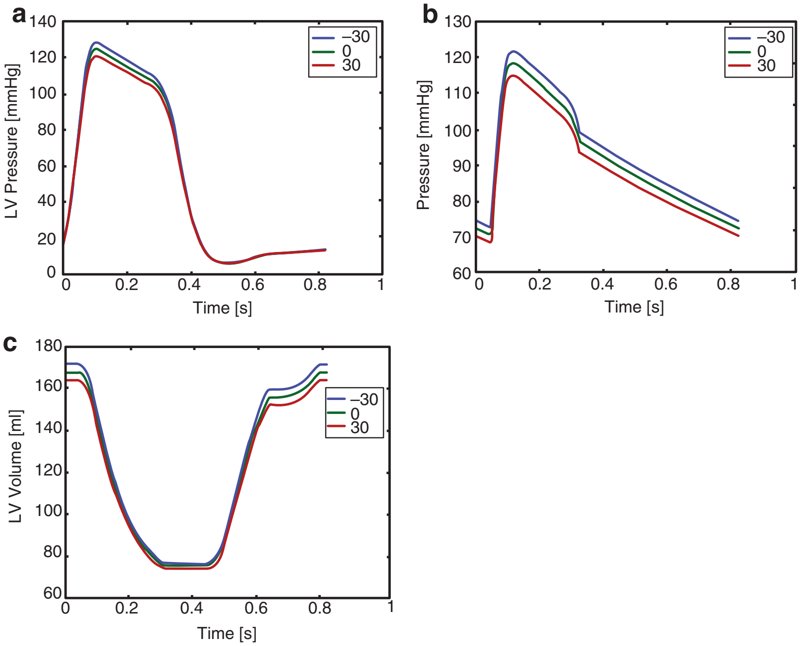


Fig. 5.17 Effect of LA dead volume on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Volume is expressed in [ml]

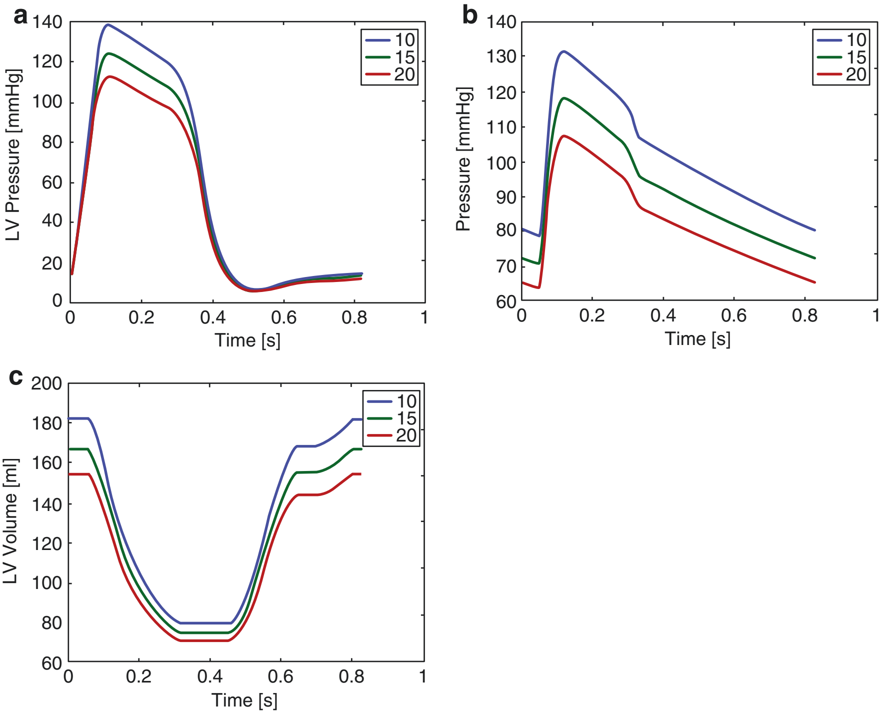


Fig. 5.18 Effect of systemic venous compliance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Compliance is expressed in $[10^{-6}\ \mathrm{cm}^4\ \mathrm{s}^2/\mathrm{g}]$

Figure 5.19 displays the effects of a variation in the aortic valve resistance (only a healthy valve is considered): effects are negligible.

Figure 5.20 displays the effects of a variation in the aortic valve inertance (only a healthy valve is considered): effects are negligible.

Figure 5.21 displays the effects of a variation in the mitral valve resistance (only a healthy valve is considered). The mitral valve resistance mainly affects the slope of the ventricular filling phase.

Figure 5.22 displays the effects of a variation in the mitral valve inertance (only a healthy valve is considered). The mitral valve inertance affects the shape of the ventricular filling phase.

Figure 5.23 displays the effects of a variation in the RV maximum elastance. A lower RV maximum elastance leads to a decrease of LV pressure, aortic pressure and LV volume but the effects are overall small.

Figure 5.24 displays the effects of a variation of the time when RV elastance is maximum: effects are negligible.

Figure 5.25 displays the effects of a variation in the dead volume of the RV: the effects are negligible.

Figure 5.26 displays the effects of a variation in the pulmonary aortic resistance: the effects are negligible.

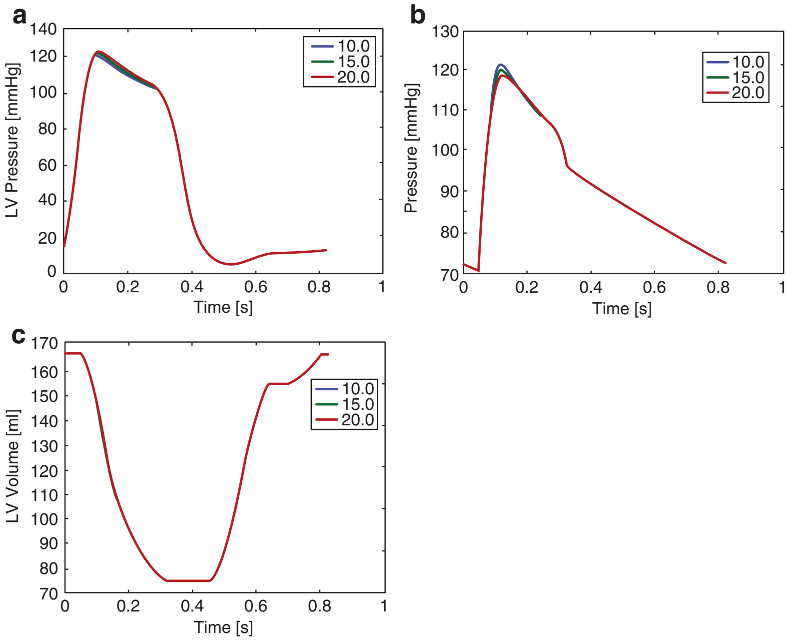


Fig. 5.19 Effect of aortic valve resistance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Resistance is expressed in [g/(cm^4 s)]

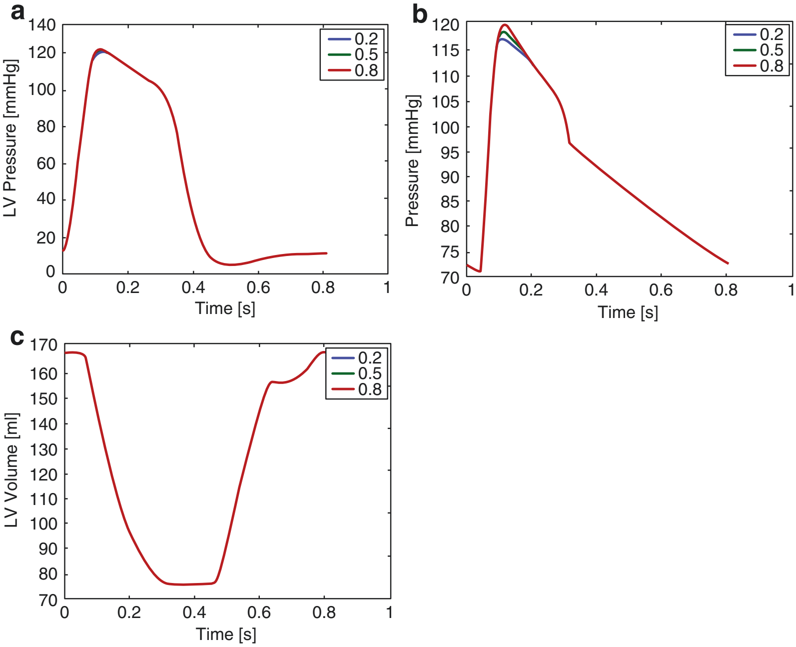


Fig. 5.20 Effect of aortic valve inertance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Inertance is expressed in [cm^2/s]

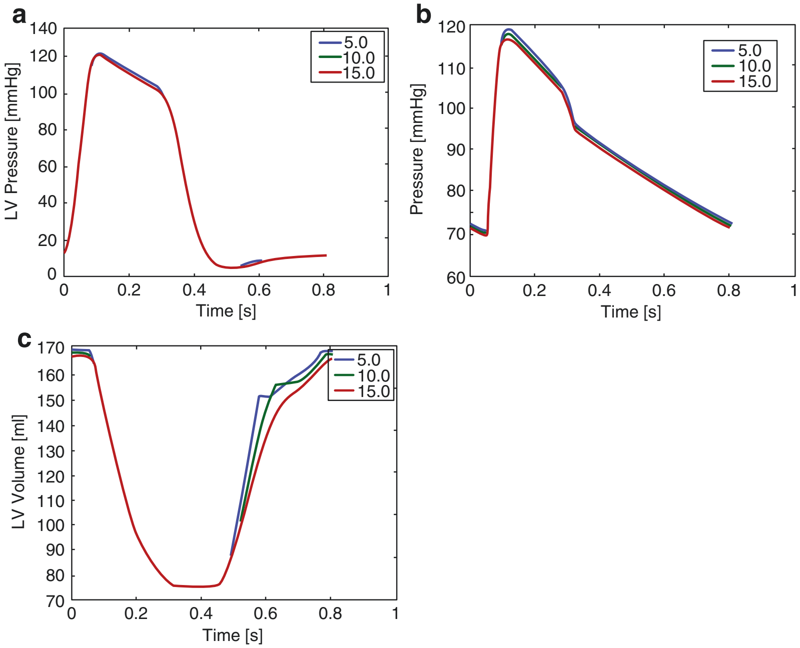


Fig. 5.21 Effect of mitral valve resistance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Resistance is expressed in [g/(cm^4 s)]

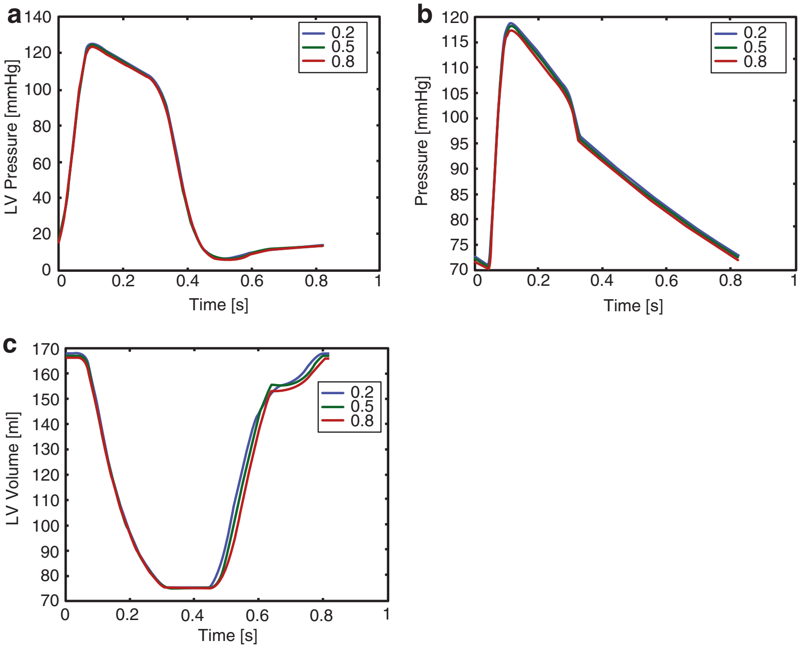


Fig. 5.22 Effect of mitral valve inertance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Inertance is expressed in [cm^2/s]

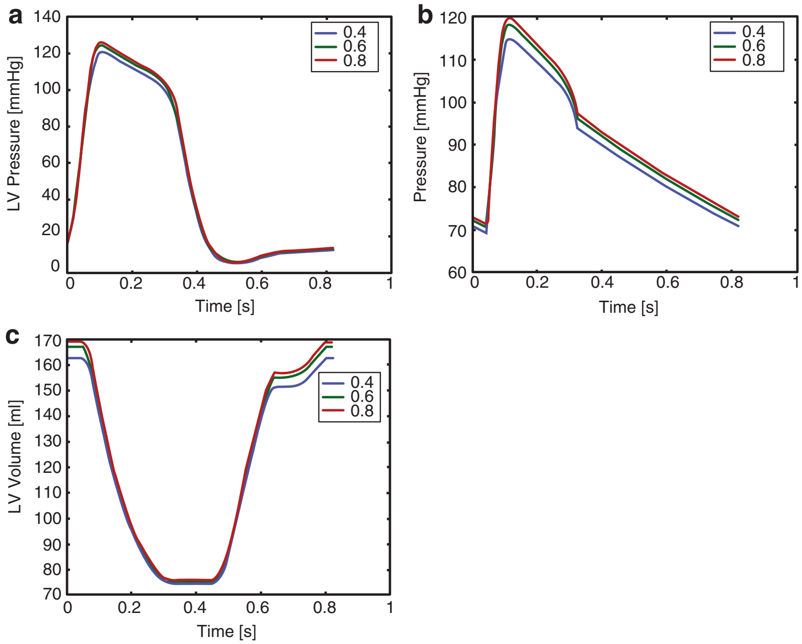


Fig. 5.23 Effect of RV maximum elastance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Elastance is expressed in [mmHg/ml]

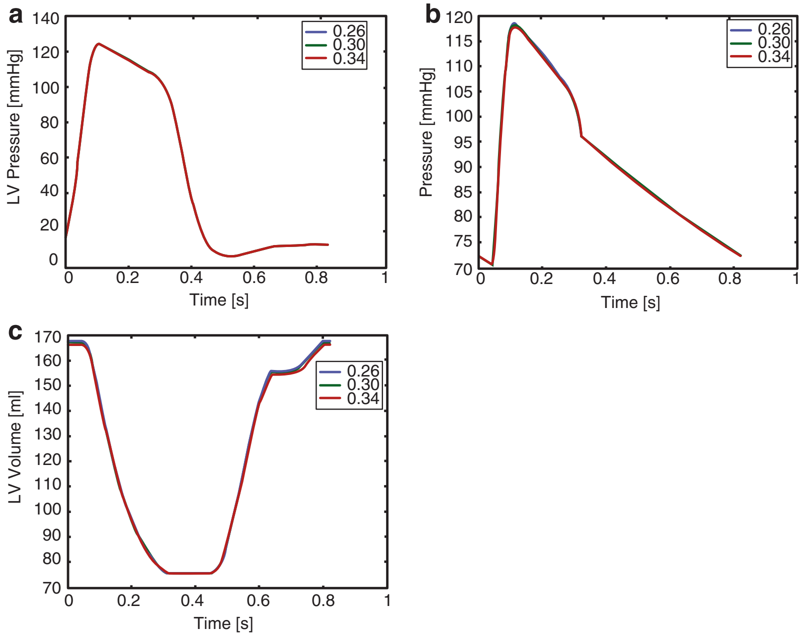


Fig. 5.24 Effect of time when RV elastance is maximum on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Time is expressed in [s]

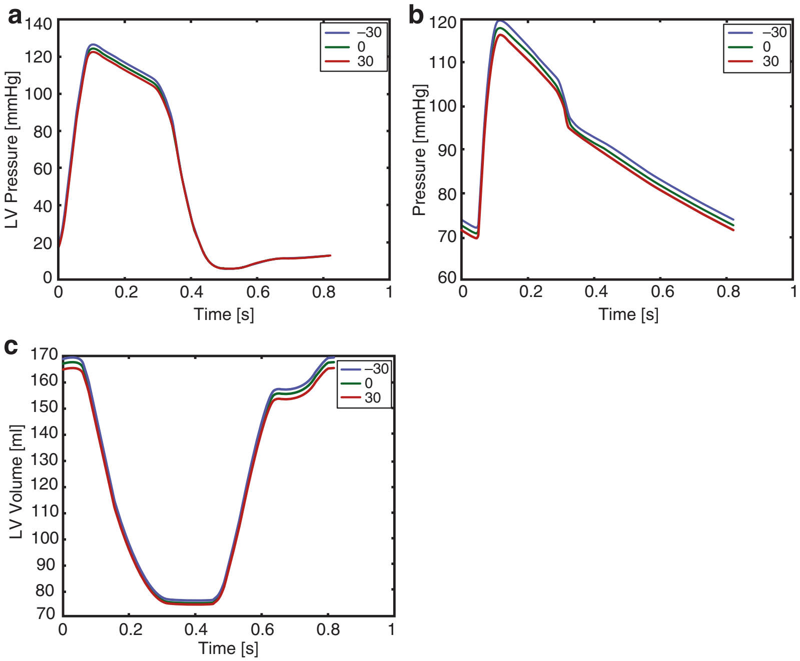


Fig. 5.25 Effect of RV dead volume on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Volume is expressed in [ml]

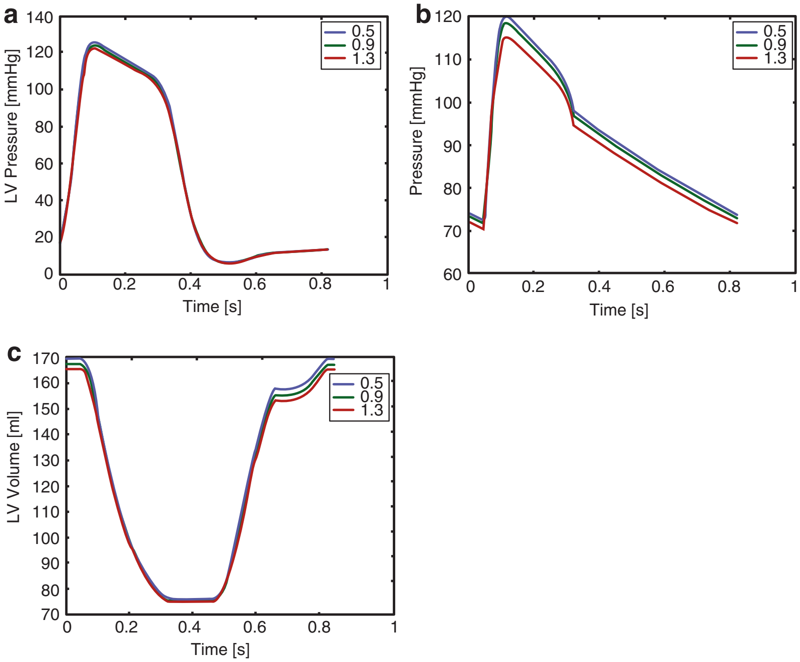


Fig. 5.26 Effect of pulmonary aortic resistance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Resistance is expressed in [g/(cm^4 s)]

Figure 5.27 displays the effects of a variation in the pulmonary aortic compliance: the effects are negligible.

Figure 5.28 displays the effects of a variation in the pulmonary venous compliance. Since the pulmonary venous compliance is very large and stores a large amount of blood, a decrease of this compliance will lead to higher overall levels of pressure and LV volume.

Figure 5.29 displays the effects of a variation in the pulmonary valve resistance (only a healthy valve is considered): effects are negligible.

Figure 5.30 displays the effects of a variation in the pulmonary valve inertance (only a healthy valve is considered): effects are negligible.

Figure 5.31 displays the effects of a variation in the tricuspid valve resistance (only a healthy valve is considered): effects are negligible.

Figure 5.32 displays the effects of a variation in the tricuspid valve inertance (only a healthy valve is considered):

Figure 5.33 displays the effects of a variation in the heart rate. With lower HR value, the ejection time decreases, but the ratio of systolic to diastolic time increases. Stroke volume decreases slightly. Hence, also the overall LV and aortic pressure level decreases slightly. Note that no autoregulation mechanism is used.

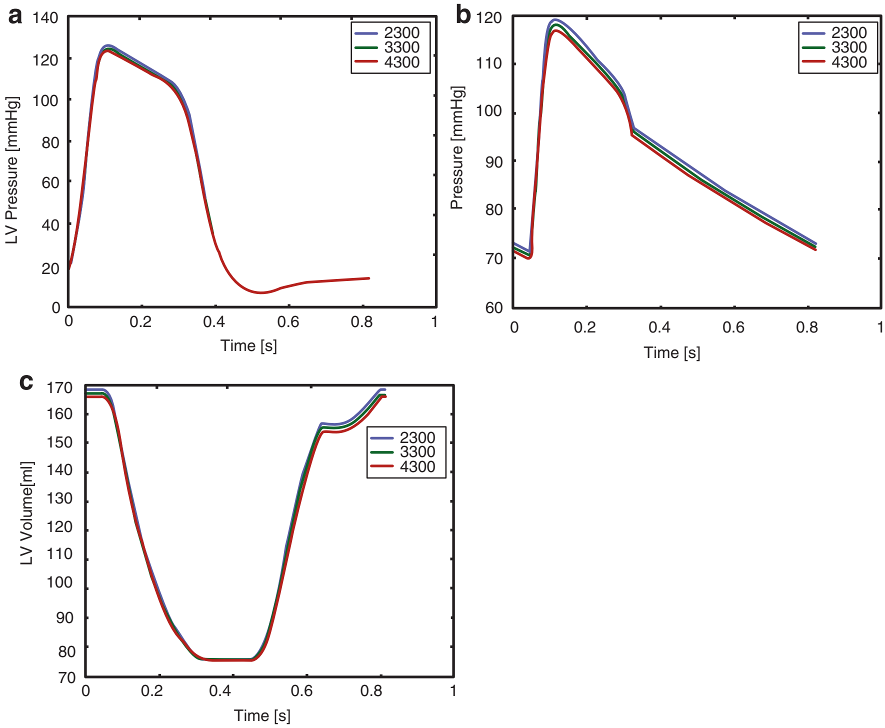


Fig. 5.27 Effect of pulmonary aortic compliance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Compliance is expressed in [10^{-6} cm^4 s^2/g]

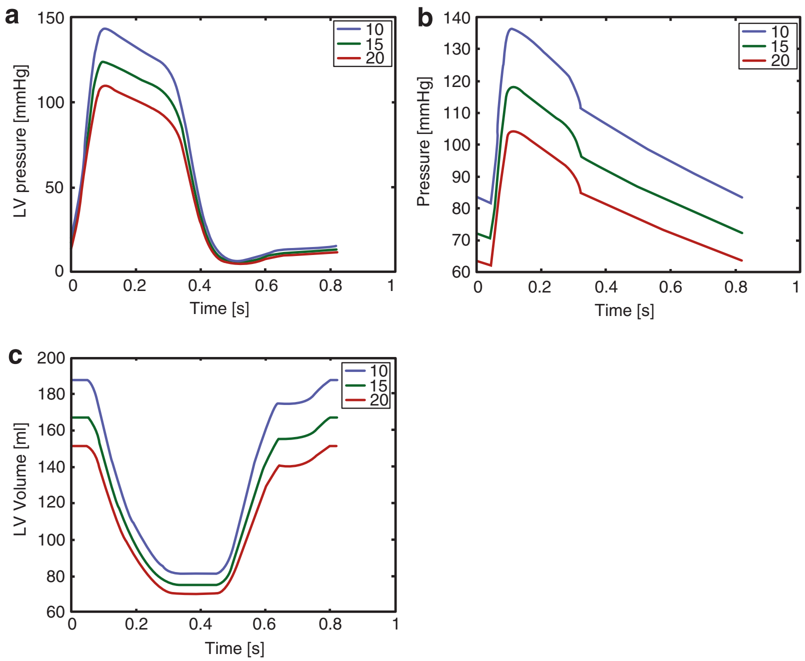


Fig. 5.28 Effect of pulmonary venous compliance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Compliance is expressed in [10^{-6} cm^4 s^2/g]

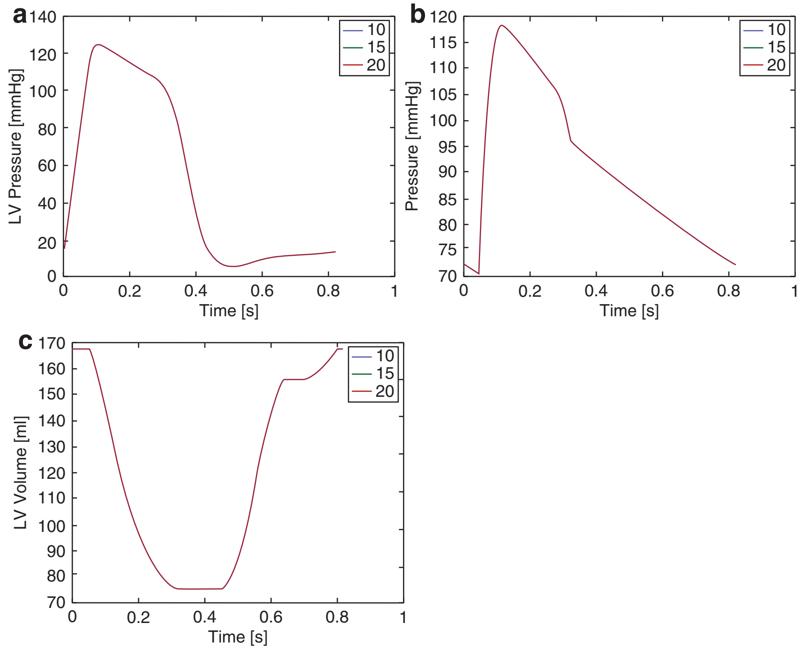


Fig. 5.29 Effect of pulmonary valve resistance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Resistance is expressed in [g/(cm^4 s)]

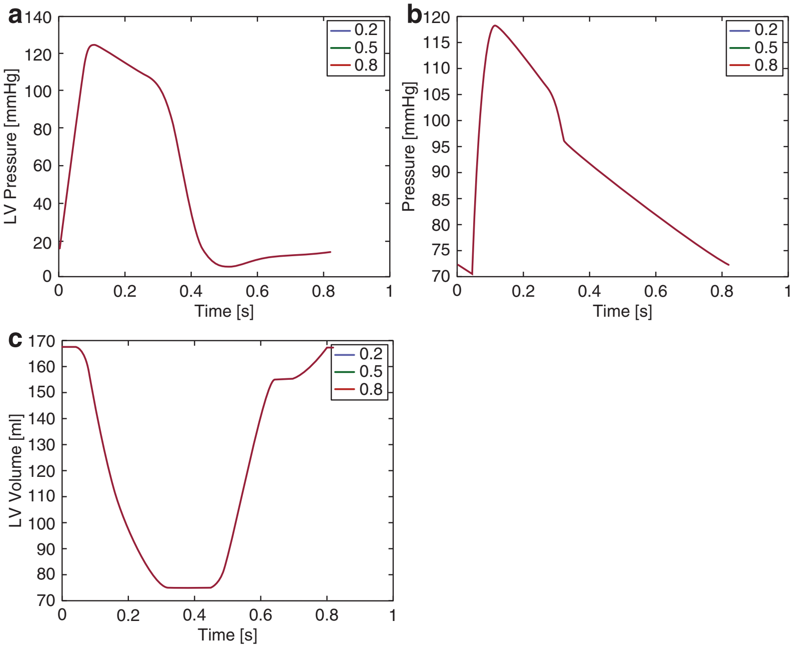


Fig. 5.30 Effect of pulmonary valve inertance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Inertance is expressed in [cm^2/s]

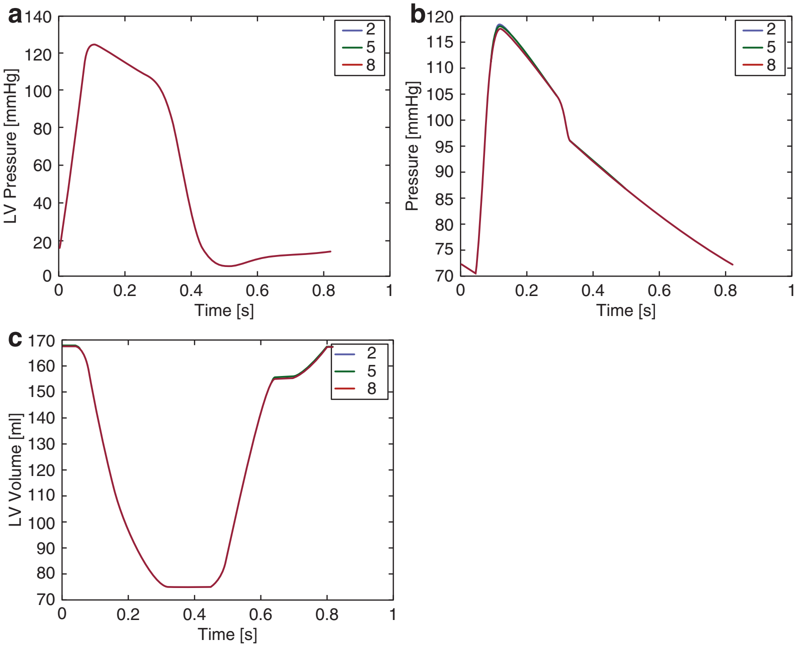


Fig. 5.31 Effect of tricuspid valve resistance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Resistance is expressed in [g/(cm^4 s)]

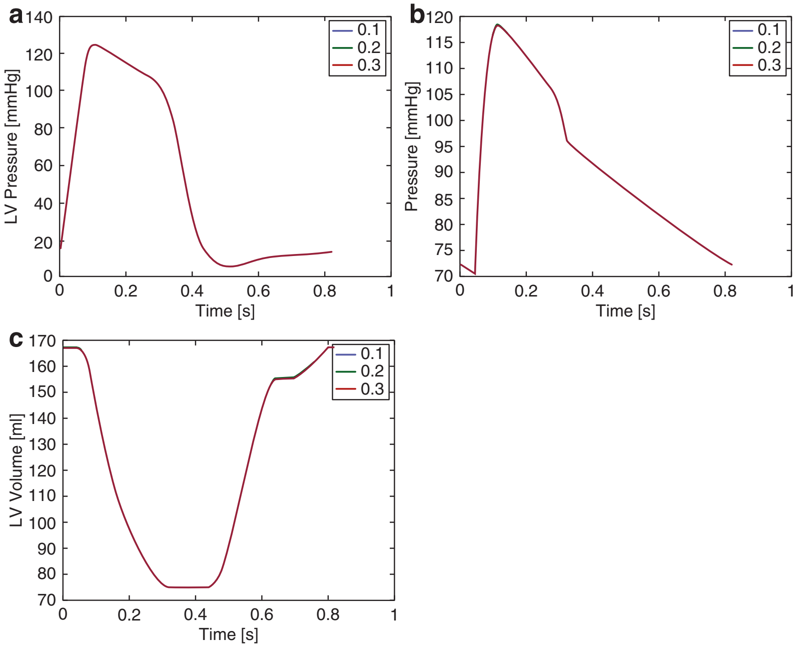


Fig. 5.32 Effect of tricuspid valve inertance on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Inertance is expressed in [cm^2/s]

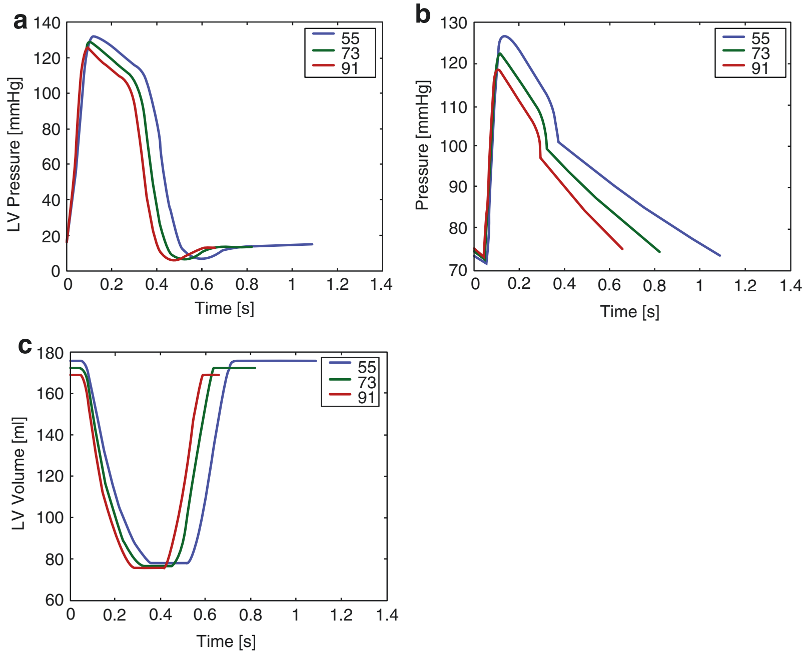


Fig. 5.33 Effect of heart rate on (**a**) LV pressure, (**b**) aortic pressure, (**c**) LV volume. Heart rate is expressed in rpm

5.3.2 Personalization

A set of 12 basic objectives are defined, six for the systemic circulation and six for the pulmonary circulation:

- Maximum aortic pressure
- Minimum aortic pressure
- Average aortic pressure
- Maximum ventricular volume
- Minimum ventricular volume
- Duration of time interval during which the aortic valve is open

The parameters that need to be adapted so as to match the basic objectives are:

- The initial volume in the closed loop model
- The total resistance of the systemic/pulmonary windkessel model
- The compliance of the systemic/pulmonary windkessel model
- The ratio of proximal to distal resistance for the systemic/pulmonary windkessel model
- Timing of the maximum ventricular elastance
- Dead volume of the ventricle

Beyond the set of 12 objectives, the following advanced objectives are employed (both for the systemic and the pulmonary circulation):

- Pressure based objectives (Fig. 5.34)
  - Slope$_1$: the slope of the systolic pressure during early systole (before the valve opens)
  - P_{Dia}: the ventricular pressure at mid diastole
  - Slope$_2$: the slope of the aortic pressure decay during diastole

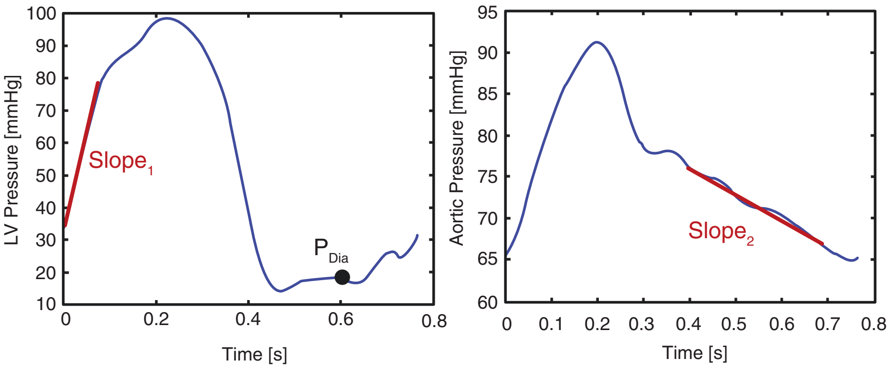


Fig. 5.34 Advanced pressure based objectives

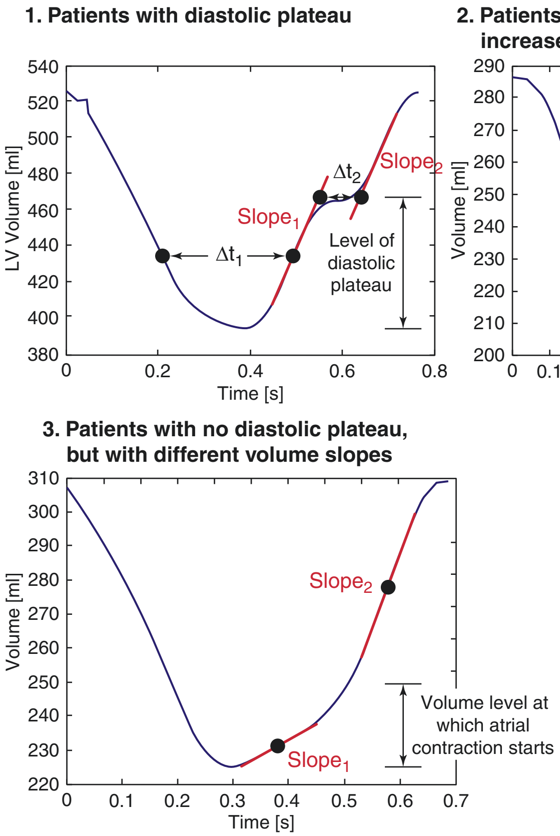


Fig. 5.35 Advanced volume based objectives and classification of patients into three categories

- Volume based objectives: Based on the shape of the volume curve, the patients are classified into three categories (Fig. 5.35).

  For the patients with diastolic plateau the following objectives are defined:

- Δt_1: the delta time on the volume curve between two identical volume values (one in systole and one in diastole), at a volume level of 30% between minimum and maximum volume
- Slope$_1$: the slope of the volume curve in diastole at the above defined volume value
- Δt_2: the duration of the diastolic plateau
- Slope$_2$: the slope of the volume curve when the atrium contracts
- Level of diastolic plateau, expressed in percentages between the min and max volume value

For the patients with diastolic volume increase during early diastole the following objectives are defined:

- Δt_1: the delta time on the volume curve between two identical volume values (one in systole and one in diastole), at a volume level of 30% between minimum and maximum volume
- Slope_1: the slope of the volume curve in diastole at the above defined volume value

For the patients with no diastolic plateau but with different slopes during diastole the following objectives are defined:

- Slope_1: the slope of the volume curve in early diastole, before atrial contraction
- Slope_2: the slope of the volume curve in late diastole, after atrial contraction
- Volume level at which atrial contraction starts

The parameters that need to be adapted so as to match the advanced objectives are:

- Timing of ventricular elastance curve after maximum elastance is reached
- Mitral/tricuspid valve resistance
- Onset of atrial contraction
- Maximum atrial elastance
- Minimum atrial elastance
- Maximum ventricular elastance
- Minimum ventricular elastance

The personalization is performed as described in Sect. 5.5.2 for the open-loop model.

5.4 Advanced Modeling Aspects

5.4.1 Modeling Valve Regurgitation/Stenosis

The valve model used for the computations described in the following is based on the model previously in Mynard et al. (2012). The main concepts for this valve model are:

- the effective area of the valve is introduced as dynamic variable
- the variation of the area of the valve is determined through a valve state variable, which varies between 0 (closed valve) and 1 (open valve)
- the closure and the opening of the valve are driven by the pressure gradient through the valve
- the model of the valve is a pressure drop model

The model and implementation details are described in detail in the paper mentioned above. Using this type of valve, one can model both healthy and pathologic states (aortic valve stenosis, aortic valve regurgitation, mitral valve stenosis, mitral valve regurgitation).

In the following some computational results are presented for both healthy and pathologic states, using the open-loop systemic circulation lumped parameter model.

5.4.1.1 Healthy Valves

For the healthy state the following input data was used: SBP = 120 mmHg, DBP = 70 mmHg, HR = 86 bpm, EF = 70%, EDV = 108 ml.

As a result the maximum elastance value of the left ventricle, the systemic compliance and the dead volume of the left ventricle were estimated. The results are displayed in Fig. 5.36.

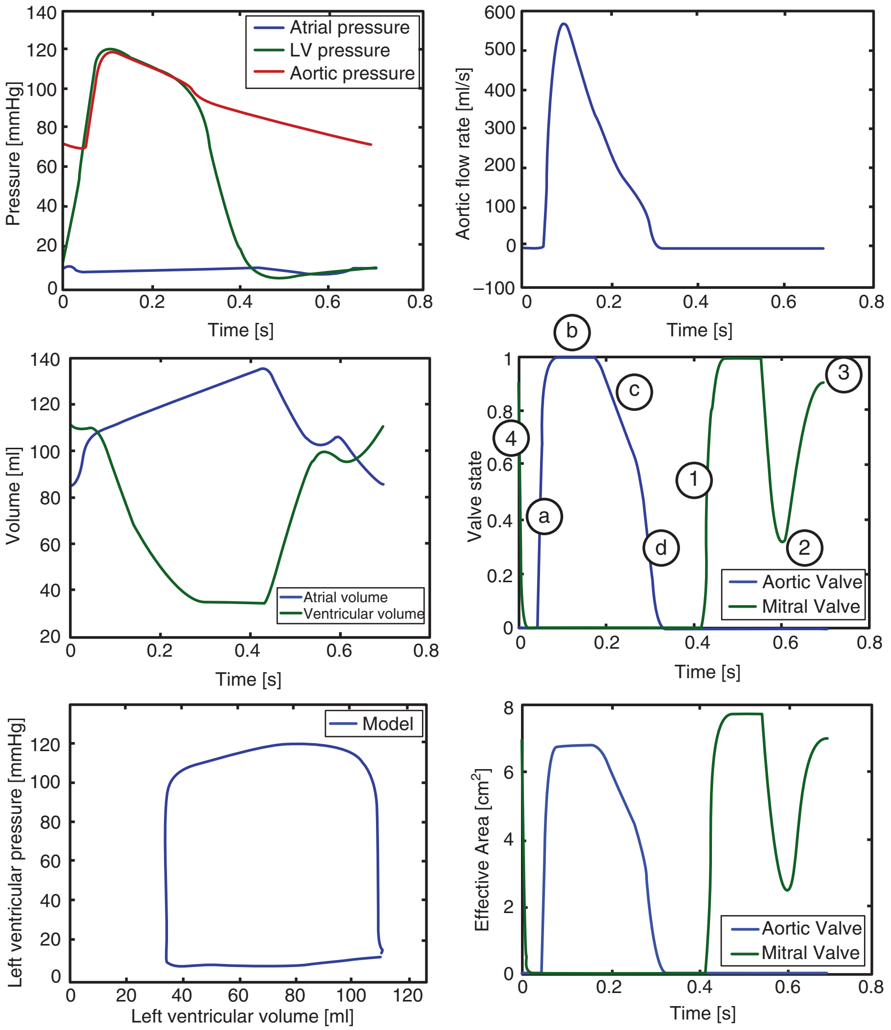


Fig. 5.36 Computational results under healthy conditions

The valve model ensures smooth valve motion (the rate of opening/closure approaches zero as the valve approaches a fully open/closed position). Two constants are used for controlling the speed of valve closure/opening.

Under normal conditions, the **aortic valve** exhibits four phases that have been observed in vivo and are predicted by the model: (a) rapid valve opening when LV pressure initially rises above aortic pressure; (b) a short period in which the valve remains fully open; (c) a slow closure phase during the second-half of systole; and (d) rapid closure when the ventricle begins relaxing.

In accord with in vitro, experimental and clinical data, the simulated normal **mitral valve** undergoes: (1) rapid opening in early diastole when left ventricular pressure falls below left atrial pressure; (2) partial closure during diastasis when the transmitral pressure difference reverses and then oscillates around zero; (3) full re-opening when left atrial contraction re-establishes a positive pressure gradient; and (4) rapid closure due to left atrial relaxation and left ventricular contraction.

The effective valve area of the aortic and mitral valves varies between a minimum value (0.0 cm^2 under healthy conditions) and a maximum value (6.8 cm^2 for the aortic valve and 7.7 cm^2 for the mitral valve under healthy conditions).

5.4.1.2 Aortic Valve Stenosis

To simulate aortic valve stenosis the maximum effective area of the aortic valve is reduced. The same parameter values as for the healthy state were used and no further tuning was performed.

As displayed in Fig. 5.37, with aortic stenosis, the pressure difference across the aortic valve increases significantly, leading to a much higher LV pressure. Three main changes to valve dynamics are notable with aortic stenosis compared with the normal valve. First, the maximal opening area and rate of opening/closure are reduced as a result of the changes prescribed to the maximum effective area of the aortic valve. Second, the slow closure phase is absent, consistent with clinical findings. Third, although the prescribed time course of the ventricular elastance curve is identical to the healthy case, an increasingly restricted aortic valve leads to progressively longer ejection periods, as has been shown clinically.

The changes in system dynamics are clearly visible in the PV loop, when compared to the healthy case.

5.4.1.3 Aortic Valve Regurgitation

To simulate aortic valve regurgitation the minimum effective area of the aortic valve is increased to a non-zero value. The same parameter values as for the healthy state were used and no further tuning was performed.

As displayed in Fig. 5.38, with aortic regurgitation, one can observe that the aortic flow rate is negative during diastole (as a result of the backflow into the LV). The minimal opening area of the aortic valve is non-zero.

Significant changes are also visible in the PV loop, where the isovolumetric relaxation and contraction phases are missing, due to the fact that the aortic valve is never fully closed.

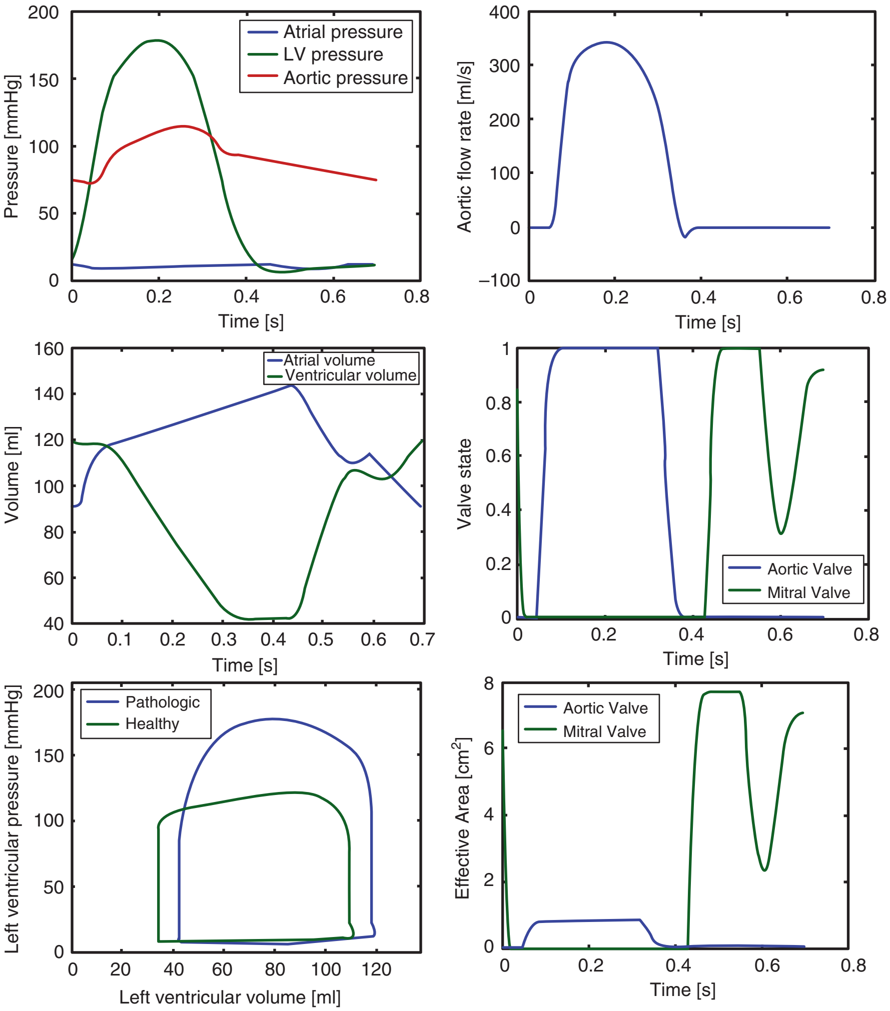


Fig. 5.37 Computational results under aortic valve stenosis

5.4.1.4 Mitral Valve Stenosis

To simulate mitral valve stenosis the maximum effective area of the mitral valve is reduced. The same parameter values as for the healthy state were used and no further tuning was performed.

As displayed in Fig. 5.39, with mitral valve stenosis, atrial volume increases, a restricted mitral valve opening is obtained, and the partial closure phase during diastasis is absent (as has been demonstrated in patients) due to a positive pressure gradient persisting throughout diastole.

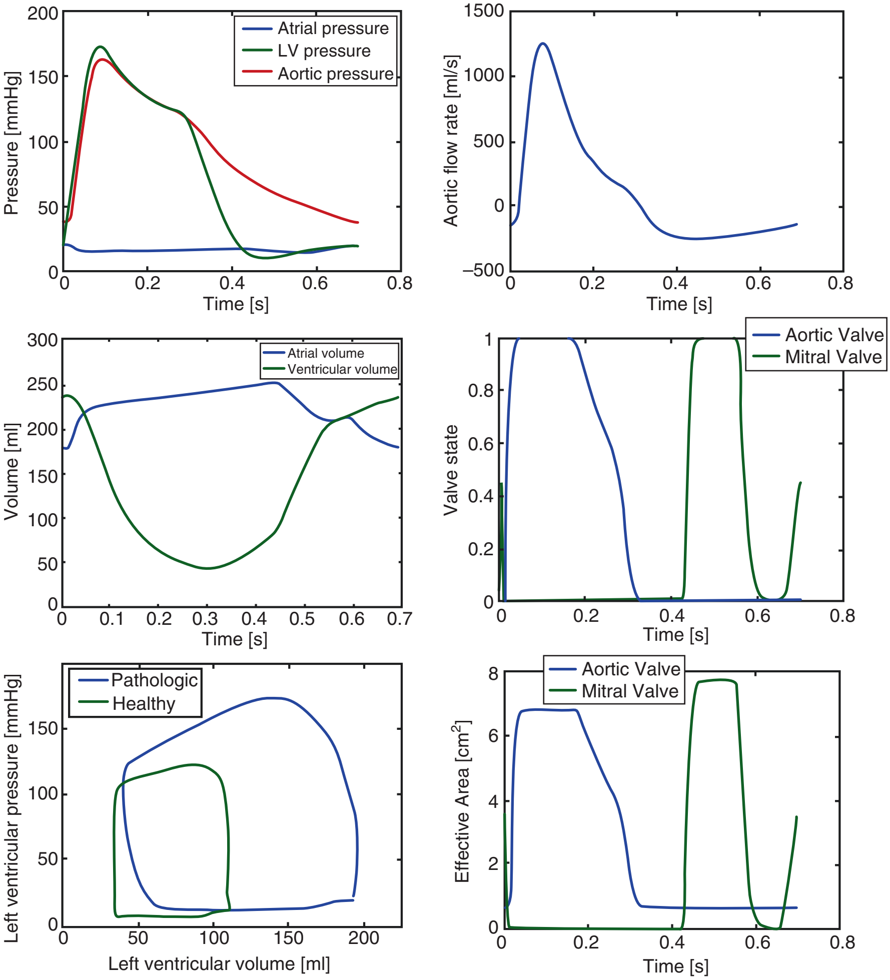


Fig. 5.38 Computational results under aortic valve regurgitation

5.4.1.5 Mitral Valve Regurgitation

To simulate mitral valve regurgitation the minimum effective area of the mitral valve is increased to a non-zero value. The same parameter values as for the healthy state were used and no further tuning was performed.

As displayed in Fig. 5.40, with mitral regurgitation, the minimal opening area of the mitral valve is non-zero. Significant changes are visible in the PV loop, where the isovolumetric relaxation and contraction phases are missing, due to the fact that the mitral valve is never fully closed.

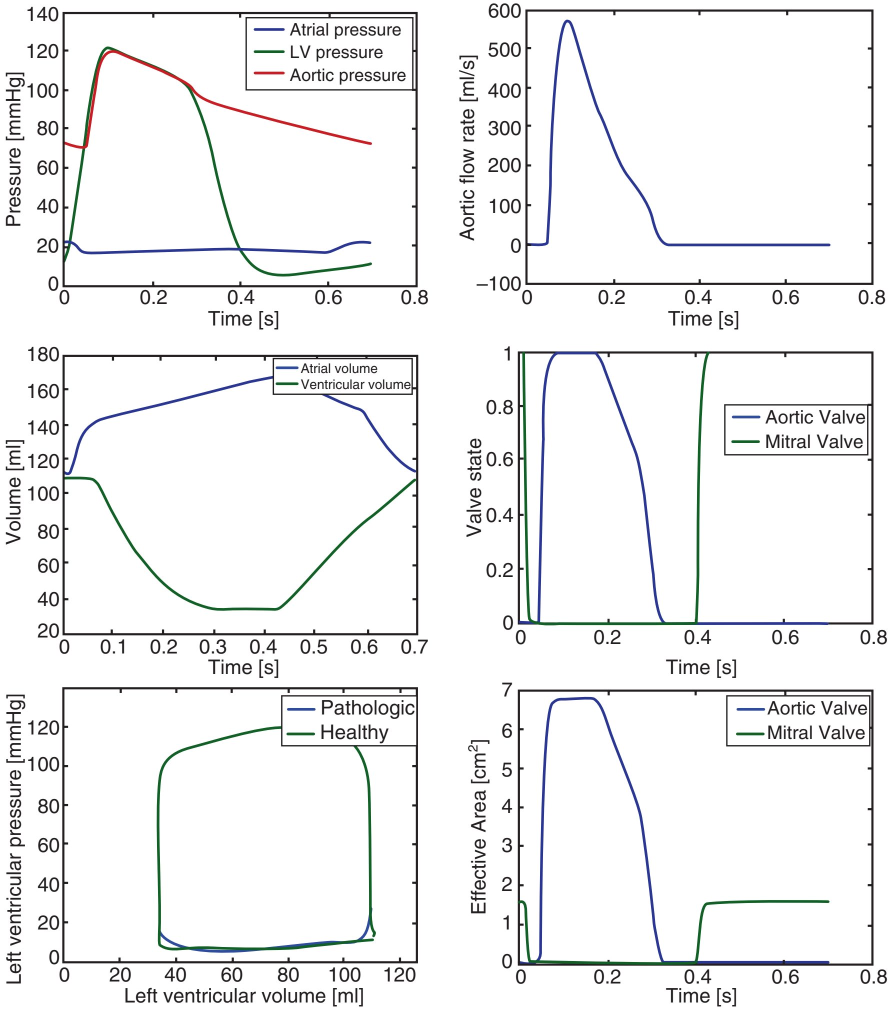


Fig. 5.39 Computational results under mitral valve stenosis

In contrast to the original valve model that was used for computing patient-specific PV loops (composed of a diode, a resistance and an inertance), the herein used valve model is able to compute patient-specific PV loops under pathologic conditions. The main advantage compared to the original valve model is that realistic pressure drop values between the chamber/aorta can be computed.

To be able to compute truly patient-specific results additional patient-specific parameter values are required for the aortic and mitral valve:

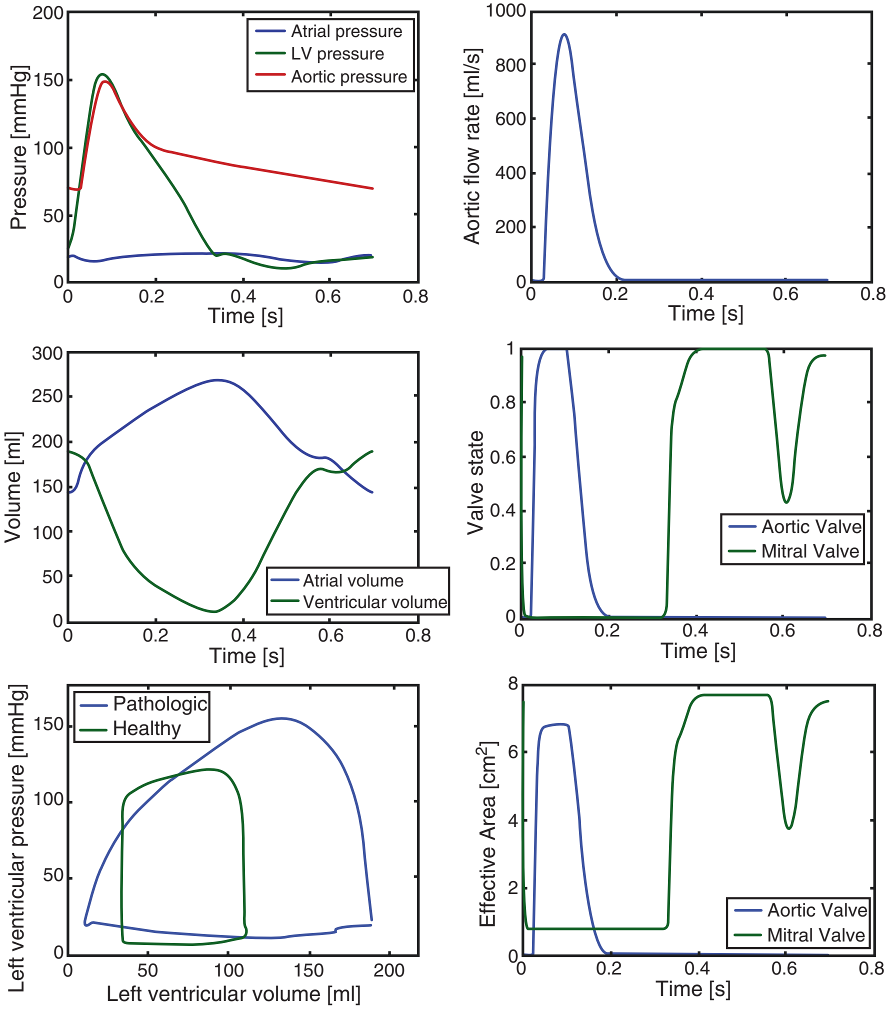


Fig. 5.40 Computational results under mitral valve regurgitation

- maximum annulus area
- minimum annulus area
- dimensions of the LV outflow tract (length and radiuses), dimensions for the ascending aorta (radius)
- information regarding valve opening/closure timing

The original patient-specific values were: SBP, DBP, HR, end diastolic volume, and end systolic volume (or ejection fraction).

5.4.2 Modeling the KG Diaphragm

Another model enhancement is the modeling of the influence of the KG diaphragm on the flow. The KG diaphragm is a soft tissue including the annulus fibrosus and the four heart valves, which undergoes periodic displacement into the atrioventricular chambers under the combined action of several forces, including:

- the pressure force due to the pressure difference across the valves and surrounding tissue
- the tissue strain forces from both the atrium and the ventricle sides that act on the base of the annulus fibrosus
- the frictional force from blood flow
- the elastic force due to the elasticity of the KG diaphragm

The KG diaphragm dynamics depends on the balance of all these forces. The MRI measurements of heart kinetics in several heart cycles in an adult native healthy heart show that in the systolic phase the KG diaphragm moves into the ventricular chamber, and in the end diastolic phase it moves into the atrium (due to the atrial contraction). The total displacement along the long axis of the heart is of about 2–3 cm.

The KG diaphragm dynamic equation (for the left heart) is (Korakianitis et al. 2006):

$$\mathrm{Msav}\frac{\mathrm{d}^2 \mathrm{l}}{\mathrm{dt}^2} = \mathrm{Kst,la}\cdot\mathrm{ela} - \mathrm{Kst,lv}\cdot\mathrm{elv} + \left(\mathrm{Plv} - \mathrm{Pla}\right)\cdot\mathrm{Asav} - \mathrm{Kf,sav}\frac{\mathrm{dl}}{\mathrm{dt}} - \mathrm{Ke,savl} \tag{5.18}$$

The motion of the KG diaphragm redefines the location of the atrioventricular boundary at each time instant and introduces volume changes to the two left chambers (and similarly to the right cardiac chambers):

$$\mathrm{Vlv} = \mathrm{Vlv} + \mathrm{Asav}\cdot\mathrm{l} \quad \text{and} \quad \mathrm{Vla} = \mathrm{Vla} - \mathrm{Asav}\cdot\mathrm{l}. \tag{5.19}$$

Acknowledgments and Disclaimer The authors would like to thank Bogdan Georgescu, Ali Kamen, Constantin Suciu and Dorin Comaniciu for their input.

This feature is based on research, and is not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

References

Acierno L (1994) The history of cardiology. The Parthenon Publishing Group, New York

Burkhoff D (2013) Pressure-volume loops in clinical research: a contemporary view. J Am Coll Cardiol 62:1173–1176

Coogan JS, Pak Chan F, Taylor CA, Feinstein JA (2011) Computational fluid dynamic simulations of aortic coarctation comparing the effects of surgical- and stent-based treatments on aortic compliance and ventricular workload. Catheter Cardiovasc Interv 77:680–691

Dankelmann J, Vergroesen I, Han Y, Spaan JAE (1992) Dynamic response of coronary regulation to heart rate and perfusion changes in dogs. Am J Phys 263:447–452

Drake-Holland AJ, Laird JD, Noble MIM, Spaan JAE, Vergroesen I (1984) Oxygen and coronary vascular resistance during autoregulation and metabolic vasodilation in the dog. J Physiol 348:285–300

Ganong W (1975) Review of medical physiology, 17th edn. Lange Medical Publications, Los Altos

Guyton A (1991) Textbook of Medical Physiology, 8th edn. W.B. Saunders Company, Philadelphia

Hall J (2011) Guyton and Hall textbook of medical physiology, 12th edn. Saunders Elsevier, Philadelphia

Itu L et al (2014) Model based non-invasive estimation of PV loop from echocardiography. In: 36th annual international conference of the IEEE Engineering in Medicine and Biology Society—EMBC 2014, Chicago, August 26–30, pp 6774–6777

Kim HJ, Jansen KE, Taylor CA (2010) Incorporating Autoregulatory mechanisms of the cardiovascular system in three-dimensional finite element models of arterial blood flow. Ann Biomed Eng 38:2314–2330

Korakianitis T et al (2006) A concentrated parameter model for the human cardiovascular system including heart valve dynamics and atrioventricular interaction. Med Eng Phys 28:613–628

Lakin W et al (2007) Whole-body mathematical model for simulating intracranial pressure dynamics. US 7182602 B2

Leaning M, Pullen H, Carson E, Finkelstein L (1983) Modelling a complex biological system: the human cardiovascular system-1. Methodology and model description. Trans Inst Meas Control 5:71–86

Miyashiro JK, Feigl E (1995) A model of combined feedforward and feedback control of coronary blood flow. Am J Phys 268:895–908

Mynard JP, Davidson MR, Penny DJ, Smolich JJ (2012) A simple, versatile valve model for use in lumped parameter and one-dimensional cardiovascular models. Int J Numer Method Biomed Eng 28:626–641

Nocedal J, Wright SJ (2006) Numerical optimization, 2nd edn. Springer, New York

Ottesen J (1997) Nonlinearity of baroreceptor nerves. Surv Math Ind 7:187–201

Ottesen JT et al (2003) Applied mathematical models in human Physiology. Siam Publishing, Philadelphia

Ottesen JT, Olufsen MS, Larsen JK (2004) Applied mathematical models in human physiology. Siam publications, Philadelphia

Paeme S et al (2011) Mathematical multi-scale model of the cardiovascular system including mitral valve dynamics. Application to ischemic mitral insufficiency. Biomed Eng Online 10:86

Quarteroni A et al (2001) Coupling between lumped and distributed models for blood flow problems. Comput Vis Sci 4:111–124

Segers P, Stergiopulos N, Westerhof N, Wouters P, Kolh P, Verdonck P (2003) Systemic and pulmonary hemodynamics assessed with a lumped-parameter heart-arterial interaction model. J Eng Math 47:185–199

Shroff SG, Janicki JS, Weber KT (1985) Evidence and quantitation of left ventricular systolic resistance. Am J Phys 249:H358–H370

Spevack DM, Karl J, Yedlapati N, Goldberg Y, Garcia MJ (2013) Echocardiographic left ventricular end-diastolic pressure volume loop estimate predicts survival in congestive heart failure. J Card Fail 19:251–259

Suga H (1971) Theoretical analysis of a left-ventricular pumping model based on the systolic time-varying pressure/volume ratio. IEEE Trans Biomed Eng 18:47–55

Taher M, Cecchini A, Allen M, Gobran S, Gorman R, Guthrie B, Lingenfelter K, Rabbany S, Rolchigo P, Melbin J, Noordergraaf A (1988) Baroreceptors responses derived from a fundamental concept. Ann Biomed Eng 16:429–443

Tortora G, Anagnostakos N (1990) Principles of anatomy and physiology, 6th edn. Harper and Row Publishers, New York

Ursino M (1999) A mathematical model of the carotid-baroreflex control in pulsatile conditions. IEEE Trans Biomed Eng 46:382–392

Ursino M (2000) Modelling the interaction among several mechanism in the short-term arterial pressure control. In: Danielsen M, Ottesen J (eds) Mathematical Modelling in Medicine. IOS Press, Amsterdam, pp 139–162

Van Slochteren FJ, Grundeman PF, Teske AJ, Bovendeerd PHM (2012) Left ventricular PV loop estimation using 3D echo. Eindhoven University of Technology, Internal Report BMTE 07.42. http://www.mate.tue.nl/mate/pdfs/8738.pdf.

Three Dimensional Reconstruction and Hemodynamic Information Extraction from Monoplane Angiography

6

Irina Andra Tache

Abstract

The visual extension of the 2D angiographic data to 3D is called 3D reconstruction and improves clinical routine. Herein we present results of the application of one of the most frequently used techniques in reconstruction, i.e. based on a pinhole camera. The basic theoretical aspects are provided in the first part: pinhole theory, epipolar geometry, and geometry of the C-arm angiographic system. Secondly, the rotation and translation matrices are adapted to the image acquisition equipment, and the algorithm of 3D reconstruction from two different views is described. Thirdly, the quantitative results of the algorithm are presented for a vessel segment, which was manually tagged by a user in both views. To improve the initial results, i.e. to minimize the reconstruction error, an optimization of key parameters is performed, and the outputs are evaluated.

Finally, an algorithm for estimating the transit time of blood flow from angiographic image series is presented. Starting from the time density curve, which is derived from the pixel grey intensities of a region of interest in an image, more than a dozen of methods for computing the transport time are extracted from the literature and compared for an image series.

6.1 Introduction

The human eye requires additional information to be able to interpret a two dimensional (2D) scene. An alternative for the 2D images, are the 3D volume renderings which imply advanced technology for storing and visualizing the data. This however

---

I.A. Tache
Department of Automatic Control and Systems Engineering, University Politehnica of Bucharest, Splaiul Independentei No. 313, Bucharest 060042, Romania
e-mail: irina.andra@gmail.com

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_6

requires superior computing power and costs. For real time data manipulations and acquisition in the daily clinical workflow, the 2D images are preferred. Still, for the radiographic images, the limitations are increased by the phenomena of overlapping and foreshortening of the vessels.

The vessels' foreshortening means that the vessel appears to be shorter than it is in reality because its direction is not perpendicular to the X-ray beam. When the vessel is parallel to the X-ray source, the foreshortening is maximum and it will appear on an image with a darker color (denser) because the beam has to travel longer along the vessel.

The overlapping effect appears when two vessels appear to intersect in the 2D plan.

The identification of the vascular diseases on grayscale medical images creates real difficulties for the physicians, especially for the inexperienced ones. Typically, in case of cardiac angiograms, around 6–7 imaging series are acquired from different angles in order to help the physician to better assess the severity of the vascular lesions.

Usually, in daily clinical practice, a group of specialists in angiography interpretation meets to deal with the difficult cases. This implies increased costs, more radiation and burdens for the patients. In addition to all these, there are the restrictions of the projection angles due to the mechanics of the C-arm equipment.

The visualization of a vessel segment in 3D (Sect. 6.2) provides an estimation of the cross-section, which can be combined with the transit time estimation in order to compute the blood flow. This measure helps the physicians in appreciating the severity of the stenosis without the need of fractional flow reserve acquired invasively with a catheter.

The transit time is computed from the fluoroscopic angiography based on the indicator dilution theory presented in the Sect. 6.3.

6.2 Three Dimensional Reconstruction of Monoplane Angiography

A main achievement in the computer vision field is the restoration of the stereoscopic vision described by the following mechanism: the two eyes are assimilated with two cameras placed nearby, and their acquired images are recomposed by the computer, as the brain does, by using a dedicated algorithm for generating as output a three dimensional image.

The 3D reconstruction was intensively investigated during the last three decades, simultaneously with the development of the digital medical imaging. Many methods have been presented in the scientific literature for solving this problem, each one dedicated to a specific angiography type, such as: the monoplane, the biplane or the rotational system.

A geometric calibration is typically used in the above mentioned algorithms, which means that an angiographic acquisition is performed for an object with known dimensions. The images are analyzed and the intrinsic and the extrinsic parameters of the medical equipment are extracted. The pixels' correspondences between the different views are performed by the epipolar geometry, which are further included into the pinhole camera model by triangulating the point correspondences (Yang et al. 2009).

Still in most cases of standard angiographic acquisition, the image series are uncalibrated, and the extrinsic and intrinsic parameters are taken from the meta-data of the DICOM file. A spatial representation of the vessels is difficult to be obtained because of the lack of fix references and the algorithm uses only a few pixels' correspondences, identified at bifurcations and/or end points.

In this chapter, a 3D reconstruction algorithm from two different uncalibrated and monoplane angiograms is presented, based on the work of (Brost et al. 2009a, b). The inaccuracies in the results, especially for the coronary cases, are solved with an optimization algorithm where new parameters of the mathematical model of the reconstruction are identified, based on the work of (Yang et al. 2009).

6.2.1 The Fundamentals of 3D Reconstruction

Two main types of coordinates are typically encountered in computer vision: Cartesian and homogenous.

The homogenous coordinates are introduced in order to support the writing of the camera projection matrix.

Considering an image point $p_{2D} = [p_x, p_y]$, it can be represented in the homogenous coordinates—$\widetilde{p_{2D}} \in \mathbb{R}^3$ as follows:

$$\widetilde{p_{2D}} = \varepsilon . \left[p_x \; p_y \; 1 \right] \tag{6.1}$$

where ε is a scaling factor with the property $\varepsilon \epsilon \mathbb{R} \backslash \{0\}$.

An Euclidian transform is, by definition, $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$, where $g(x) \rightarrow RP \bullet x + t$, $R_P \epsilon \, \mathbb{R}^{mxm}$ is the rotation matrix, with the following properties: $R_P^{-1} = R_P^T$, and its determinant is equal to 1 and $t \epsilon \, \mathbb{R}^m$ is the translation vector.

For the homogenous coordinates, the Euclidian transform can be formulated as:

$$p_{3D} = R_p . \widetilde{p_{2D}} + t \tag{6.2}$$

where p_{3D} is the spatial point in Euclidian coordinates. By translating Eq. (6.2) into homogenous coordinates:

$$\widetilde{p_{3D}} = \begin{bmatrix} R_P & t \\ 0 & 1 \end{bmatrix} . \widetilde{p_{2D}} \tag{6.3}$$

To simplify the process of 2D projection, researchers have defined the pinhole camera model. In fact, it describes the perspective vision and consists of a simple camera, whereas the light is passing through a small aperture which projects an inverted version of the object. In radiography this inversion operation is replaced by a magnification. Except for this aspect, both procedures share the same perspective projection (Zhang 1996).

There are three important entities in the pinhole camera model: the x-ray source, the patient and the detector, each one with its own 3D coordinates system.

There are two different types of coordinate system:

- The camera coordinates,
- The real coordinates.

In Fig. 6.1 the geometry of the perspective projection is described. The image is placed between the camera and the object. The focal distance is computed as the distance between the projection center and the image center.

The object will not suffer any inversion for this configuration. Euclidean geometry is a special case of perspective geometry, which consists of projecting the 3D object through lines which cross the small pinhole and form the 2D representation.

Therefore, the image coordinates are linearly projected to the world coordinates. In consequence, a spatial point is mapped to a 2D point using Eq. (6.4):

$$(x_1\, x_2\, x_3)' \rightarrow \left(f.\frac{x_1}{x_3}+o_x,f.\frac{x_2}{x_3}+o_y\right)' \tag{6.4}$$

Moreover, the spatial point expressed in homogenous coordinates $\widetilde{p_{3D}}\left(x_1,,,x_2,,,x_3,,,1\right)$ is converted into Cartesian coordinates p_{3D} using Eq. (6.5):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} f.x_1 \\ f.x_2 \\ x_3 \end{pmatrix} \tag{6.5}$$

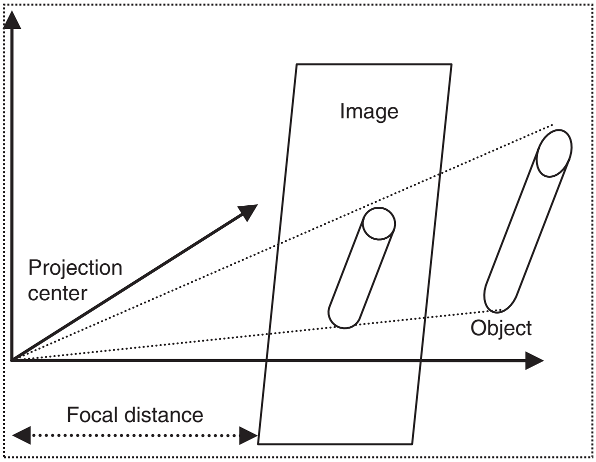


Fig. 6.1 The geometry of the perspective projection

Considering the image center location (o_x, o_y), Eq. (6.5) is inserted into the matrix form:

$$\begin{pmatrix} f.x_1 \\ f.x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} f & 0 & o_x & 0 \\ 0 & f & o_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} . \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \tag{6.6}$$

By tagging the perspective projection matrix as K_p:

$$K_p = \begin{pmatrix} f & 0 & o_x \\ 0 & f & o_y \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.7}$$

the final matrix of the 2D to 3D mapping can be written as:

$$\widetilde{p_{2D}} = K_p . \left[I_{3x3} | 0_3 \right] . \widetilde{p_{3D}} \tag{6.8}$$

The coordinate system of the camera and the coordinate system of the world are related through the extrinsic parameter matrix. It comprises the rotations (R_p) and the translations (t) of the system axes. The rotation matrix, for example, converts the world coordinates— $\widetilde{p_{3D}}$ to the camera coordinates— $\widetilde{p_{3D}^c}$ as follows:

$$\widetilde{p_{3D}^c} = \begin{bmatrix} R_P & t \\ 0 & 1 \end{bmatrix} . \widetilde{p_{3D}} \tag{6.9}$$

The mapping of the spatial point to the image point in case of rotations and translations of the world reference system to the camera system becomes:

$$\widetilde{p_{2D}} = K_p . \left[R_p | t \right] . \widetilde{p_{3D}} \tag{6.10}$$

Having at least two different projections of the same object, and the orientation and the position of the camera for capturing the views are known, the epipolar geometry can be used for generating the 3D reconstruction.

The principle of correspondence of two points located in two different images acquired with two cameras located at known distance which tag the same 3D point is employed by the epipolar geometry. For angiograms, the analyzed object is the centerline of the vessel, which corresponds to a skeleton of the vessel tree. Considering two different projected images (I_1 and I_2 in Fig. 6.2), from which the vessel tree skeleton was extracted.

Point A from image I_1 and the centers of the two X-ray sources O_1 and O_2 form the epipolar plane. The intersection of the epipolar plane with the image plane I_2 forms the epipolar line. Point B from image I_2 must lie on the epipolar line as in Fig. 6.2 in order to assure the correspondences of the pixels between the two views.

Fig. 6.2 Exemplification of epipolar geometry

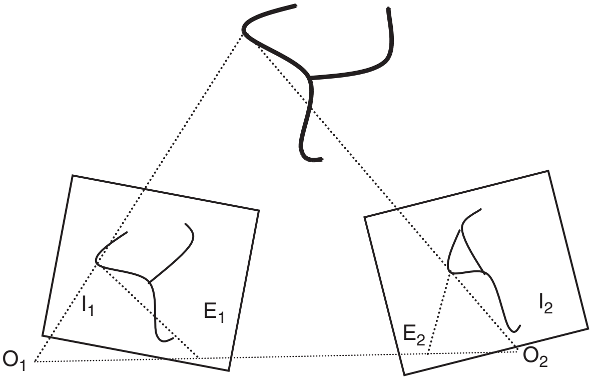
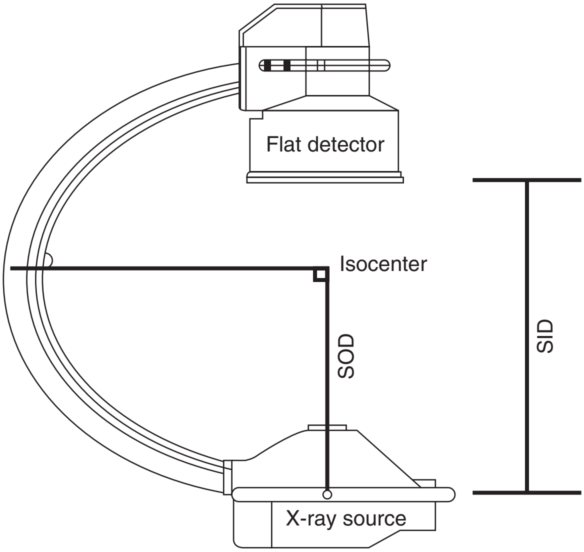


Fig. 6.3 The X-ray angiograph: principal distances (adapted from (Shechter 2004))



If the 2D coordinates from two views manually selected by an expert are known, it enables the computation of the spatial point. The errors can appear due to the incorrect matching of the correspondence points.

6.2.2 The Geometry of the C-Arm Angiography

The C-arm angiography comprises an x-ray source and a flat detector, as displayed in Fig. 6.3.

The isocenter is the center of the reference coordinate system of the equipment (OXYZ) and it is usually located in the region of interest of the patient. The X-ray source direction is parallel to one of the primary axis (OY, for example).

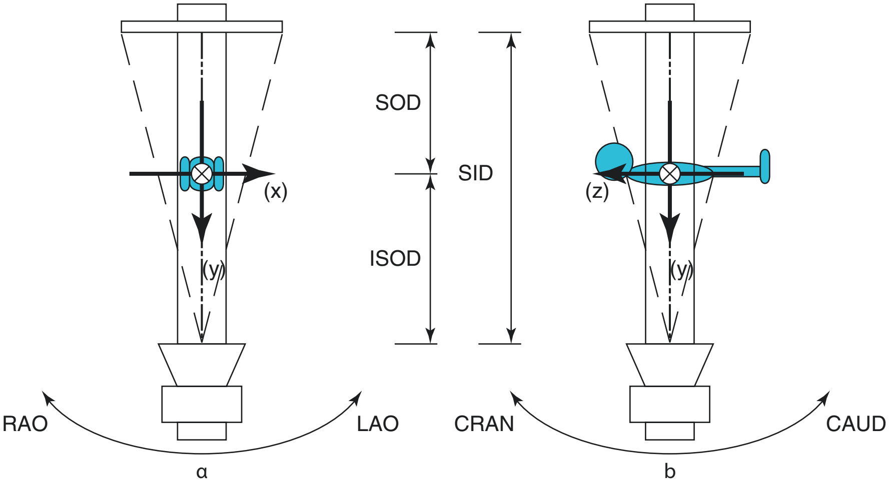


Fig. 6.4 The C-arm angiograph rotation around the axis OX and OZ, respectively (Brost et al. 2009b)

For small vessels, like the coronaries, the X-ray source can be located near the patient’s chest, to obtain a magnification of the vessels.

Therefore, the distances between source and detector (SID) and source to object (SOD) can vary during the clinical examinations. Moreover, the C-arm system rotates around the isocenter in two directions, as depicted in Fig. 6.4: one around the OX axis from the patient’s right arm (RAO) to his left arm (LAO) called primary angulation axis defined by the alpha angle and another one around the OZ axis, from the patient’s cranium (CRAN) to his feet (CAUD) called secondary angulation axis quantified by the beta angle.

In Fig. 6.4 SOD is the distance between the X-ray source and the isocenter which is related to the radius of the source’s motion sphere and ISOD is the distance between the isocenter and the detector which is related to the source motion sphere.

6.2.3 The Mathematical Modeling of the 3D Reconstruction

Starting from Eq. (6.10), the projection matrix (P) is defined as the 3D point mapping to the image points:

$$P = K_i \left[R_P | t \right] \tag{6.11}$$

where $\in \mathbb{R}^{3\times4}$, $K_i \in \mathbb{R}^{3\times3}$ is the matrix of the camera intrinsic parameters and comprises the perspective projection matrix. $R_P \in \mathbb{R}^{3\times3}$ is the rotational matrix and $t \in \mathbb{R}^3$ is the translation vector, which both form the matrix of the camera extrinsic parameters and describe the orientation of the imaging apparatus relative to the world coordinate system.

Starting from the perspective projection matrix, all the parameters are adapted for the angiography geometry. The focal distance—*f* is replaced by the ratio between SID and the pixel size in millimeters—*s*, which has usually the same value for both axes of the image detector, as in Eq. (6.12):

$$f \to \frac{SID}{s} \tag{6.12}$$

The matrix of the camera intrinsic parameters is described in (Brost et al. 2009a, b):

$$K_i = \begin{pmatrix} \frac{SID}{s} & 0 & o_x \\ 0 & \frac{SID}{s} & o_y \\ 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{6.13}$$

whereas o_x and o_y are the isocenter coordinates of the C-arm system into the image plane. The translation vector—*t* is dependent on SOD:

$$t = \begin{pmatrix} 0 \\ SOD \\ 0 \end{pmatrix} \tag{6.14}$$

The rotation around the OY axis, which is related to the primary angle is described by the matrix R_α:

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.15}$$

The rotation around the OX axis, which is related to the secondary angle is described by R_β:

$$R_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & \sin(\beta) \\ 0 & -\sin(\beta) & \cos(\beta) \end{pmatrix} \tag{6.16}$$

Because the C-arm system does not have a rotation around the OZ axis, the global rotation matrix becomes:

$$R_P = R_\alpha . R_\beta \tag{6.17}$$

Knowing that the last element in the vector $\widetilde{p_{3D}}$ is 1, Eq. (6.18) can be written as:

$$\widetilde{p_{2D}} = K_i \left(R_P . p_{3D} + t \right) \tag{6.18}$$

The spatial point in Cartesian coordinates is defined in Eq. (6.19):

$$p_{3D} = R_P^{-1}.\left(K_i^{-1}.\widetilde{p_{2D}} - t\right) \tag{6.19}$$

where $p_2D = [x_2, y_2]$ contains the coordinates of the image point.

A scale factor (γ) is introduced to perform the conversion from homogeneous to normal coordinates as in Eq. (6.20):

$$\widetilde{p_{2D}} = \begin{pmatrix} \gamma.x_2 \\ \gamma.y_2 \\ \gamma \end{pmatrix} = \gamma.\begin{pmatrix} p_{2D} \\ 1 \end{pmatrix} = \gamma.\widetilde{p_{2D}}^{*} \tag{6.20}$$

where $\widetilde{p_{2D}}^{*} = \left(p_{2D}, 1\right)^T$.

The equation of the straight line which connects the image point and the spatial point, also called the epipolar line, is:

$$d_\gamma = \gamma.R_P^{-1}.K_i^{-1}.\widetilde{p_{2D}}^{*} - R_P^{-1}t \tag{6.21}$$

If the origin of the line is $O = RP^{-1} \bullet t$ and its direction is $r = RP^{-1} \bullet Ki^{-1}$, the Eq. (6.21) can be rewritten as: $d\gamma = O + \gamma \bullet r$.

The user can manually match the points from two projections. The spatial point must lie on both epipolar lines described by:

$$\begin{aligned} p_{3D} &= O_1 + \gamma_1.r_1 \\ p_{3D} &= O_2 + \gamma_2.r_2 \end{aligned} \tag{6.22}$$

After determining the scaling factors and considering that the spatial point must be found at the intersection of the two epipolar lines, its spatial coordinates are computed as:

$$p_{3D}^{*} = \frac{O_1 + \gamma_1.r_1 + O_2 + \gamma_2.r_2}{2} \tag{6.23}$$

For performing the reconstruction, the geometrical configuration of the imaging system is required. In the DICOM files, the following parameters are stored for every image series: the distance between the source and detector (SID), the distance between the source and the patient, the pixel spacing in mm, the primary and secondary angles (alpha and beta).

The distance between the source and isocenter (SOD) is generally fixed for the imaging equipment, even though the accuracy of these parameters is debatable, especially due to the motion of the table during the acquisitions.

6.2.4 The Optimization Problem of the Parameters

The reconstruction error is the distance between the actual coordinates of the 2D points tagged in both images and the coordinates from the back projection of the spatial point.

The error can become important, especially in case of uncalibrated angiograms, because the extrinsic and intrinsic parameters provided by the DICOM file are not accurate, e.g. due to the possible motion of the table during the image acquisition. Issues may also be caused by the voluntary or involuntary patient motions such as during the respiration or the contractions during the cardiac cycle. Furthermore, the isocenters of the two angiographic views don't coincident and their imaging axes are not orthogonal, as it is stated in (Yang et al. 2009). The gravitational and inertial forces will influence the spherical rotation of the C-arm system around its isocenter (Wächter 2009).

To deal with all these aspects, a reevaluation of the reconstruction parameters is needed, by performing a minimization of an objective function. It is constructed as the Euclidean distance error between the image coordinates of the vessel skeleton and their corresponding back projection coordinates of the 3D reconstruction.

For solving the optimization problem, the nonlinear algorithm of Levenberg–Marquardt is used due to its good computational times for solving the sum of squared residuals functions (Lourakis 2005).

Yang et al. in 2009 introduced in the objective function the influence of the table motion modeled through a translation vector.

There are 11 parameters proposed for the optimization: SID and SOD (sometimes replaced by the source-to-patient distance), the rotation angles for both projections, the coordinates of the image plan center (o_x, o_y) and the table motion translation vector—m_t.

The proposed algorithm for the 3D reconstruction optimization is:

- the parameters are initialized with values extracted from the DICOM file. The central point of the image plan is chosen as the image center.
- the minimization of the objective function (O) is formulated as:

$$\min O\left(SID_i, o_x, o_y, \alpha_i, \beta_i, SOD_i, \overrightarrow{m_t}\right) = \\ \sum\nolimits_{i=1}^{n}\left[\left\| p_{2D1,i} - P_1.\left(p_{3Di} + \overrightarrow{m_t}\right)\right\|^2 + \left\| p_{2D2,i} - P_2.(p_{3Di} + \overrightarrow{m_t}\right\|^2\right] \tag{6.24}$$

- the distance error computed for the new optimization projection matrix is evaluated.

6.2.5 Results and Discussion

The complex motion of the coronaries during the cardiac cycle complicates the problem. For testing the algorithm, two different cases are presented, one when the images do not correspond to the same cardiac phase and another one when the images correspond to the late diastolic phase, whereas the vessels are supposed to have less motion.

The 3D reconstruction of the centerline of a circumflex coronary tagged in two projections pertaining to different cardiac phases is presented in the following.

The radiologic expert selected the intermediate points of the vessel centerline and its final coordinates are obtained from the mathematical modeling of the segment with a cubic spline, with a resolution of 100 fractions.

To quantify the quality of the 3D reconstruction, the distance error is computed and an optimization algorithm is proposed and the adjusted parameters are determined. On one hand there are the specific parameters for each projection, such as SID and SOD, for which the initialization is performed with the distance source-to-patient. On the other hand, the global parameters are: the coordinates of the center of the image plan (o_x, o_y), and the translation vector reflecting the table motion with the coordinates [mx, my, mz].

In Table 6.1 all the above mentioned parameters are provided. By varying the coordinates of the table motion vector, different optimization versions are created.

The initial version corresponds to the default 3D reconstruction, without optimization, version 1 implies the variations of all three coordinates of the vector, version 2 varies only the OX and OZ coordinates (mx, mz), version 3 implies the variation of the OX and OZ components, for version 4 only the OX component is taken into account, as opposed to version 5, for which only the OZ component is considered.

Interesting observations can be made based on Table 6.1, regarding the important changes of the isocenter coordinates and the rotation angles (alpha is changing with a mean of −32.6° for the first projection and with a mean of 7.8° for the second projection; beta is changing with a mean of −32° for the first projection and with a mean of 35.7° for the second projection) during the optimization process.

The isocenter constitutes the center of the reference coordinate system, and its variation confirms the complexity of the acquisition process.

The C-arm system doesn’t rotate strictly around the isocenter, the table is moved during the two acquisitions by the physician, which tries to obtain the best views from the heart of the patient. Furthermore, the vessel motion during the cardiac cycle can also be accounted for during the modification of the translation vector.

Table 6.1 Optimization parameters for 3D reconstruction (Tache 2015)

Vers. opt.	Pr.	SID [mm]	SOD [mm]	Alpha [°]	Beta [°]	ox	oy	mx	my	mz
Initial	1	1112	785	1.6	−39	256	256	–	–	–
	2	1045	783	−27	−15					
v1—[mx, my, mz]	1	1112	784	5.9	−13	256	258	2	−10	−5
	2	1044	785	−30	−40					
v2—[mx, mz]	1	1112	784	32	−13	255	245	−1	–	−4
	2	1037	794	−43	−64					
v3—[mx, my]	1	1112	785	39	1.5	255	266	−3	−3	–
	2	1036	794	−43	−56					
v4—[mx]	1	1112	785	60	8.3	259	255	−4	–	–
	2	1024	810	−24	−62					
v5—[mz]	1	1112	785	34	−19	241	238	–	–	−5
	2	1032	800	−42	−67					

Table 6.2 Quantification of the optimization process (Tache 2015)

	Pr1	Pr2
Vers. Opt.	Error [mm]	Error [mm]
Initial	24.42	28.63
v1—[mx, my, mz]	0.50	0.60
v2—[mx, mz]	0.79	0.99
v3—[mx, my]	0.51	0.62
v4—[mx]	1.63	2.24
v5—[mz]	1.88	2.46

Another observation is related to the very small adjustments of SID and SOD. Only for the second projection, SOD is corrected more significantly, with a mean of 13.6 mm.

The center of the image plan moved its initial coordinates with a mean of 2 pixels on the OX axis and with a mean of 3 pixels on the OY axis.

In Table 6.2 the distance errors computed as the sums of Euclidian distances between consecutive points of the centerlines are given in millimeters.

As can be observed in Table 6.2, the error of the initial 3D reconstruction (without optimization) is augmented, for the first projection it is of around 24.4 mm and for the second projection it is of approx. 28.6 mm. Considering the small dimensions of the coronaries these errors are significant.

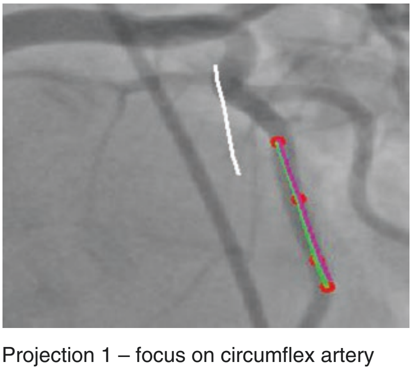
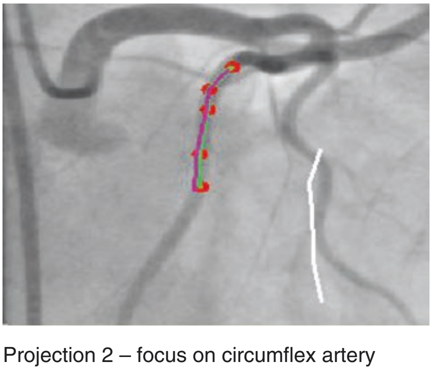
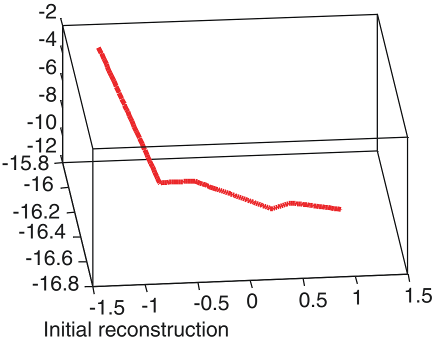
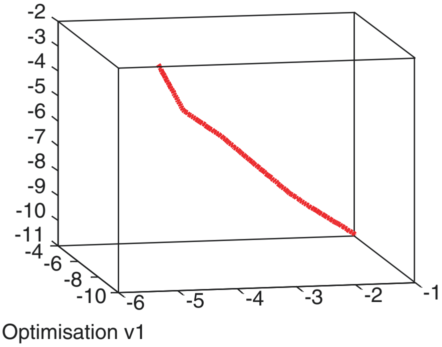
For the optimized versions, the minimum errors were encountered for the variations of the motion table vector on all the three axis with a mean of 0.55 mm for both projections, followed by version 3 with a mean of 0.57 mm for both projections. This is probably due to the fact that the displacement of the isocenter on the OY axis is more important. The maximum error is obtained for version 5, where only the motion on OZ axis was considered: the isocenter is not moving significantly in this direction.

In Table 6.3, the vessel segment centerline was selected in two projections with a green line, the corresponding back projections of the 3D centerline is marked with white lines for the initial reconstruction and with magenta for the optimized version 1. Finally, on the bottom row, the 3D centerlines are plotted for both projections. The best match between the centerlines can be observed for version 1, whereas the back projections are almost perfectly superimposed over the initial 2D centerline.

In Table 6.4 only a subset of the optimization versions are depicted. For version 2, its back projections are marked with yellow, for version 3 the lines are marked with red (a good matching with the initial centerlines is observed) and for version 4, the lines are marked with blue. Version 5 was removed from this figure, because it had large errors.

The algorithm is sensitive to a good matching of the vessels on both projections.

Table 6.3 The visualization of the initial and optimized 3D reconstruction (Tache 2015)

	
Projection 1—Focus on circumflex artery	Projection 2—Focus on circumflex artery
	
Initial reconstruction	Optimization v1

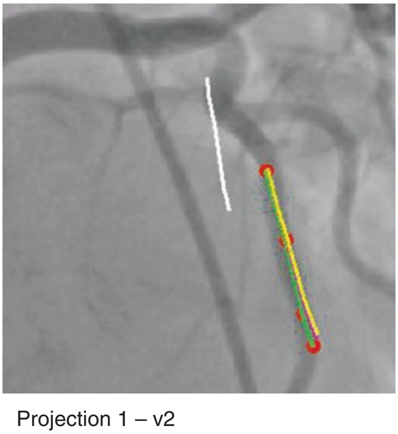
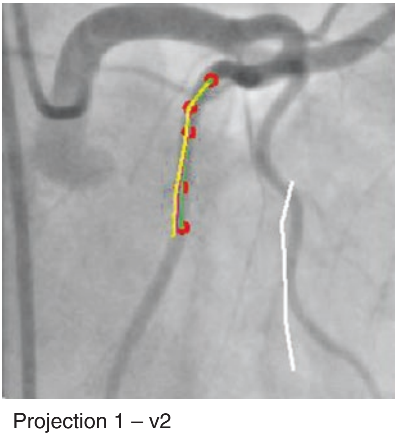
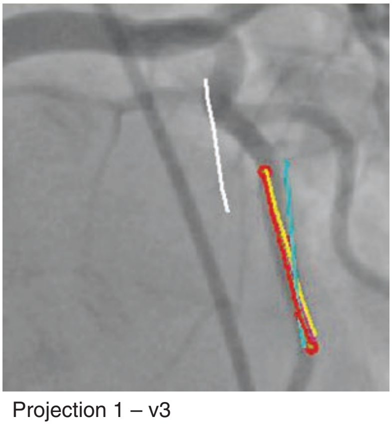
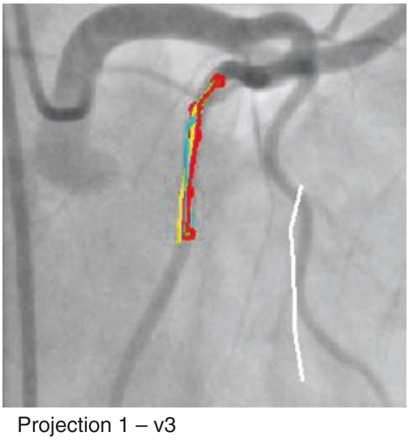
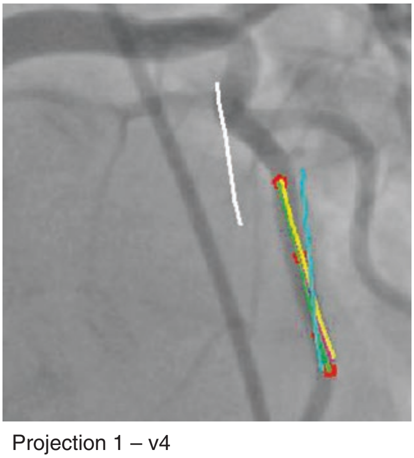
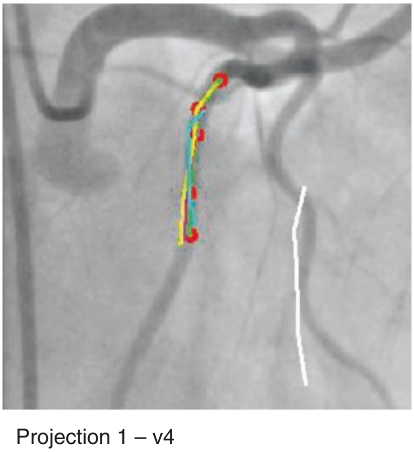
6.3 Bolus Transit Time Method

One of the advantages of angiogram imaging data is the temporal acquisition. The more images are acquired per second, the better the time resolution will be, and more information related to the contrast agent propagation may be extracted.

The temporal resolution depends on the type of angiogram and on the organ circulation time. For cardiac angiograms, the acquisition typically lasts for a few seconds and the contrast quickly fills all vessels. In the cerebral circulation, blood flow velocities are smaller and the acquisition time typically doubles.

The integration of the pixel intensities for a region of interest is called time density curve. Because of noises, the generated curve must be post-processed with filtering and fitting operations.

Table 6.4 Visualization of the optimization versions (Tache 2015)

	
Projection 1—v2	Projection 1—v2
	
Projection 1—v3	Projection 1—v3
	
Projection 1—v4	Projection 1—v4

6.3.1 Indicator Dilution Theory

The time density curve (TDC) was associated with the indicator dilution curve, which was employed in clinical diagnostics. There are a few hemodynamic parameters, such as the blood volumes and the cardiac output, which can be detected from this curve (Zierler 2000).

The estimation of the mean transit time is related to the blood volume, perfusion bed computation, and, finally, it is a prerequisite for the determination of the blood flow parameters such as velocity and flow rate. Their accurate estimation has a major clinical importance (Mischi et al. 2008).

For a constant region volume, the pixel intensity is proportional to the amount of the contrast agent in that region. Assuming a ROI in an image, the integral of the pixel intensities can be computed for every frame (see Fig. 6.5). A mathematical formulation of the TDC is offered by (Schrijver 2002) which takes into consideration the DSA:

$$D(t) = D_{roi}(t) = \frac{-1}{\mu_0} \iint\limits_{ROI} I_{dsa}(x, y, t)\, dA \tag{6.25}$$

where $D(t)$ is the time density curve, D_{roi} is the amount of contrast agent, $I_{dsa}(x, y)$ are the pixel values of the DSA, μ_o is the attenuation coefficient.

For a more accurate approximation of the gray levels with the contrast density, the logarithmic DSA can be first computed for all images from the series:

$$D(t) = \frac{-1}{\mu_0} \iint_{ROI} I_{\log\_dsa}(x, y, t)\, dA \tag{6.26}$$

where $I_{log\_dsa}(x, y)$ are the pixel values of the logarithmic DSA.

The major problem of the DSA application for the coronary arteries is the motion of the vessels during the respiration and systolic contraction of the heart muscle. There are also external sources of movement (the patient, the table and the C-arm gantry). An approach for managing this aspect is the cardiac gated image acquisition, as presented in Eq. (6.27):

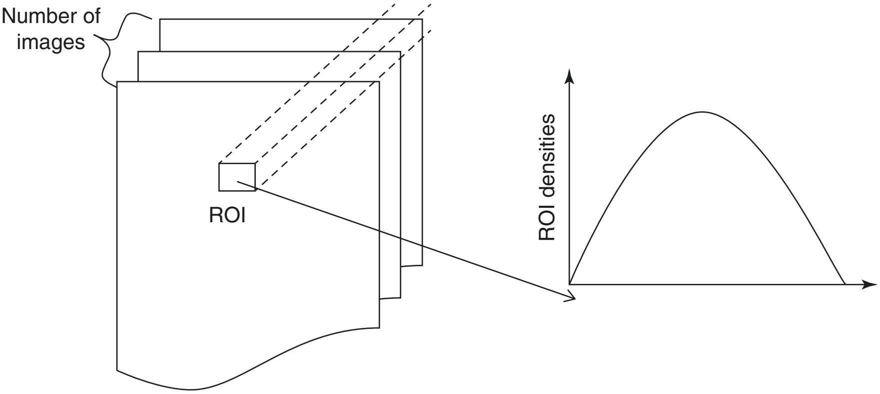


Fig. 6.5 Time density curve generation

$$D\left(kT_{ecg}\right)=\frac{-1}{\mu_0}\iint_{ROI} I_{\log\_dsa}\left(x,y,kT_{ecg}\right)dA \tag{6.27}$$

where T_{ecg} corresponds to the late diastole phase in the cardiac cycle and it leads to good results for patients with normal heart rate and with no involuntary motions during the acquisition (Schrijver 2002).

Another problem of the DSA appears in the case of a filled epicardial vessel overlaid on a variable background generated by the myocardium perfusion with contrast material. A solution to this problem is offered by (Schrijver 2002) and consists in the selection of a ROI in close proximity to the artery in order to appreciate the average background intensity which eliminates the need for the mask image subtraction.

The most relevant measures related to the filling of the coronary arteries can be determined during the first cardiac cycles before the installation of the venous washout phase.

Still, in practice, we face with different TDCs and the most important sources of inadvertent effects are (Schrijver 2002):

- Recirculation of the contrast material;
- Extravascular accumulation of the contrast material, which leads to a lower peak and a slower washout;
- The fluctuations in amplitudes due to the manual injection;
- Non-steady flows: the contrast agent does not fully mix with the blood;
- Opacification of the background structures (bones) which influence the shape of the curve.

6.3.2 Processing of Time Density Curves

Data filtering is a key signal processing operation for smoothing and removing the noise from the time density curve, before proceeding with the data fitting.

The smoothing filters remove the high frequency components, which are related to noises and preserve the medium and low frequency components. This filter resembles with the low pass filters and captures the key elements in the data, but it tends to filter out some portion of the high frequency elements along with the noise.

The Savitzky-Golay filtering uses a least square smoothing filter and it has a finite impulse response, which reduces the noises by preserving the relevant high frequency components, the maxima, the minima and the width of the curves (Savitzky and Golay 1964). It actually minimizes the least-squares error between a fitting polynomial and the local frames of the noisy data.

In Matlab, the function *sgolayfilt* implements this algorithm and it uses as input the following parameters:

- the frame size, $f = 2 * N + 1$, which is odd,
- the polynomial order M which must be smaller than f.

If M = f − 1, the filter performs no smoothing.

An example of TDC filtering with a 9th order polynomial and a frame size f of 11 produces the results depicted in the Fig. 6.6.

The filtering operation flattened only the small fluctuations in the curve, which reflects the manual injection of the contrast agent. In the following, the curve fitting is performed to smooth the remaining fluctuations and to model the data.

A mathematical model of the TDC is required in order to overcome the limitation of the temporal resolution and the noises which cannot be eliminated through smooth filtering. The resulted curve must resemble with an ideal indicator dilution curve.

(Thompson et al. 1964) discovered that, for the non-branching vessels, the time-density curves can be represented by a complex exponential called Gamma variate function. This model is a representation of the impulse response of the dilution system (Mischi et al. 2008). With only minor changes, this function can be expressed as follows:

$$G(t) = A(t - ta)^{\alpha}\, e^{\frac{-(t-ta)}{\beta}} \tag{6.28}$$

where t is the time after injection, $G(t)$ is an indicator of the concentration at the time t, $A > 0$ is a constant scale factor which reflects the intensity scaling values of the x-ray device, ta is the appearance time when the contrast agent has been injected into the ostium, α and β represent the scaling parameters of the curve which reflect the nature of the contrast agent propagation in the considered ROI.

In Fig. 6.7 an exemplification of the fitting algorithm for some raw and filtered curves samples are displayed. The initial value for alpha is set to 2.

There are also other techniques which may employed for fitting the time density curves, such as the Gaussian or the polynomial functions (Alassi 2012).

Sadasivan et al. (2002) have developed a mathematical model for the time density curves with clinical applications for predicting the likelihood of thrombus apparition after stent placement in cerebral aneurysm treatment. For the cardiac angiograms, the Gamma function sufficed.

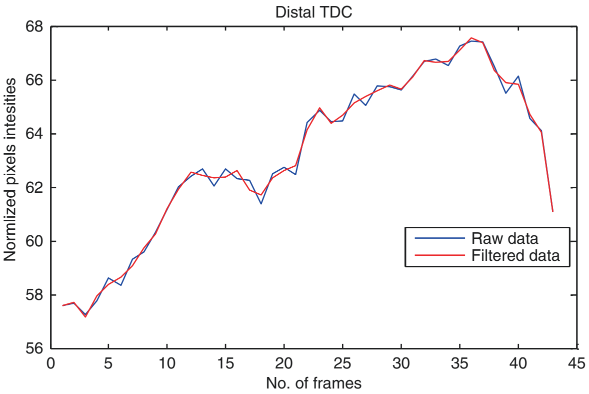


Fig. 6.6 Savitzky-Golay filtering of a TDC: *in blue*—the raw data and *in red*—the filtered data

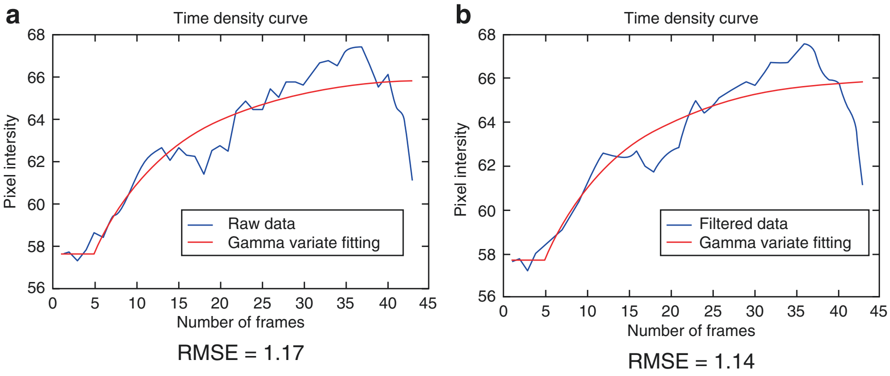


Fig. 6.7 Gamma variate fitting of the TDC for (**a**) raw data and (**b**) filtered data represented in *blue* and the Gamma fitted data in *red*

6.3.3 Computing the Transit Time

Bolus transit time or time-density algorithms are based on the tracking of the contrast material, as it travels along the blood vessel. The time required for the bolus to travel the distance between two regions of interest represents the transit time.

From the TDC key timing parameters can be extracted for the transit time estimation. These methods and their applications are summarized in Table 6.5, and can be used for transit time estimations.

Pijls et al. (1990) concluded that the mean transit time best correlates with the Doppler blood flow measurements of the circumflex artery, followed by the peak opacification time.

The transport or transit time is determined as the time difference between corresponding timing parameters determined based on the distal and proximal curve. This is mathematically expressed as:

$$\bar{t} = t_{distal} - t_{proximal} \tag{6.29}$$

where $\bar{t}$ is the transit time, t_{distal} is the bolus transport time for the distal ROI and t_{proximal} is the reference time which is usually taken as the bolus transport time of the proximal ROI, or 0.

6.3.4 Results

The above mentioned methods for estimating the transit time have been implemented and a dedicated GUI was developed to easily manipulate the data. It enables the manual placement of both ROIs, one at a proximal site and another one at a distal site of the vessel of interest. The interface offers the possibility to delete or to reiterate the selection step for an image.

Table 6.5 Alternative timing parameters and their applications (Tache 2015)

Method	Mathematical description	Applications
Mean transit time	$t_{mtt} = \dfrac{\int_0^\infty t \cdot D(t)\,dt}{\int_0^\infty D(t)\,dt}$	Internal carotid, coronaries, cerebral vessels
Mean transit time after curve thresholding	$D(t) = \begin{cases} D(t) - \Delta & \text{if } D(t) \hat{>} \Delta \ \ \textit{with} \ \ \Delta = \alpha \cdot \max D(t) \\ 0 & \text{otherwise} \end{cases}$	Coronaries
Mean transit time difference	$tmttD = tmtt - tfa$	Cerebral vessels
Time of peak opacification	$D(tmax) = Dmax$	Carotid, cerebral vessels
Time to half peak	$D(thalf_{\_\max}) = 0.5 \bullet Dmax$	Carotid and coronaries
First appearance time	$D(tfa) = 0.05 \bullet Dmax$	Coronaries, cerebral vessels
Rise time	$\Delta trt = tmax - tfa$	Coronaries
Mean concentration time	$D(tmct) = Dmean$	Iliac arteries for animals and large arteries for humans
Mean arrival time	$t_{mat} = \dfrac{1}{D_{\max}} \int_{t_{ref}}^{t_{\max}} \left[D_{\max} - D(t) \right] dt$	Coronaries
Time of peak gradient	$D(tpg) = \max(\mathrm{grad}(D(t)))$	Coronaries and renal arteries
Cross correlation method	$\phi(\Delta t) = \int_0^{t_{end}} D_{ROI1}(t - \Delta t) \cdot D_{ROI2}(t)\,dt$	Arteries

Furthermore, the user must tag the corresponding ROIs in all the frames of the series in order to generate the TDCs and to compute the distance. Next, a filtering method can be chosen, and a fitting algorithm can be employed. Their performances are visually or quantitatively evaluated using RMSE. For this study, eight patients were chosen, with different stenoses degrees.

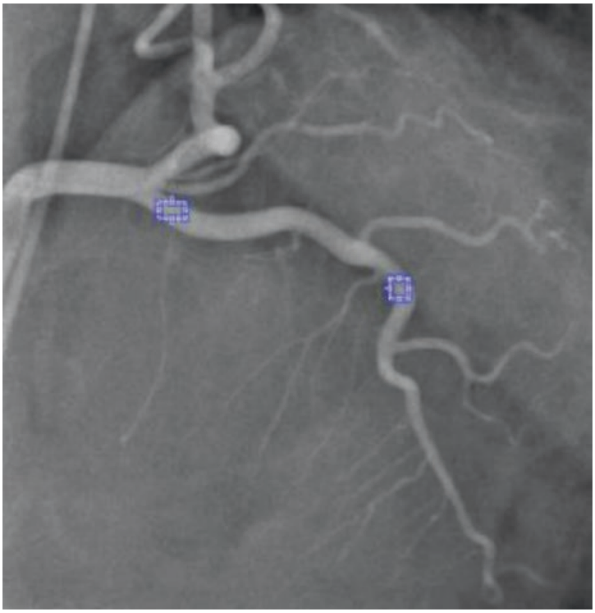


Fig. 6.8 The selection of the proximal and distal ROIs on a vessel

The two ROIs were placed on large coronary vessels, such as the left anterior descending coronary artery—LAD and/or the left circumflex coronary artery LCX and/or the right coronary artery—RCA (Fig. 6.8).

The time density curves were generated from images pertaining to the entire cardiac cycle.

In Tache et al. (2014) it was concluded that the minimal variation from the average values computed for different alternative timing parameters were found for the filtered and fitted curves. Therefore, in the current study the TDCs were filtered with a Savitzky-Golay filter and fitted with the Gamma variate function. Finally, the transit time estimations were computed for two different projections of the same vessel segment. The average transit time for the two views is presented in Table 6.6.

For the filtering step a frame size of 11 and a polynomial order of 9 were typically employed. For the Gamma variate fitting, the frame in which the contrast agent appears is manually selected by the user, due the pronounced noises of the raw curves. The alpha parameter is initially set to 2.

For a better comparison of the results, the deviation from the mean values for all timing parameters computed in the previous table, are presented in the Table 6.7.

Table 6.6 Average transit time

No. patient	1		2	3	4	5	6	7		8
Vessel methods (s)	RCA	LCX	LAD	RCA	RCA	RCA	LAD	LAD	LCX	LCX
Mean transit time	0.30	0.32	0.28	0.66	0.79	0.47	0.30	0.15	0.23	0.55
Time of peak opacification	0.93	0.70	0.93	1.20	1.40	0.87	0.70	0.30	0.50	1.53
Time to half peak	0.31	NAN	0.26	0.58	0.72	0.56	0.54	0.55	0.30	0.82
Mean concentration time	0.39	0.22	0.35	0.55	0.46	0.40	0.36	0.33	0.24	0.66
Time of peak gradient	0.35	0.25	0.32	0.44	1.05	0.60	0.40	0.60	0.47	0.99
First appearance time	0.45	−0.23	1.13	2.46	1.32	0.54	0.22	0.57	0.52	1.06
Mean arrival time	0.04	0.59	0.57	0.65	1.54	1.00	0.91	0.87	0.67	1.35
Rise time	1.87	0.96	0.41	−0.84	1.04	1.02	1.09	0.53	0.55	0.93
Cross correlation method	0.23	0.20	0.23	0.40	0.60	0.43	0.37	0.10	0.27	0.73
Mean transit time after curve thresholding	0.49	0.33	0.28	0.62	0.81	0.45	0.36	0.03	0.17	0.67
Downslope time	0.23	−0.07	−0.45	0.13	0.17	0.03	−0.45	−0.42	−0.32	NaN

Both mean transit time methods, with and without curve thresholding, downslope time and the cross correlation method underestimate the average transit time, while the rise time significantly overestimates it.

The rise time and the mean arrival time have the highest deviation from the mean value, making them less accurate.

In the case of arrival time methods (peak opacification time, half to peak time, etc.) the shape of the injection bolus influences the performances of the outcome and require a powerful injection rate of the contrast material to ensure a good mixing with the blood (Schrijver 2002).

Table 6.7 Deviation from the mean average transit time

No. patient	1		2	3	4	5	6	7		8
Vessel	RCA	LCX	LAD	RCA	RCA	RCA	LAD	LAD	LCX	LCX
Mean transit time	−0.21	−0.01	−0.11	0.04	−0.11	−0.11	−0.13	−0.18	−0.10	−0.38
Time of peak opacification	0.42	0.37	0.54	0.58	0.50	0.29	0.26	−0.03	0.17	0.60
Time to half peak	−0.19	Nan	−0.13	−0.05	−0.18	−0.02	0.10	0.22	−0.03	−0.11
Mean concentration time	−0.12	−0.10	−0.04	−0.07	−0.44	−0.18	−0.07	0.00	−0.08	−0.27
Time of peak gradient	−0.16	−0.07	−0.08	−0.19	0.15	0.02	−0.04	0.27	0.14	0.06
First appearance time	−0.06	−0.55	0.74	1.84	0.43	−0.04	−0.22	0.24	0.19	0.13
Mean arrival time	−0.47	0.26	0.18	0.03	0.64	0.42	0.48	0.54	0.35	0.42
Rise time	1.36	0.63	0.02	−1.46	0.14	0.44	0.65	0.20	0.22	0.00
Cross correlation method	−0.28	−0.13	−0.16	−0.22	−0.30	−0.15	−0.07	−0.23	−0.06	−0.20
Mean transit time after curve thresholding	−0.02	0.00	−0.12	0.00	−0.09	−0.13	−0.08	−0.29	−0.15	−0.26
Downslope time	−0.28	−0.39	−0.84	−0.49	−0.73	−0.55	−0.89	−0.75	−0.65	Nan

Conclusions

A variety of transit time methods from different types of TDC processing approaches corresponding to the coronary arteries were evaluated, and a new comparison of these timing parameters was performed for the Gamma variate fitting of the filtered time density curve.

Methods such as peak opacification time, half peak time, first appearance time and rise time are computed at a single point of the curve. Hence, they are more sensitive to noises and they are only considering straight vessels, while the other methods (mean transit time, mean concentration time, peak gradient, mean arrival time, cross correlation) are taking into the consideration the evolution of amplitudes in the curve and the results indicate that they are more robust with respect to noises, and they can account for the pulsatile flows in the large vessels.

The estimation of the transit time is an important measure for computing the contrast agent velocity and for estimating various flow indices from angiography, such as the fractional flow reserve. To obtain the absolute length or the radius, a three dimensional reconstruction of the vessel segment is required. Therefore, a method for reconstructing the vessel centerline was implemented. Because there are errors in the estimation of intrinsic parameters of the camera, and the tagging of vessel in two different views, an optimization algorithm was also presented in this chapter along with a validation and comparison.

References

Alassi S (2012) Estimating blood flow based on 2D angiographic image sequences. Master thesis, Friedrich-Alexander Universität Erlangen-Nürnberg

Brost A, Strobel N, Yatziv L, Gilson W, Meyer B, Hornegger J, Lewin J, Wacker F (2009a) Accuracy of x-ray image-based 3D localization from two C-arm views: a comparison between an ideal system and a real device. In: Medical imaging 2009: visualization, image-guided procedures, and modeling, vol 726, Lake Buena Vista, FL, SPIE, pp 72611Z-72611Z-10

Brost A, Strobel N, Yatziv L, Gilson W, Meyer B, Hornegger J, Lewin J, Wacker F (2009b) Geometric accuracy of 3-D X-ray image-based localization from two C-arm views. In: Workshop on geometric accuracy in image guided interventions—Medical Image Computing and Computer Assisted Interventions, MICCAI, pp 12–19

Lourakis MIA (2005) A brief description of the Levenberg-Marquardt Algorithm Implemened by levmar, free paper, Institute of Computer Science Foundation for Research and Technology—Hellas (FORTH), Crete, Greece

Mischi M, den Boer JA, Korsten HH (2008) On the physical and stochastic representation of an indicator dilution curve as a gamma variate, physiological measurement, 29(3), 281–294

Pijls NH, Uijen GJ, Hoevelaken A, Arts T, Aengevaeren WR, Bos HS, Fast JH, van Leeuwen KL, van der Werf T (1990) Mean transit time for the assessment of myocardial perfusion by videodensitometry. Circulation 81:1331–1340

Sadasivan C, Lieber BB, Gounis MJ, Demetrius K, Lopes DK, Hopkins LN (2002) Angiographic quantification of contrast medium washout from cerebral aneurysms after stent placement. AJNR Am J Neuroradiol 23:1214–1221

Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36:1627–1639

Schrijver M (2002) Angiographic image analysis to assess the severity of coronary stenosis. Ph.D. thesis, University of Twente

Shechter G (2004) Respiratory motion of the heart: implications for magnetic resonance coronary angiography. Ph.D. thesis, Johns Hopkins University, Baltimore, MD

Tache IA, Itu L, Niculescu R (2014) Transit time estimations from coronary angiograms. In: The 18th international conference on system theory, control and computing, 978-1-4577-1173-2, Sinaia, Romania

Tache IA (2015) Estimation of hemodynamic parameters of the cardiac and cerebral vessels from angiograms. Ph.D. thesis, University Politehnica of Bucharest

Thompson H, Starmer C, Whalen R, Mcintosh H (1964) Indicator transit time considered as a gamma variate. Circ Res 14:502–515

Wächter I (2009) 3D reconstruction of cerebral blood flow and vessel morphology from X-ray rotational angiography. Ph.D. thesis, Department of Medical Physics and Bioengineering, University College London

Yang J, Wang Y, Liu Y, Tang S, Chen W (2009) Novel approach for 3-D reconstruction of coronary arteries from two uncalibrated angiographic images. IEEE Trans Image Process 18(7):1563–1572

Zhang Z (1996) Epipolar geometry in stereo, motion and object recognition: a unified approach. Springer, New York

Zierler K (2000) Indicator dilution methods for measuring blood flow, volume, and other properties of biological systems: a brief history and memoir. Ann Biomed Eng 28:836–848

GPU-Based High Performance Computing: Employing Massively Parallel Processors for Speeding-Up Compute Intensive Algorithms

7

Constantin Suciu, Lucian Itu, Cosmin Nita, Anamaria Vizitiu,
Iulian Stroia, Laszlo Lazăr, Alina Gîrbea, Ulrich Foerster,
and Viorel Mihalef

Abstract

One-dimensional blood flow models have been used extensively for hemodynamic computations in the human arterial circulation. In this chapter we introduce a high performance computing solution based Graphics Processing Units (GPU). Novel GPU only and hybrid CPU-GPU solutions are proposed and evaluated. Physiologically sound periodic (structured tree) and non-periodic (windkessel) boundary conditions are considered, in combination with both elastic and viscoelastic arterial wall laws, and different second-order accurate numerical solutions schemes. Both the GPU only and the hybrid solutions lead to significantly smaller execution times.

Parts of Sect. 7.2 have been published before in the paper 'Graphics Processing Unit Accelerated One-Dimensional Blood Flow Computation in the Human Arterial Tree', International Journal on Numerical Methods in Biomedical Engineering, Vol. 29, December, 2013, pp. 1428–1455.

C. Suciu (✉) • L. Itu • C. Nita • A. Vizitiu • I. Stroia • A. Gîrbea
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

Automation and Information Technology, Transilvania University of Brasov,
Mihai Viteazu nr. 5, Brasov 5000174, Romania
e-mail: constantin.suciu@siemens.com

L. Lazăr
Corporate Technology, Siemens SRL, B-dul Eroilor nr. 3A, Brasov 500007, Romania

U. Foerster
Corporate Technology, Siemens AG, Otto-Hahn-Ring 6, Munich 81739, Germany

Harz University of Applied Sciences, Friedrichstrasse 57-59, Wernigerode 38855, Germany

V. Mihalef
Medical Imaging Technologies, Siemens Healthcare, 755 College Road,
Princeton, NJ 08540, USA

L.M. Itu et al. (eds.), *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*,
DOI 10.1007/978-3-319-56853-9_7

Moreover, we introduce GPU-based computationally efficient implementations for the voxelization of a mesh, for the solution of large linear system of equations using both the preconditioned conjugate gradient method, and for random forest based classification. The GPU based implementations can lead to significant execution time improvements for all these applications. This feature is based on research, and is not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

7.1 Introduction

Graphics Processing Units (GPUs) are dedicated processors, designed originally as graphic accelerators. Since CUDA (Compute Unified Device Architecture) was introduced in 2006 by NVIDIA as a graphic application programming interface (API), the GPU has been used increasingly in various areas of scientific computations due to its superior parallel performance and energy efficiency, leading to the definition of a new concept: general-purpose computing on graphics processing units (GPGPU) (Ryoo et al. 2008; Zou et al. 2009). GPGPU pipelines initially developed due to the speed-up requirements of scientific computing applications. Due to its very efficient performance-cost ratio, and its widespread availability, the GPU is currently the most used massively parallel processor.

Multiple graphics cards may be used in one computer, or GPU clusters may be employed. Nevertheless even a single CPU-GPU framework typically outperforms multiple CPUs due to the specialization of each processor (Mittal and Vetter 2015).

The GPU is a stream processor, developed for performing a large number of floating point operations (FLOPS) in parallel, by using a large number of processing units. Recent GPUs deliver the same performance as that of a cluster at 10% of the cost. The most important development has been the introduction of a programming interface for GPUs, which has transformed them into General Purpose GPUs (GPGPU) (Owens et al. 2008). The possibility of running jobs, which typically require clusters, on a single computer equipped with one or more GPUs, has drawn the attention of researchers from various research areas.

Since graphics applications require significant computing power, the GPU, as opposed to the CPU, uses the majority of its transistors for data processing and not for data cache or for execution control. This characteristic coincides with the requirements of many scientific computing applications (Chen et al. 2009). Furthermore, the GPU is well suited for parallel computations. Its stream processors work independently and concurrently at high frequencies when a program is being executed, diminishing thus significantly the need for complex execution control.

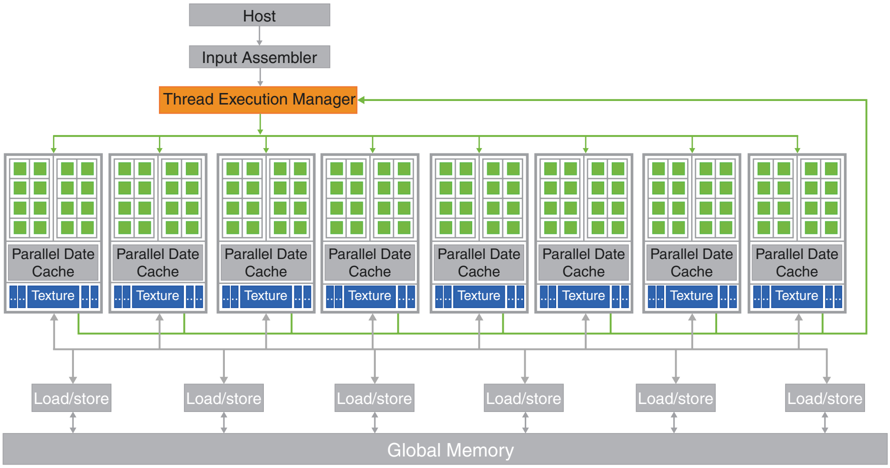


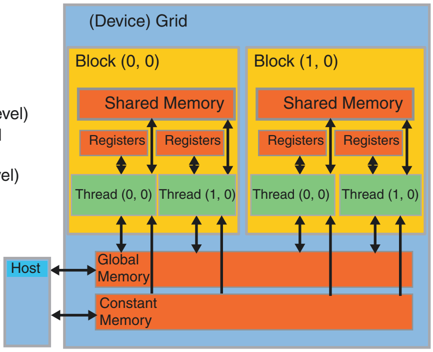
Fig. 7.1 Architecture of a GPU programmed through CUDA (Kirk and Hwu 2010)

Figure 7.1 depicts the architecture of a GPU which can be programmed through CUDA. It is composed of several streaming multiprocessors, which can execute in parallel a large number of threads. In this example, two multiprocessors form a structural block (this specific organization varies from one GPU generation to the next). Furthermore, each multiprocessor has a number of processing units which have a common control logic and instruction cache.

The G80 GPU, which was introduced together with the CUDA language, had a memory bandwidth of 86 GB/s a CPU-GPU bandwidth of 8 GB/s. Furthermore the G80 has 128 processing units (16 multiprocessors, each one with 8 processing units). The peak performance was over 500 gigaflops. Current GPUs have over 1000 processing units and exceed several teraflops in terms of processing performance.

Each processing unit can run thousands of threads in a single application. A typical GPU-based application executes 5000–15,000 threads. The G80 allows the simultaneous execution of up to 768 threads per streaming multiprocessor, and over 12,000 threads for the entire GPU. The GT200 allows the simultaneous execution of up to 1024 threads per streaming multiprocessor, and over 30,000 threads for the entire GPU. The maximum level of parallelism at hardware level has increased with the release of the Fermi, Kepler and Maxwell architectures.

In the CUDA architecture, the host and the CUDA device have separate memory spaces: CUDA devices are separate hardware components with their own DRAM. To run a kernel on the device, the programmer needs to allocate memory on the device and transfer the data of interest from the host to the device. Similarly, once the execution on the GPU has finalized, the results need to be transferred back to the CPU, and the device memory that is no longer required has to be deallocated.



The device is able to:
- read / write to / from registers (allocated at thread level)
- read / write to / from local memory (allocated at thread level)
- read / write to / from shared memory (allocated at thread block level)
- read / write to / from global memory (allocated at grid level)
- read from constant level (allocated at grid level)

The host is able to:
- read / write to / from global and constant device memory

Fig. 7.2 CUDA memory model (Kirk and Hwu 2010)

Figure 7.2 depicts an overview of the CUDA memory model. On the lower side of the figure one can observe the global and the constant memory. These are the two types of memory that can be accessed by the host in both directions (the constant memory can only be read by the CUDA device).

At thread block level, shared memory is available, which can be accessed by all threads of the same block. Typically, shared memory is employed to reduce the number of redundant accesses of the global memory. Unlike the global memory, shared memory can be accessed with low latency for reading and writing data.

Furthermore, each thread has access to a set of localized registers. These are high speed memory locations, but the number of such locations is limited.

GPUs are currently extensively being used in biomedical engineering (for fast diagnosis, simulation of surgical procedures and prediction, etc.). Image-guided therapy systems, which are gaining more traction in clinical treatment and interventions applications, are based on programs with a large level of parallelism (Shams et al. 2010). Thus, radiology departments may profit in future from GPU-based high performance computing, at reasonable costs. GPUs have been used in research activities related to the cardiovascular system: modelling and simulation of the heart (offline—to reduce the simulation time), or for the real-time visualization of 3D/4D sequences acquired through MRI or CT. Sato et al. have described the acceleration of simulations performed for understanding the propagation of electrical waves in the cardiac tissue, and have compared CPU and GPU based execution times (Sato et al. 2009). The approach based on a single GPU has shown to be 20–40 times than a single CPU based implementation. Furthermore, an interactive simulation for cardiac interventions has been proposed, which also takes into account the collision between the catheter and the inner wall of the heart (Yu et al. 2010). The GPU-based implementation is used for cardiac modelling, visualization and interactive simulation. The simulation of the cardiovascular system can also be performed with the Lattice Boltzmann method which is readily parallelizable on GPUs (Nita et al. 2013).

Tanno et al. have introduced a GPU-based implementation of the artificial compressibility method (Tanno et al. 2011). A speed-up of approx. 8× was achieved, when compared to the CPU based implementation.

Moreover, a multi-GPU simulator based on the three-dimensional Navier-Stokes equations has been introduced, which can simulate the interaction between different fluids, and which is based on a level-set approach (Zaspel and Griebel 2013). High-order finite differences schemes and the projection method were employed for the discretization in space and time. The speed-up, when compared to a CPU-based implementation, is of around 3×, while the energy consumption is approx. two times smaller.

Another research study has analyzed the GPU-based speed-up potential for a hemodynamic simulation performed for the abdominal aorta (Malecha et al. 2011). Only a part of the code has been transferred to the GPU (the solution of the linear system of equations), and a speed-up of 3–4× was obtained.

In a different approach, blood was modeled as a mix of plasma and red blood cells, and hemodynamic simulations were run (Rahimian et al. 2010). Up to 260 millions deformable red blood cells were used, and the code was developed so as to support parallelism at all levels, including internodal parallelism, intranodal parallelism through shared memory, data parallelism (vectorization), as well as a very large number of threads for efficient GPU-based execution. The most complex simulation employed 256 CPU-GPU systems, and 0.7 petaflops were achieved.

In conclusion, GPU-based approaches are already employed widely in many compute-intensive research areas, like:

- Computational fluid dynamics (Itu et al. 2013a, b)
- Machine learning (Garcia et al. 2008)
- Molecular modeling (Hasan Khondker et al. 2014)
- Astrophysics (Klages et al. 2015)
- Bioinformatics (Schatz et al. 2007; Manavski and Valle 2008)
- Medical imaging (Shams et al. 2010; Shen et al. 2009)
- Lattice Boltzmann methods (Nita et al. 2013)
- Monte Carlo simulations (Alerstam et al. 2008)
- Weather forecasting (Schalkwijk et al. 2015)
- Fuzzy logic (Cococcioni et al. 2011)

The first generation of CUDA enabled GPUs was limited to single precision floating point computations, and hence could not be used efficiently for applications with high accuracy requirements, e.g. CFD. Double precision computations were then made available in the second generation of CUDA enabled GPUs (Tesla architecture). Furthermore, the Fermi and Kepler architectures increased the double precision performance. Hence, GPUs are used nowadays widely for the acceleration of scientific computations.

The GPU is viewed as a compute device which is able to run a very high number of threads in parallel inside a kernel (a function, written in C language, which is executed on the GPU and launched by the CPU). The threads of a kernel are organized at three levels: blocks of threads are organized in a three dimensional grid at the top level, threads are organized in three-dimensional blocks at the middle level, and, at the lowest levels, threads are grouped into warps (groups of 32 threads formed by linearization of the three-dimensional block structure along the x, y and z axes respectively).

The concepts and information presented in this chapter are based on research and are not commercially available. Due to regulatory reasons its future availability cannot be guaranteed.

7.2 GPU Accelerated One-Dimensional Blood Flow Computation in the Human Arterial Tree

In previous chapters we have focused on computational approaches for modeling the flow of blood in the human cardiovascular system.

For three-dimensional blood flow models, due to the extremely high computational requirements, there has been a lot of interest in exploring high performance computing techniques for speeding up the algorithms. Although one-dimensional blood flow models are generally at least two orders of magnitude faster, the requirement of short execution times is still valid. Thus, when blood flow is modeled in patient-specific geometries in a clinical setting, results are required in a timely manner not only to potentially treat the patient faster, but also to perform computations for more patients in a certain amount of time.

As described in previous chapter, it is crucial to match the patient-specific state in a hemodynamic computation. The tuning procedure requires repetitive runs on the same geometry, with different parameter values (e.g. for inlet, outlet or wall boundary conditions (BC)), until the computed and the measured quantities match. This increases the total execution time for a single patient-specific geometry.

Another type of application for which the acceleration of the hemodynamic computations in general, and of the one-dimensional blood flow in particular, is important, has been reported in (Vardoulis et al. 2012): a total of 1000 different hemodynamic cases have been simulated (by varying geometry, heart rate, compliance or resistance) in order to find a correlation between total arterial compliance and aortic pulse wave speed. The speed-up of the execution time can either significantly reduce the time required to perform all computations for a certain correlation result, or it can allow one to perform even more computations in the same amount of time, with different configurations, to obtain better final results.

The high performance computing techniques have focused on the parallelization of the algorithms. For the implementation of the parallel algorithms cluster-based approaches, graphics processing unit (GPU) based approaches, or even a combination of the two have been explored. Since the execution time of a three-dimensional model is generally at least two orders of magnitude higher than the execution time of a one-dimensional model, most of the parallelization activities have focused on three-dimensional models (Habchia et al. 2003; Tanno et al. 2011; Zaspel and Griebel 2013). Nevertheless, the speed-up requirements mentioned above are equally valid for one-dimensional and three-dimensional models.

A previous research activity focused on the parallelization of the one-dimensional blood flow model, employing a cluster-based approach, has been introduced in (Kumar et al. 2003), with speed-up values of up to 3.5×. On the other hand, with the advent of CUDA (Compute Unified Device Architecture), several researchers have identified the potential of GPUs to accelerate biomedical engineering applications in general (Kirk and Hwu 2010), and computational fluid dynamics (CFD) computations in particular to unprecedented levels (Jiang et al. 2011).

In this section, we focus on the GPU based acceleration of the one-dimensional blood flow model and present two algorithms: a novel Parallel Hybrid CPU-GPU algorithm with Compact Copy operations (PHCGCC) and a Parallel GPU Only (PGO) algorithm (Itu et al. 2013a, b). We use a full body arterial model composed of 51 arteries and the speed-up of the two approaches is evaluated compared to both single-threaded and multi-threaded CPU implementations. The computations are performed using two different second order numerical schemes, with an elastic or viscoelastic wall model, and windkessel or structured tree boundary conditions as representative examples of physiological non-periodic and respectively periodic outlet boundary conditions.

7.2.1 Methods

The one-dimensional blood flow model is derived from the three-dimensional Navier-Stokes equations based on a series of simplifying assumptions (Formaggia et al. 2003). The governing equations ensuring mass and momentum conservation are as follows:

$$\frac{\partial A(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0, \tag{7.1}$$

$$\frac{\partial q(x,t)}{\partial t} + \frac{\partial}{\partial x}\left(\alpha \frac{q^2(x,t)}{A(x,t)}\right) + \frac{A(x,t)}{\rho}\frac{\partial p(x,t)}{\partial x} = K_R \frac{q(x,t)}{A(x,t)}, \tag{7.2}$$

where x denotes the axial location and t denotes the time. $A(x,t)$ is the cross-sectional area, $p(x,t)$ the pressure, $q(x,t)$ the flow rate, and ρ is the density. Coefficients α and K_R account for the momentum-flux correction and viscous losses due to friction respectively. For a parabolic velocity profile, $K_R = -8\pi\nu$ and $\alpha = 4/3$, with ν being the kinematic viscosity of the fluid.

A state equation, which relates the pressure inside the vessel to the cross-sectional area, is used to close the system of equations. When the vessel wall is modeled as a pure elastic material, the following relationship holds:

$$p(x,t) = \Psi_{el}(A) + p_0 = \frac{4}{3}\frac{Eh}{r_0}(x)\left(1 - \sqrt{\frac{A_0(x)}{A(x,t)}}\right) + p_0, \tag{7.3}$$

where E is the Young modulus, h is the wall thickness, r_0 is the initial radius corresponding to the initial pressure p_0, and A_0 is the initial cross-sectional area. The elastic wall properties are estimated using a best fit to experimental data (Olufsen et al. 2000).

Alternatively, a viscoelastic wall model can also be used. To include viscoelasticity, the vessel wall is considered to be a Voigt-type material (Fung 1993), for which the tensile stress depends on both the tensile strain and the time-derivative of the strain (Malossi et al. 2012):

$$p(x,t) = \Psi_{el}(A) + \Psi_{v}(A) + p_0 = \frac{4}{3}\frac{Eh}{r_0}\left(1 - \sqrt{\frac{A_0}{A(x,t)}}\right) + \frac{\gamma_S}{A\sqrt{A}}\frac{\partial A}{\partial t} + p_0, \tag{7.4}$$

where γ_S is the viscoelastic coefficient, defined by:

$$\gamma_S = \frac{T_S \cdot \tan\Phi_S}{4\pi}\frac{hE}{1-\sigma^2}. \tag{7.5}$$

Here T_S is the wave characteristic time (usually taken equal to the systolic period ~0.24 s), Φ_S is the viscoelastic angle (10°), while σ is the Poisson ratio (the material is considered to be incompressible for $\sigma = 0.5$). As in the case of elastic modeling, the viscoelastic coefficient is considered to be non-uniform in space (i.e. $\gamma_S = \gamma_S(x)$).

The presence of the viscoelastic component in Eq. (7.4) introduces an additional term in the momentum conservation equation:

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\alpha\frac{q^2}{A}\right) + \frac{A}{\rho}\frac{\partial \Psi_{el}}{\partial x} + \frac{A}{\rho}\frac{\partial \Psi_{v}}{\partial x} = K_R\frac{q}{A}. \tag{7.6}$$

Spatial and temporal dependencies of the quantities have been omitted for notational clarity. At each bifurcation, the continuity of flow and total pressure is imposed,

$$q_p = \sum_i (q_d)_i, \tag{7.7}$$

$$p_p + \frac{1}{2}\rho\frac{q_p^2}{A_p^2} = (p_d)_i + \frac{1}{2}\rho\frac{(q_d^2)_i}{(A_d^2)_i}, \tag{7.8}$$

where subscript p refers to the parent, while subscript d refers to the daughter vessels.

7.2.1.1 Boundary Conditions

Depending on the availability of in-vivo measurements and the underlying assumptions used in the modeling, researchers typically use one of the following inlet boundary condition: (1) time-varying flow profile, (2) a lumped model of the heart coupled at the inlet (Formaggia et al. 2006), or (3) a non-reflecting boundary condition like a forward running pressure wave (Mynard et al. 2012a, b), (Willemet et al. 2011). A time-varying velocity profile (or flow rate profile) can be consistently determined in a clinical setting, and is often part of the diagnostic workflow (2D/3D Phase-contrast MRI (Magnetic Resonance Imaging), Doppler ultrasound). The parameters of the lumped heart model can be computed based on non-invasively acquired flow rate and pressure values (Senzaki et al. 1996), while the third type of inlet boundary condition is generally not used in patient-specific computations.

Outlet boundary conditions may be classified as either periodic or non-periodic boundary conditions. Whereas periodic boundary conditions can only be used in steady-state computations (e.g. the patient state does not change from one heart cycle to the next—the same inlet flow rate profile is applied for each heart cycle) and require the flow information from the previous heart cycle, non-periodic boundary conditions do not have these restrictions (e.g. they can be used to model the transition from a rest state to an exercise state for a patient).

We consider two physiologically motivated boundary conditions:

1. The three-element windkessel model (WK), as a non-periodic boundary condition (Westerhof et al. 1971):

$$\frac{\partial p}{\partial t} = R_p \frac{\partial q}{\partial t} - \frac{p}{R_d \cdot C} + \frac{q\left(R_p + R_d\right)}{R_d \cdot C}, \tag{7.9}$$

where R_p is the proximal resistance, R_d is the distal resistance and C is the compliance, while p and q refer to the pressure and respectively the flow rate at the inlet of the windkessel model.

2. The structured tree model (ST) (Olufsen et al. 2000), as a periodic boundary condition. The structured tree is a binary, asymmetrical vascular tree computed individually for each outlet, composed of a varying number of vessel generations. It is terminated once the radius decreases below a preset minimum radius and its root impedance, $z(t)$, is computed recursively. The root impedance is applied at the outlet of the proximal domain through a convolution integral:

$$p\left(x,t\right) = \int_{t-T}^{t} q\left(x,\tau\right) z\left(x,t-\tau\right) d\tau, \tag{7.10}$$

where T is the period. To apply a periodic boundary condition, the flow history is stored and a multiply-sum scan operation is performed at each time-step, leading to considerably higher execution times than for a non-periodic boundary condition.

We emphasize the fact that this choice is not mandatory, the windkessel model can also be applied as a periodic boundary condition, and, as recently described, even the structured tree boundary condition can be applied as a non-periodic boundary condition (Cousins et al. 2013).

7.2.1.2 Numerical Solution of the One-Dimensional Blood Flow Model

When an elastic wall model is used, Eqs. (7.1)–(7.3) represents a hyperbolic system of equations. When a viscoelastic wall model is used, the hyperbolic nature of the equations is lost due to the additional term in the pressure-area relationship. The approaches for the numerical solution of the one-dimensional equations can be divided into two main categories:

1. Methods which do not exploit the original hyperbolic nature of the equations: discontinuous finite element Galerkin method with stabilization terms (Raghu et al. 2011); implicit finite difference/spectral element method where the non-linear terms are solved iteratively at each time-step using the Newton method (Reymond et al. 2011), (Bessems et al. 2008), etc.;
2. Methods which exploit the hyperbolic nature of the equations in case an elastic wall law is used (Olufsen et al. 2000), (Mynard and Nithiarasu 2008), or which recover the original hyperbolic nature of the equations in case a viscoelastic wall law is used, by employing an operator-splitting scheme for the momentum equation. This method has been originally proposed in Formaggia et al. (2003) for a single vessel and subsequently used in Passerini (2009) and Alastruey et al. (2011).

Implicit methods, although solvable with larger time-steps, are slower since they require the solution of a system of equations at each time step and, additionally, require the application of the Newton method for the non-linear terms (Reymond et al. 2011). On the other hand, the methods that exploit the hyperbolic nature of the equations are explicit. They are computationally faster, in spite of the time-step limitation imposed by the CFL condition (named after Courant, Friedrich and Lewy (Courant et al. 1928)).

In addition to the fact that the explicit methods are faster in a sequential implementation, their explicit nature also enables them to be parallelized, and thus be implemented on a cluster (Kumar et al. 2003), or on a GPU.

The explicit methods are based either on a first order method (the method of characteristics), or on a second order method (two-step Lax Wendroff (Olufsen et al. 2000), or expansion in Taylor series (Mynard and Nithiarasu 2008)). Due to their higher accuracy, second-order methods are preferred and have been used for the current study. The method of characteristics is used at the inflow, bifurcation and outflow points.

7.2.1.3 Numerical Solution of the Elastic One-Dimensional Model

First, Eqs. (7.1) and (7.2) are written in conservation form:

$$\frac{\partial \mathbf{U}}{\partial t}+\frac{\partial \mathbf{R}}{\partial x}=\mathbf{S}, \tag{7.11}$$

$$\mathbf{U}=\begin{pmatrix} A \\ q \end{pmatrix} \quad \mathbf{R}=\begin{pmatrix} R_1 \\ R_2 \end{pmatrix}=\begin{pmatrix} q \\ \alpha \dfrac{q^2}{A}+B \end{pmatrix} \quad \mathbf{S}=\begin{pmatrix} S_1 \\ S_2 \end{pmatrix}=\begin{pmatrix} 0 \\ -K_R \dfrac{q}{A}+\dfrac{\partial B}{\partial r_0}\dfrac{dr_0}{dx} \end{pmatrix},$$
$$B\left(r_0\left(x\right),p\left(x,t\right)\right)=\frac{1}{\rho}\int_{p_0}^{p(x,t)}\Psi_{el}^{-1}\left(p'\right)dp', \tag{7.12}$$

where ***U*** is the vector of the unknown quantities, ***R*** is the flux term, and ***S*** is the right hand side (RHS).

The Lax-Wendroff (LW) scheme consists of two main steps:

Step 1. Computation of the half step-points: these values are computed between the grid points, hence there are only interior half-step values:

$$\mathbf{U}_j^{n+1/2}=\frac{\mathbf{U}_{j+1/2}^n+\mathbf{U}_{j-1/2}^n}{2}+\frac{\Delta t}{2}\cdot\left(-\frac{\mathbf{R}_{j+1/2}^n-\mathbf{R}_{j-1/2}^n}{\Delta x}+\frac{\mathbf{S}_{j+1/2}^n+\mathbf{S}_{j-1/2}^n}{2}\right), \tag{7.13}$$

where $j=m\pm 1/2$ and m refers to the grid points;

Step 2. Computation of the full-step-points: this step uses values both from the previous time step and from the half-step points:

$$\mathbf{U}_m^{n+1}=\mathbf{U}_m^n-\frac{\Delta t}{\Delta x}\left(\mathbf{R}_{m+1/2}^{n+1/2}-\mathbf{R}_{m-1/2}^{n+1/2}\right)+\frac{\Delta t}{2}\cdot\left(\mathbf{S}_{m+1/2}^{n+1/2}+\mathbf{S}_{m-1/2}^{n+1/2}\right)\cdot\left(\mathbf{S}_{m+1/2}^{n+1/2}+\mathbf{S}_{m-1/2}^{n+1/2}\right). \tag{7.14}$$

The expansion in Taylor series (TS) scheme consists of a single step:

$$\frac{\mathbf{U}^{n+1}-\mathbf{U}^n}{\Delta t}=\mathbf{S}^n-\frac{\partial \mathbf{R}^n}{\partial x}-\frac{\Delta t}{2}\left[\frac{\partial}{\partial x}\left(\mathbf{R}_{\mathbf{U}}^n\mathbf{S}^n-\mathbf{R}_{\mathbf{U}}^n\frac{\partial \mathbf{R}^n}{\partial x}\right)-\mathbf{S}_{\mathbf{U}}^n\frac{\partial \mathbf{R}^n}{\partial x}-\mathbf{S}_{\mathbf{U}}^n\mathbf{S}^n\right], \tag{7.15}$$

where all the spatial derivatives are discretized using central difference schemes, and:

$$\mathbf{R}_{\mathbf{U}}^n=\frac{\partial \mathbf{R}}{\partial \mathbf{U}}; \quad \mathbf{S}_{\mathbf{U}}^n=\frac{\partial \mathbf{S}}{\partial \mathbf{U}}. \tag{7.16}$$

Both numerical schemes require the apriori computation of the flux and RHS terms.

7.2.1.4 Numerical Solution of the Viscoelastic One-Dimensional Model

An operator splitting scheme is employed for the momentum equation in order to recover the hyperbolic nature of the equations. Thus, Eq. (7.6) is rewritten as:

$$\frac{\partial q}{\partial t}+\frac{\partial R_2}{\partial x}-\frac{A}{\rho}\frac{\partial}{\partial x}\left(\gamma\frac{\partial q}{\partial x}\right)=S_2. \tag{7.17}$$

The equation is no longer hyperbolic and cannot be cast into conservative form. The splitting scheme assumes that the contribution of the viscoelastic term is small compared to the contribution of the elastic term. The flow rate is considered to be composed of an elastic and a viscoelastic component ($q=q_e+q_v$), and Eq. (7.17) is split into two equations:

$$\frac{\partial q_e}{\partial t}+\frac{\partial R_2}{\partial x}=S_2. \tag{7.18}$$

$$\frac{\partial q_v}{\partial t}-\frac{A}{\rho}\frac{\partial}{\partial x}\left(\gamma\frac{\partial q}{\partial x}\right)=0. \tag{7.19}$$

Consequently, the numerical solution at each step is composed of two sequential sub-steps:

Step 1. The system composed of Eqs. (7.1) and (7.18) is solved, yielding the quantities $A(x,t)$ and $q_e(x,t)$.

Step 2. Equation (7.19) is solved to obtain $qv(x,t)$ and thus the total flow rate $q(x,t)$:

$$\frac{q_v^{n+1}-q_v^{n}}{\Delta t}-\frac{A^{n+1}}{\rho}\frac{\partial}{\partial x}\left(\gamma\frac{\partial\left(q_e^{n+1}+q_v^{n+1}\right)}{\partial x}\right)=0. \tag{7.20}$$

Equation (7.20) is discretized using a central difference scheme, leading to a tridiagonal system of equations, which can be readily solved using the Thomas algorithm in a sequential program. For the viscoelastic component of the flow, homogeneous Dirichlet boundary conditions are imposed at the boundaries of each vessel.

Both the LW and the TS schemes can be used to compute $A(x,t)$ and $qe(x,t)$, but because the LW method is composed of two steps, it requires the computation of the viscoelastic correction term twice for each time step. Since this would significantly increase the total execution time, we applied only the TS scheme when a viscoelastic wall law was enforced. The various computational setups, for which different parallelization strategies have been adopted, are displayed in Table 7.1. For performance comparison, we considered both a Single-threaded CPU only (SCO) algorithm and a Multi-threaded CPU only (MCO) algorithm. The MCO algorithm represents a parallel version of the SCO algorithm, implemented using openMP.

Table 7.1 Computational setups for which the speed-up obtained through GPU-based parallel implementations is investigated

Case	Numerical scheme	Wall law	Outlet BC
1	Lax-Wendroff	Elastic	Windkessel
2	Lax-Wendroff	Elastic	Structured tree
3	Taylor series	Elastic	Windkessel
4	Taylor series	Elastic	Structured tree
5	Taylor series	Viscoelastic	Windkessel
6	Taylor series	Viscoelastic	Structured tree

Table 7.2 Execution time and corresponding percentage of total execution time for the computational steps of the numerical solution of the one-dimensional blood flow model

	Case 5		Case 6	
Computational step	Time (s)	Perc. of total time (%)	Time (s)	Perc. of total time (%)
Interior grid points	357.12	46.12	357.67	30.32
Inflow grid point	0.08	0.01	0.14	0.01
Bifurcation grid points	27.66	3.57	27.61	2.34
Outflow grid points	4.96	0.64	401.72	34.06
Viscoelastic comp.	367.43	47.45	367.55	31.16
Other operations	17.12	2.21	24.87	2.11

Before introducing the parallel implementation of the one-dimensional blood flow model, we first analyze the execution time of the SCO algorithm. Table 7.2 displays the execution time of the different parts of the SCO algorithm, for cases 5 and 6 from Table 7.1 (these two cases were chosen because they contain all computational steps and both types of outlet boundary conditions are considered). Execution times are obtained for the arterial tree described in Sect. 7.3 and correspond to the computation for ten heart cycles. When the WK boundary condition is used, approx. 93% of the time is spent on the computation at the interior grid points and on the viscoelastic terms of the flow rate. Since the numerical solution for the interior grid points is explicit, this part can be efficiently parallelized on a manycore architecture (like the one of a GPU device). Though the computation of the viscoelastic terms employs a sequential algorithm, it can also be efficiently parallelized on a manycore architecture, as shown in Zhang et al. (2010). Furthermore, the computation at the bifurcation and outflow grid points is also parallelizable, but due to the low number of grid points of these types, usually below 100 for an arterial tree, the implementation on a manycore architecture is not efficient. Other operations (initialization activities, writing results to files during the last heart cycle, etc.) account for 2.21% of the total execution time and are not parallelizable. As a result, operations which occupy 93.57% of the total execution time for case 5 are efficiently parallelizable. The difference in terms of execution time between case 5 and case 6 is primarily due to the outlet boundary condition, which requires a multiply-sum scan operation at each time step. Since this operation is efficiently parallelizable on a manycore architecture (Sengupta et al. 2008), the operations which occupy 95.54% of the total execution time for case 6 are efficiently parallelizable.

For the MCO algorithm, the computation on the interior, bifurcation and outflow points, as well as the computation of the viscoelastic component of the flow rate, can be efficiently parallelized since the number of cores is much smaller for a multicore architecture than for a manycore architecture. This is achieved by associating different arterial segments and bifurcation points to distinct cores.

The results in Table 7.2 show that if the LW scheme were used for a viscoelastic wall law (case in which the viscoelastic correction term would be computed twice at each time step), the total execution would increase by 30–50%, depending on the computational setup.

We conclude that the implementation of the numerical solution of the one-dimensional blood flow model is efficiently parallelizable on a manycore architecture (like the one of a GPU device), regardless of the computational setup.

7.2.1.5 Parallelization of the Numerical Solution

We propose a parallel implementation of the one-dimensional blood flow model, based on a GPU device, programmed through CUDA (Kirk and Hwu 2010).

The numerical scheme of the interior grid points of each vessel is efficiently parallelizable and we considered two different implementation approaches:

1. A Parallel Hybrid CPU–GPU (PHCG) algorithm, whereas the unknown quantities at the interior points are computed on the GPU and the inflow/bifurcation/outflow points (called in the following junction points) are computed on the CPU. The advantage is that each device is used for computations for which it is best suited (CPU—sequential, GPU—parallel), but the disadvantage is that memory copies are required at each time step in order to interchange the values near the junction points;
2. A Parallel GPU Only (PGO) algorithm, whereas all grid points are computed on the GPU and the CPU is only used to initialize and to control the execution on the GPU. The advantage is that no memory copies between the CPU and the GPU are required, but the disadvantage is that less parallelizable operations need to be performed on the GPU.

7.2.1.6 Parallel Hybrid CPU-GPU (PHCG) Algorithm

First we refer to the implementation used in case an elastic wall law is applied. Starting from the numerical schemes described in the previous section, we applied the general workflow displayed in Fig. 7.3a.

The CPU is called host, while the GPU is called device. First, the arterial model is initialized (host memory is allocated for each grid point, initial radius, initial cross-sectional area, wall elasticity, derivatives of radius and wall elasticity are computed) and the device memory is allocated and initialized. Next, a while loop is started which advances the entire model in time for a given number of iterations and heart cycles. Inside the while loop, the host and device thread are executed in parallel until a synchronization barrier is reached. During the parallel activities, the

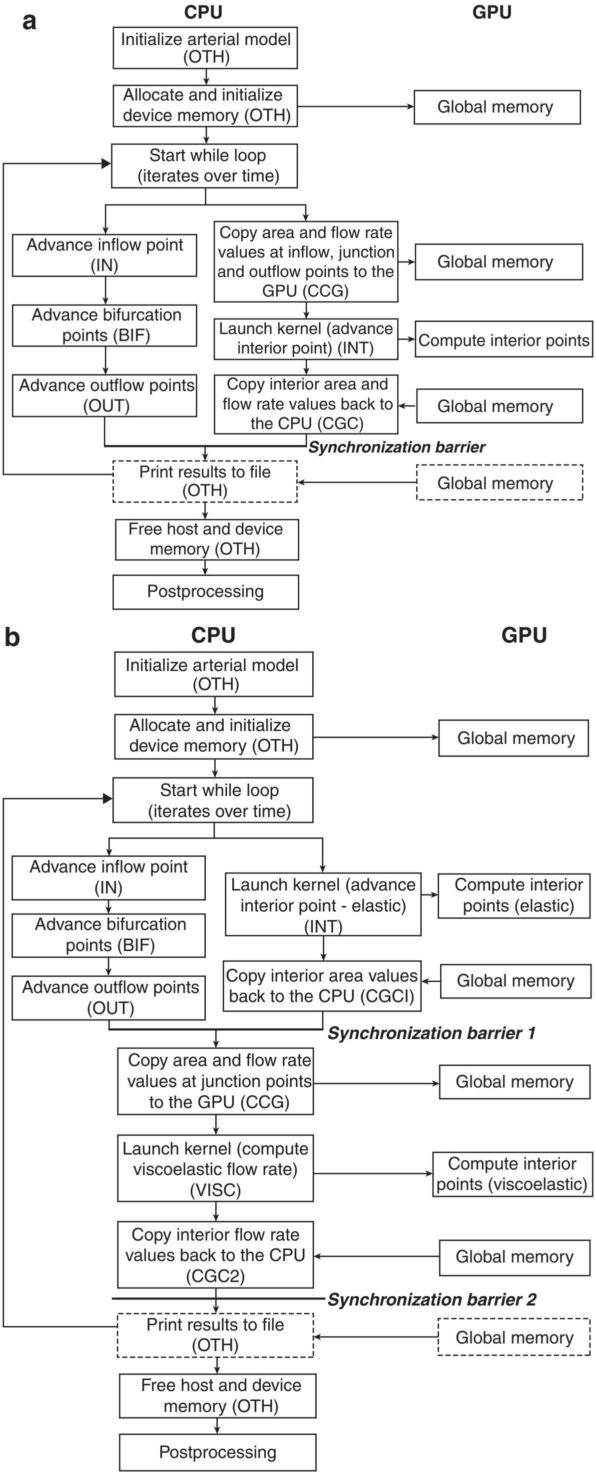


Fig. 7.3 PHCG workflow in case (**a**) an elastic wall law (Itu et al. 2012a, b), or (**b**) a viscoelastic wall law is used. Junction points are solved on the CPU, while interior points are solved on the GPU. Memory copies are required at each iteration in order to exchange the values near and at the junction points (Itu et al. 2013a, b)

CPU computes the new values at the junction points and the device performs the computations for the interior points (Eqs. (7.13), (7.14) for the LW scheme and (7.15) for the TS scheme). Since the device operations are asynchronous, no special approach is required to achieve the task level parallelism between the CPU and the GPU. The computation of the junction points on the CPU is parallelized using openMP for all PHCG implementations. An acronym is displayed in Fig. 7.3 for each operation to easily match the execution times discussed in the next section with the operations (e.g. OTH stands for *Other operations*, which comprise several activities).

To compute the junction points, the host code requires the values at the grid points next to the junction points, from the previous time step. To compute the values at the grid points next to the junction points, the device code requires the values at the junction points, also from the previous time step. Hence, to exchange the values at or next to the next junction points, memory copy operations between the device and the host are performed at the beginning and the end of each iteration. A synchronization barrier is introduced after each iteration to ensure that the copy operations have finished. During the last cycle of the computation, after convergence has been reached, the results are saved to files for visualization or post processing. Since the number of iterations for each heart cycle is very high (18,000 for a grid space of 0.1 cm), the results are saved only after a certain number of iterations (every 20–50 iterations).

To improve the execution time on the GPU, the kernel has been optimized. The specific goal has been to lower the global memory requirement. This approach is necessary to assure efficient kernel performance, even for smaller arterial trees, where parallelism is not pronounced. To reduce global memory operations, memory accesses are coalesced (the global memory accesses performed by threads of the same warp (group of 32 threads) are both sequential and aligned). To obtain aligned memory accesses all global memory arrays have been padded. Furthermore, to avoid redundant accesses performed by different threads, intermediate results are stored in the shared memory of the multiprocessor.

The execution configuration of the kernel which computes the interior grid points is organized as follows: each thread is responsible for one grid point, a block of threads is defined for each arterial segment and both the block and the thread grid are one-dimensional. The numerical solution of a one-dimensional arterial tree, as described in the previous section, is a domain decomposition approach. Hence, data is exchanged between two arterial segments only at the interfaces of the domains. Since for the PHCG algorithm the junction points are solved on the GPU and there is no communication and synchronization requirement between the thread blocks, the association between one thread block and one arterial segment is natural. Furthermore, since parallelism is limited (the number of interior grid points in an arterial tree is usually below 10,000 when a grid space of 0.1 cm is used) and the computational intensity is high (the kernel which computes the interior grid points is limited by the instruction throughput—see Sect. 7.4), an approach for which one thread may compute the unknown quantities of several grid points has not been considered. We split an arterial segment into several domains if the hardware resources of a streaming multiprocessor were insufficient to run the corresponding thread block (the solution variables at the interfaces between the domains of the same arterial segment were determined by enforcing continuity of flow rate and total pressure).

An important aspect for the PHCG algorithm is the data transfer between host and device. Although, the amount of data to be transferred is low (only the values at or next to the junction points are exchanged), the total execution time required for these data transfers is high. This is due to the high number of copy operations and the fact that the locations to be copied are scattered throughout the memory arrays. Three different approaches, displayed in Fig. 7.4, are evaluated for decreasing the total execution time and have led to three different variants of the PHCG algorithm:

1. PHCG Copy Separately (PHCGCS): each location is copied separately, resulting in a total of eight copy operations for each arterial segment at each iteration. Figure 7.4a displays the locations for which the values are exchanged at each iteration as well as the direction of the copy operations. The arrays displayed in this figure are generic and correspond to either the cross-sectional areas or the flow rates of a single blood vessel;
2. PHCG Copy All (PHCGCA): the entire arrays used for cross-sectional area and for flow rate are transferred at each iteration: four copy operations at each iteration (Fig. 7.4b);
3. PHCG Copy Compact (PHCGCC): additional arrays are allocated for the locations which are copied: four copy operations at each iteration. Figure 7.4c displays the additional arrays which need to be allocated and the locations which read/write to these arrays. For an arterial network, the values of all dependent variables of one type (cross sectional area or flow rate) are stored in a single array.

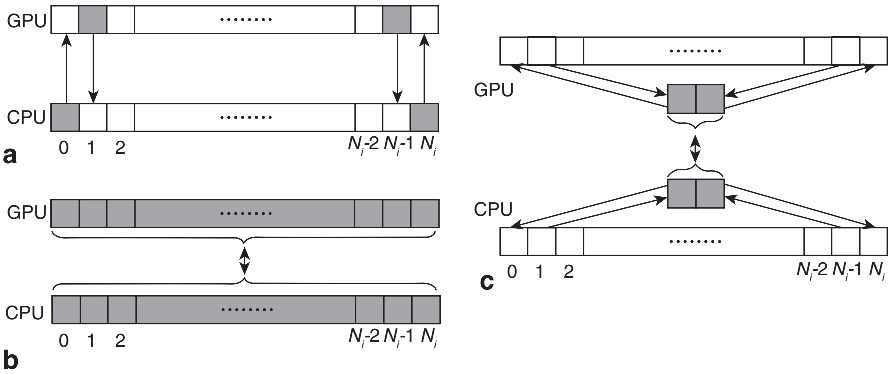


Fig. 7.4. Host ↔ Device Memory copy variants: (**a**) separate copy operation for each location, (**b**) copy entire array, (**c**) copy compact additional arrays (Itu et al. 2013a, b)

The first two memory copy strategies were introduced previously (Itu et al. 2012a, b), while the third one is developed during the current research activity and represents a combination of the first two strategies. The PHCGCS variant minimizes the amount of memory to be copied; the PHCGCA minimizes the number of copy operations, while the PHCGCC variant minimizes both aspects by trading kernel performance for data transfer performance (some threads of the kernel populate the additional arrays displayed in Fig. 7.4c).

Figure 7.5 displays the kernel operations and the shared memory arrays used for the two previously described numerical schemes (LW and TS). Since neighboring threads access the same $q/A/\boldsymbol{R}/\boldsymbol{S}$ values, shared memory is used to avoid redundant global memory reads and redundant computations. The operations of the LW scheme (Fig. 7.5a) are based on Eqs. (7.13) and (7.14) and require only four shared memory arrays (the shared memory is dynamically allocated and the size of the arrays is equal to the number of grid points of the longest vessel). The operations of the TS scheme (Fig. 7.5b) are based on Eq. (7.15) and use 11 shared memory arrays. The shared memory requirement is much higher for the TS scheme since: (1) the computations are performed in a single step (the arrays cannot be reused), and (2)

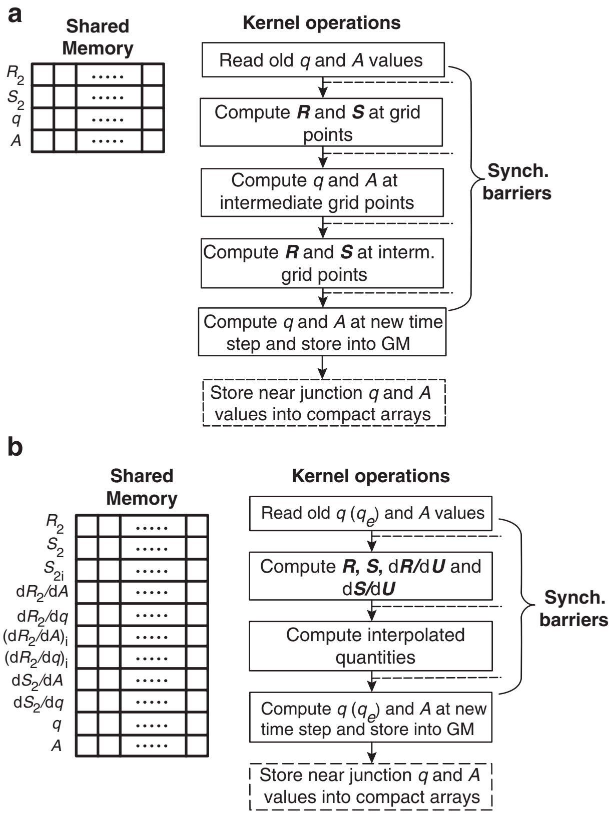


Fig. 7.5 Kernel operations and shared memory arrays used for the computation of new values at the interior grid point using (**a**) the LW scheme (Itu et al. 2012a, b), and (**b**) the TS scheme (Itu et al. 2013a, b)

Eq. (7.15) uses the derivates of $\boldsymbol{R}$ and $\boldsymbol{S}$ with respect to q and A (the quantities terminated with subscript i are computed by interpolation at locations between the grid points). If a viscoelastic wall law is enforced, the kernel displayed in Fig. 7.5b is used to compute the cross-sectional area values and the elastic component of the flow rate. The last operation of each kernel is displayed in a dashed rectangle since it is only performed for the PHCGCC variant of the PHCG algorithm. If the PHCGCC algorithm is used, the values corresponding to the last time step are read either from the regular arrays, or from the compact arrays displayed in Fig. 7.4c during the first operation of each of the two kernels displayed in Fig. 7.5. Synchronization barriers are used between the individual steps if, during the subsequent step, threads access values computed by other threads (these values are typically stored in the shared memory arrays). The synchronization barriers displayed in Fig. 7.5 are inserted at GPU thread block level (using __syncthreads()), while the synchronization barriers displayed in Fig. 7.3 are inserted at CPU level (using cudaDeviceSynchronize()).

The PHCG workflow introduced previously in Itu et al. (2012a, b), and reviewed in Fig. 7.3a cannot be used with a viscoelastic wall law. This is due to the additional steps required by the operator splitting scheme employed for this type of wall law. Consequently, we have introduced a new workflow, as illustrated in Fig. 7.3b. Two different kernels are used: one for the computation of the cross-sectional area and of the elastic flow rate: Eqs. (7.1) and (7.18); and a second one for the computation of the viscoelastic flow rate: Eq. (7.20).

The execution configuration of the first kernel is the same as in the case of an elastic wall law. Host and device instructions are executed in parallel at the beginning of each iteration. After a first synchronization barrier, the values at or next to the junction points are interchanged in order to prepare the computation of the viscoelastic flow rate (in Eq. (7.20) the new values of the cross-sectional area and of the elastic flow rate at all grid points are required), followed by the computation of the viscoelastic and the total flow rate. To solve the tridiagonal system of equations on the device, we employed an optimized CR (Cyclic Reduction)–PCR (Parallel Cyclic Reduction) algorithm (Zhang et al. 2010). Finally, the new flow rate values next to the junction points are copied back to the host and a second synchronization barrier is introduced at the end of the iteration.

For the kernel which computes the viscoelastic flow rate we use an execution configuration with a number of blocks equal to the number of arterial segments. The number of threads of each block is set equal to the smallest power of two value which is higher than the number of grid points in the longest arterial segment. This enables an efficient execution of the CR-PCR algorithm on the GPU.

Figure 7.6 displays the kernel and the shared memory arrays used for the computation of the viscoelastic component of the flow rate and of the total flow rate. First the tridiagonal system is set up (i.e. the coefficients of the three diagonals and of the RHS are computed). The CR-PCR algorithm is composed of three main steps, two forward reduction (CR and PCR, respectively) and one backward substitution (CR) step. Next, the total flow rate is determined and the new flow rate values are stored in the compact arrays if the PHCGCC algorithm is used.

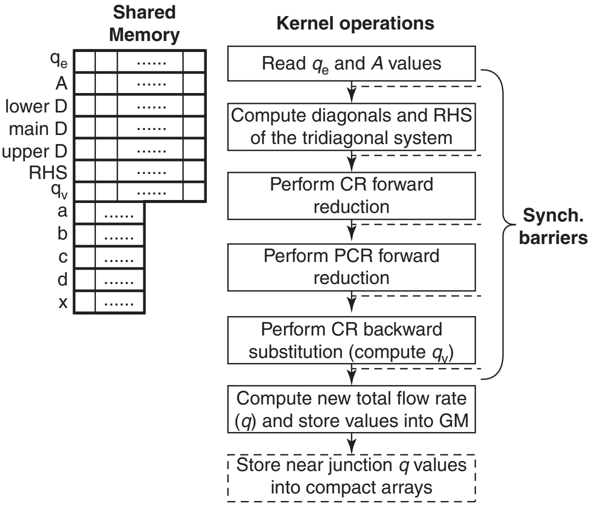


Fig. 7.6 Kernel operations and shared memory arrays used for the computation of the viscoelastic component of the flow rate and of the total flow rate (Itu et al. 2013a, b)

7.2.1.7 Parallel GPU only (PGO) Implementation

The necessity to perform copy operations at each iteration reduces significantly the overall performance of the PHCG algorithm. Hence, we performed an implementation whereas all grid points are computed on the device. This eliminates the memory copies (only the memory copies at the print iterations are required), but also forces the device to perform less parallelizable computations required for the junction points. Another disadvantage of the PGO algorithm, compared to the PHCG algorithm, is that since all operations are performed on the GPU, the task-level parallelism between CPU and GPU is lost. Figure 7.7a displays the workflow for the most complex case, namely when a viscoelastic wall law is used together with the ST boundary condition. A maximum of three kernels are executed at each iteration:

1. Computation of the convolution integral (a multiply-sum scan operation (Sengupta et al. 2008)): Eq. (7.10);
2. Computation of the new cross-sectional area and of the elastic flow rate: Eqs. (7.1) and (7.18);
3. Computation of the viscoelastic flow rate: Eq. (7.20).

The execution configuration of the first kernel is organized as follows: the number of blocks is equal to the number of arterial segments and the number of threads is set to 512. Since the number of time steps per heart cycle (which varies between 8000 and 38,000 for different grid space values) is much higher than the number of threads per block, first each thread performs multiple multiply-sum operations and stores the result in a static shared memory array (composed of 1024 double precision elements). Finally the threads perform a scan operation for the shared memory array and store the result in the global memory. The execution configuration of the other two kernels is the same as the one described in the previous section.

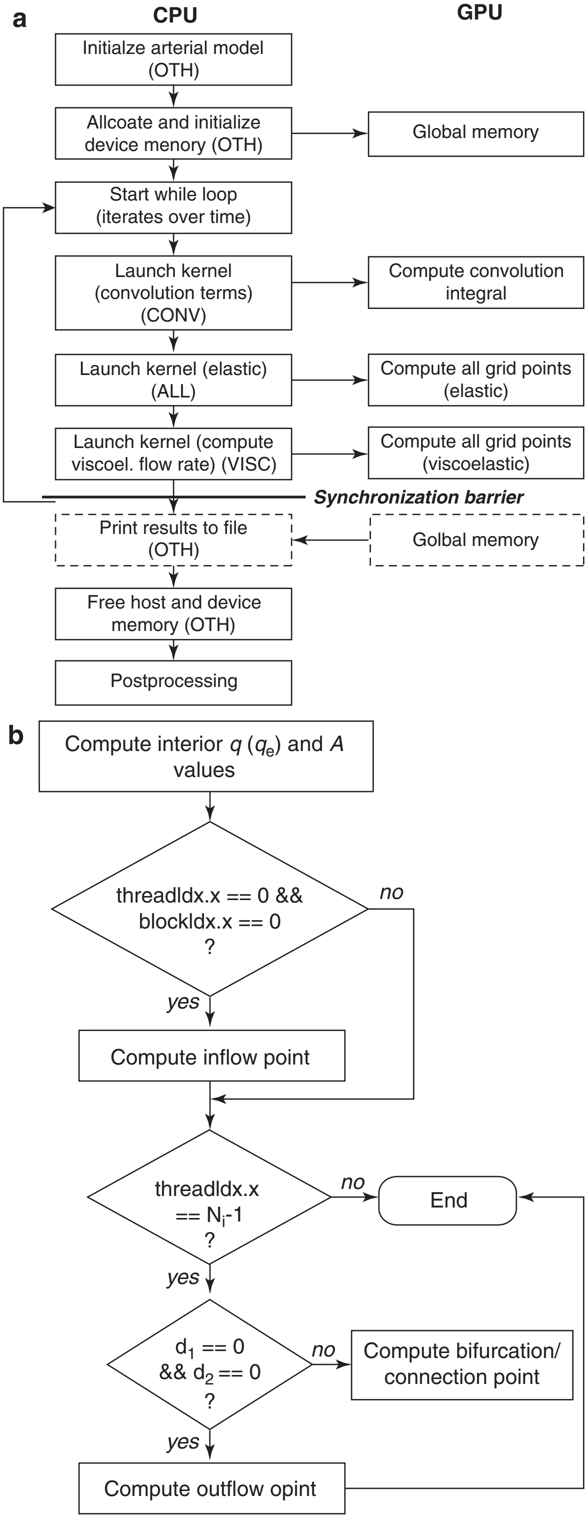


Fig. 7.7 (**a**) Generic GPU workflow when a structured tree boundary condition is used and a viscoelastic wall law is enforced. All of the computations are performed inside GPU kernels and the CPU only coordinates the operations; (**b**) Kernel operations used for the computation of the new values at all grid points (Itu et al. 2013a, b)

If the WK boundary condition is used, the first kernel is not called, and if an elastic wall law is used the third kernel is not called. An acronym is displayed in Fig. 7.7 for each operation to easily match the execution times discussed in the next section with the operations.

Figure 7.7b displays the kernel operations used to compute the new cross-sectional area and flow rate values at all grid points of a vessel segment, with a focus on the junction points. First, the interior points are computed as displayed in Fig. 7.5 (the individual operations have not been detailed). Next, the first thread of the first block solves the inlet point and the last thread of each block solves the outlet or the bifurcation/connection point (a connection point is also a junction point, which is introduced if an arterial segment is split into several domains). Thus, for the junction points, parallelism is only present at block level and not at thread level.

7.2.2 Results

Blood was modeled as an incompressible Newtonian fluid with a density of $\rho = 1.055$ g/cm^3 and a dynamic viscosity of $\mu = \nu \bullet \rho = 0.045$ dynes/cm^{2}s for all the computations.

To compare the performance of the different algorithms (SCO, MCO, three PHCG variants and PGO), the arterial tree detailed in Stergiopulos et al. (1992), and displayed in Fig. 7.8 was used. It is composed of 51 arteries. A time-varying flow rate profile was imposed at the inlet (Olufsen et al. 2000), and for the outlets, the WK and the ST boundary conditions were applied (the parameter values displayed in Table 7.3 were used). The total resistance and the compliance values were set as in Bessems (2008), and the minimum radius used for the generation of the structured tree was tuned ad-hoc so as to obtain a similar total resistance as for the WK outlet boundary condition (total resistance:1.37×10^3 dynes s/cm^5). This aspect allowed us to adequately compare the time-varying flow rate and pressure profiles obtained with the two types of physiologically motivated boundary conditions.

For the ST boundary condition we used an exponential factor equal to 2.7. The constants characterizing the asymmetry of the binary tree were set to 0.908 and 0.578, and the length-to-radius ratio was equal to 50. The elastic properties of the wall were set equal for both the proximal domain and for the structured trees. Together with the minimum radius at which the structured tree is terminated, these parameters determine the compliance of the boundary condition.

The single-threaded CPU algorithm (SCO) was executed on single Intel i7 CPU core with 3.4 GHz, the multi-threaded CPU algorithm (MCO) was executed on an eight-core i7 processor, while for the parallel algorithms (PHCG, PGO) a NVIDIA GPU GTX680 (1536 cores on 8 streaming multiprocessors with 192 cores, 48 kB of shared memory and 64K registers) was used (the GTX680 is based on the Kepler architecture). All computations were performed with double precision floating-point

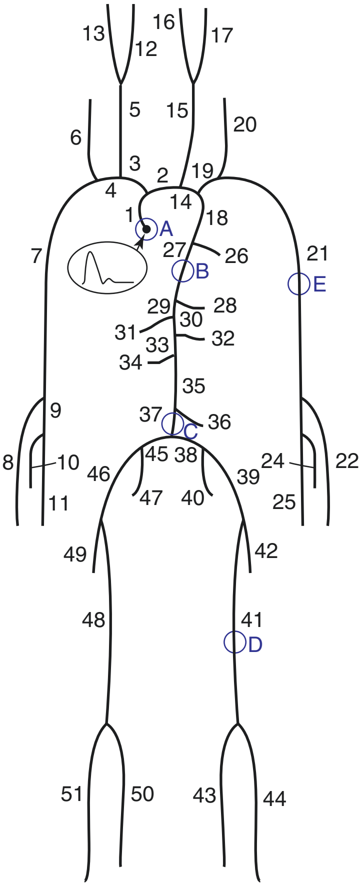


Fig. 7.8 Representation of the 51 main arteries in the human arterial system; the artery numbers of the outlet segments correspond to those displayed in Table 7.3 (Itu et al. 2013a, b).

data structures, since single precision would affect the accuracy of the results, especially at the junction points where the method of characteristics is applied based on the Newton method.

In the following subsections, we discuss the performance of the parallel algorithms and the simulation results obtained under different computational setups.

7.2.2.1 Comparison of Parallel and Sequential Computing and with Different Numerical Schemes

Taking the results determined with the SCO algorithm as reference, we computed the L_2 norms of the absolute differences between the reference numerical solution and the numerical solution obtained with the PHCG and PGO algorithms. All L_2 norm results were smaller than 10^{-13}, i.e. close to the precision of the double-type value in computer data structures (both numerical schemes, LW and TS, were used, but differences were only computed between results obtained with the same numerical scheme).

Table 7.3 Parameters of the outlet vessels used for the Windkessel boundary condition (R_p, R_d, C) and for the structured tree boundary condition (r_{min})

Art. Nr.	r_{top} (cm)	r_{bot} (cm)	Length (cm)	R_p [g/(cm^4 s)]	R_d [g/(cm^4 s)]	C (10^{-6} cm^4 s^2/g)	r_{min} (cm)
6	0.188	0.183	14.8	8.693	28.007	58.7	0.00235
8	0.174	0.142	23.5	17.165	61.434	25.9	0.00182
10	0.091	0.091	7.9	59.782	238.61	6.6	0.0012
11	0.203	0.183	17.1	8.693	28.007	59.0	0.00235
12	0.177	0.083	17.7	76.989	316.51	4.8	0.0011
13	0.177	0.083	17.7	76.989	315.81	4.8	0.0011
16	0.177	0.083	17.7	76.989	316.51	4.8	0.0011
17	0.177	0.083	17.7	76.989	315.81	4.8	0.0011
20	0.188	0.186	14.8	8.339	28.360	58.7	0.0022
22	0.174	0.142	23.5	17.165	61.434	25.9	0.00182
24	0.091	0.091	7.9	59.782	238.61	6.6	0.0012
25	0.203	0.183	17.1	8.693	28.007	59.0	0.00235
26	0.20	0.15	8.0	14.755	51.844	28.3	0.00193
28	0.30	0.30	1.0	2.796	5.504	268.0	0.0039
31	0.435	0.435	5.9	1.313	11.486	431.0	0.00007
32	0.26	0.26	3.2	3.792	9.007	162.0	0.0033
34	0.26	0.26	3.2	3.792	9.007	162.0	0.0033
36	0.16	0.16	5.0	12.378	42.521	34.0	0.00205
40	0.20	0.20	5.0	6.955	21.144	92.6	0.00255
42	0.255	0.186	12.6	8.339	26.560	62.5	0.0024
43	0.247	0.141	32.1	17.506	62.694	30.0	0.00182
44	0.13	0.13	34.3	21.969	80.330	22.1	0.0017
47	0.20	0.20	5.0	6.955	21.144	92.6	0.00255
49	0.255	0.186	12.6	8.339	26.560	62.5	0.0024
50	0.247	0.141	32.1	17.506	62.694	30.0	0.00182
51	0.13	0.13	34.3	21.969	80.330	22.1	0.0017

Furthermore, we computed the L_2 norm of the absolute differences between the numerical solution obtained with the LW scheme and the TS scheme, using the SCO algorithm. The norm results were in the order of 10^{-6} cm^2 for the cross-sectional area and of 10^{-5} ml/s for the flow rate, showing that both numerical schemes lead to the practically same results.

7.2.2.2 Comparison of the Memory Copy Strategies for the PHCG Algorithm

We evaluated the performance of the three memory copy strategies for the PHCG algorithm. Table 7.4 displays the execution times of the GPU operations, corresponding to the computation of one heart cycle with an elastic wall, the LW scheme and WK outlet boundary conditions (this is the computational setup considered in (Itu et al. 2012a, b)). For the PHCGCS algorithm, the kernel execution occupies

Table 7.4 Execution times (s) of the GPU operations obtained for the computation of one heart cycle with the three variants of the PHCG algorithm. The results correspond to a computation with elastic walls, the LW scheme and the WK outlet boundary condition. *Copy H→D* refers to a copy operation between the host (CPU) and the device (GPU), while *Copy D→H* refers to a copy operation in the opposite direction

Operation	PHCGCS	PHCGCA	PHCGCC
Copy H→D	23.7	3.76	0.85
Kernel	1.86	1.86	2.02
Copy D→H	43.1	5.29	0.89

only 2.7% of the total execution time on the GPU, making the application heavily PCI Express Bus limited. Although the amount of data to be transferred is higher, the PHCGCA algorithm represents an improvement, since the number of copy operations is reduced drastically. The best results are obtained with the PHCGCC algorithm, since the amount of data to be transferred is small as in the first case and the number of copy operations is reduced as in the second case. The only drawback is that some of the threads of the kernel need to populate the additional arrays displayed in Fig. 7.4c. This leads to an increase of 8.6% for the kernel execution time, but the increase is easily compensated by the time gained for the memory copies.

7.2.2.3 Comparison of the Performance Obtained with the SCO, MCO, PHCG and PGO Algorithms

Tables 7.5 and 7.6 summarize execution times measured for the six different computational setups displayed in Table 7.1. The execution times correspond to ten heart cycles and the highest speed-up values are displayed in bold. The grid space has been set to 0.1 cm and the time step to $5.55 \cdot 10^{-5}$s. The values are based on literature data and on the CFL-restriction respectively.

The PHCGCA algorithm cannot be applied for all computational setups investigated herein (as described in Sect. 2.3.1, the workflow in Fig. 7.4a cannot be applied for a viscoelastic wall law). The speed-up values in Tables 7.5 and 7.6 are computed based on the execution time of both the SCO and MCO algorithms.

The speed-up values vary between **5.26**× and **8.55**× compared to the SCO algorithm and between **1.84**× and **4.02**× compared to the MCO algorithm. As anticipated, the PHCGCC algorithm outperforms the PHCGCA algorithm for all cases for which the PHCGCA was applied. For an elastic wall law, in case a WK boundary condition is used, the PHCGCC algorithm performs best, while in case the ST boundary condition is used, the PGO algorithm leads to the highest speed-up. For a viscoelastic wall law, the PHCGCC algorithm performs best, regardless of the type of outlet boundary condition. Execution times are higher with a ST boundary condition because of the time spent for the computation of the convolution integral in Eq. (7.10).

For an elastic wall law, with the PHCGCC algorithm, the execution times are comparable for the LW and the TS scheme (for both outlet boundary condition types), with slight advantages for the LW scheme. For the SCO and MCO algorithms, the LW scheme is superior to the TS scheme.

Table 7.5 Execution times and speed-ups obtained for the computation of ten heart cycles with the SCO, MCO, and PGO algorithms. The first four cases correspond to an elastic wall with either the Lax-Wendroff (LW) or the Taylor series (TS) scheme and with a Windkessel (WK) or structured tree (ST) boundary condition. The last two cases correspond to a viscoelastic wall law with the TS scheme and with a WK or ST boundary condition

						PGO		
							Speed-up	
Case	Num. sch.	Wall law	Outlet BC	SCO (s)	MCO (s)	Time (s)	SCO	MCO
1	LW	Elastic	WK	273.4	81.13	101.98	2.68×	0.79×
2	LW	Elastic	ST	673.7	205.38	111.49	**6.04×**	**1.84×**
3	TS	Elastic	WK	396.3	119.37	105.10	3.77×	1.14×
4	TS	Elastic	ST	797.2	233.09	116.56	**6.84×**	**2.00×**
5	TS	Viscoel.	WK	774.4	384.43	235.14	3.29×	1.63×
6	TS	Viscoel.	ST	1179.6	501.08	241.45	4.89×	2.07×

Table 7.6 Execution times and speed-ups obtained for the computation of ten heart cycles with the PHCGC algorithms. The first four cases correspond to an elastic wall with either the Lax-Wendroff (LW) or the Taylor series (TS) scheme and with a Windkessel (WK) or structured tree (ST) boundary condition. The last two cases correspond to a viscoelastic wall law with the TS scheme and with a WK or ST boundary condition.

				PHCGCA			PHCGCC		
					Speed-up			Speed-up	
Case	Num. sch.	Wall law	Outlet BC	Time (s)	SCO	MCO	Time (s)	SCO	MCO
1	LW	Elastic	WK	68.21	4.01×	1.19×	42.4	**6.45×**	**1.91×**
2	LW	Elastic	ST	182.34	3.69×	1.13×	149.92	4.49×	1.37×
3	TS	Elastic	WK	74.81	5.30×	1.59×	46.37	**8.55×**	**2.57×**
4	TS	Elastic	ST	187.27	4.26×	1.24×	151.90	5.25×	1.53×
5	TS	Viscoel.	WK	–	–		95.64	**8.10×**	**4.02×**
6	TS	Viscoel.	ST	–	–		224.38	**5.26×**	**2.23×**

A detailed analysis of the results obtained with the best performing algorithms (PHCGCC and PGO) is presented below. Figure 7.9 displays the percentage of the execution time occupied by each operation identified in the workflows in Fig. 7.3, for cases 1, 2, 5 and 6, computed with the PHCGCC algorithm. Regarding the computations with an elastic wall law, as is displayed in Fig. 7.3a, computations on the host and on the device are performed in parallel. The operations on the device require more time than the host operations if a WK boundary condition is used. Although the copy operations were optimized, they occupy almost half of the total time spent on the GPU. Besides, a considerable time is required for other operations, which include control instructions, data exchange operations between the host arrays and the arrays used for the copy operations, and print operations during the last cycle. If a ST boundary condition is used, the computation of the convolution integral in Eq. (7.10), performed on the host, occupies most of the execution time.

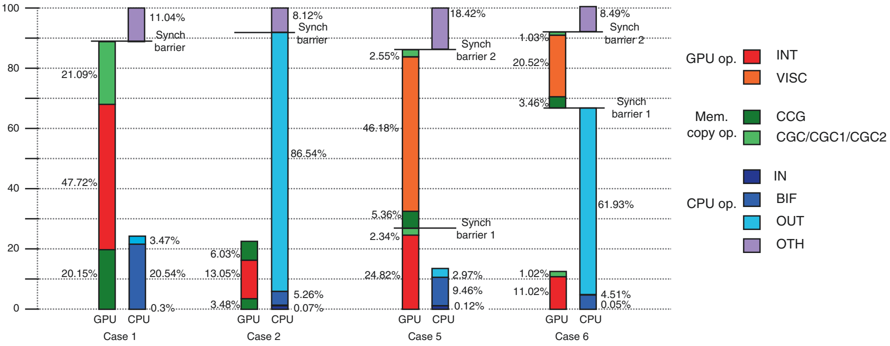


Fig. 7.9 Detailed percentage values of the execution time occupied by the operations identified in the workflow in Fig. 7.3 for the PHCGCC algorithm, for an elastic wall law (cases 1 and 2) and a viscoelastic wall law (cases 5 and 6). Acronyms are detailed in Fig. 7.3 (Itu et al. 2013a, b)

This is the primary reason behind the low speedup achieved with the PHCGCC algorithm and the ST boundary condition. In addition, it also explains the similar speed-up values obtained with the PHCGCC and PHCGCA algorithms.

Regarding the computations with a viscoelastic wall law, as is displayed in Fig. 7.3b, computations on the host and on the device are performed in parallel at the beginning of each iteration, but since the computation of the viscoelastic flow rate requires the values of the elastic flow rate and of the cross-sectional area at the junction points (from the current time step), during the second part of each iteration, only the device performs computations. As for the elastic wall law, in case a ST boundary condition is used, the computation of the convolution integral in Eq. (7.10), performed on the host, occupies most of the execution time.

Figure 7.10 displays the percentage of the execution time occupied by each operation identified in the workflow in Fig. 7.7, for cases 1 and 6, computed with the PGO algorithm. In the first case a single kernel is used, while for case 6 also the convolution integral and the viscoelastic flow rate correction are computed. The computation of the interior and junction points require more execution time for case 6 than for case 1, since, on the one side the TS scheme is used instead of the LW scheme, and on the other side additional operations are performed because of the viscoelastic wall law. Compared to case 6 in Fig. 7.9 (ST boundary condition), the execution time dedicated to the outflow points is reduced significantly since the operations are performed on the device, but because the computation of all grid points requires considerably more time, the total execution time for case 6 is higher with the PGO algorithm than with the PHCGCC algorithm.

Figure 7.11 displays a comparison of the number of heart cycles which can be computed per hour with different algorithms: the SCO and MCO algorithms, the previously introduced PHCGCA algorithm (applied only for non-periodic boundary conditions) and the best performing parallel algorithm for each computational setup as determined in the current study. The four different computational setups have been obtained by combining the different wall laws and outlet boundary conditions and by

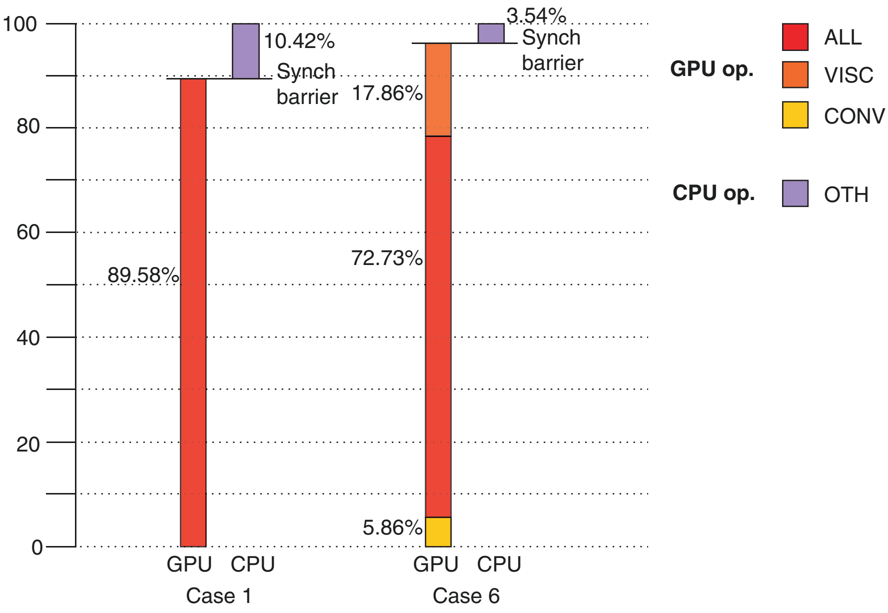


Fig. 7.10 Detailed percentage values of the execution time occupied by the operations identified in the workflow in Fig. 7.7 for the PGO algorithm, for an elastic wall law (case 1) and a viscoelastic wall law (case 6). Acronyms are detailed in Fig. 7.7 (Itu et al. 2013a, b)

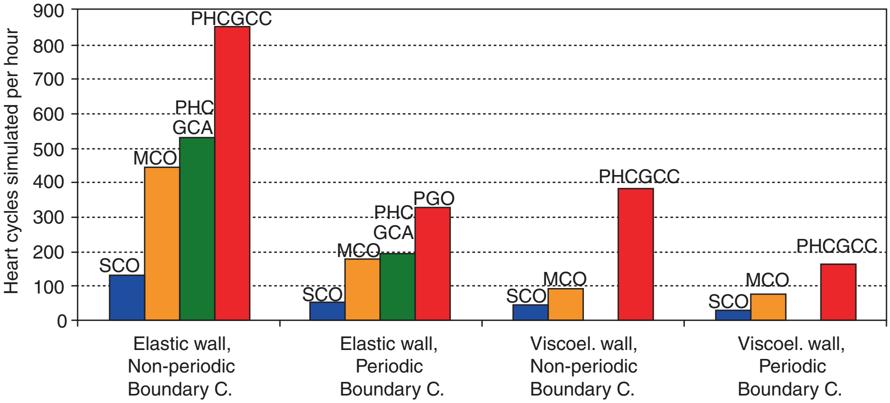


Fig. 7.11 Heart cycles computed per hour for the SCO algorithm, the MCO algorithm and the best performing parallel algorithm for each computational setup (Itu et al. 2013a, b)

choosing the best performing numerical scheme (according to the results in Tables 7.5 and 7.6). The results show that the best performing GPU based algorithms considerably increase the number of heart cycles which can be computed per hour.

To analyze the effect of the simulation parameters on the speed-up factor, we display in Fig. 7.12 the speed-up values obtained for different grid space values: Δx = 0.25 cm (5486 degrees of freedom (dofs), 8000 time steps per cycle),

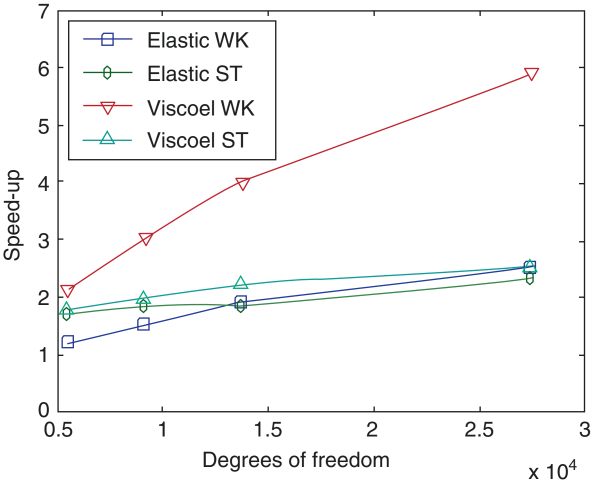


Fig. 7.12 Speed-up values obtained for different grid space configurations for the best performing parallel algorithm compared to the MCO algorithm (Itu et al. 2013a, b)

Δx = 0.15 cm (9144 dofs, 12,500 time steps per cycle), Δx = 0.1 cm (13,716 dofs, 18,000 time steps per cycle), and Δx = 0.05 cm (27,432 dofs, 37,000 time steps per cycle). The displayed values represent the speed-up obtained by the best performing GPU based algorithms compared to the MCO algorithm. The time-step values are chosen to satisfy the CFL condition for each case, and both types of wall laws and outlet boundary conditions are considered. In each case, the numerical scheme and the parallel algorithm applied for the computation correspond to the best speed up value obtained for a grid space of 0.1 cm. Figure 7.12 displays an approximately linear increase of the speed-up value, indicating that the computational power of the GPU is not fully exploited for any of the computational configurations with a grid space higher than 0.05 cm. The increase is moderate for three of the four computational setups and more pronounced in case a viscoelastic wall law is used together with a WK boundary condition. This aspect is given by the fact that the implementation of the viscoelastic wall law is more efficient for the PHCGCC algorithm compared to the MCO algorithm. On the other hand, when a viscoelastic wall law is used together with the ST boundary condition, most of the time is spent for computing the outlet grid points and the difference in execution time for the viscoelastic component becomes less important.

Figure 7.13 displays the time-varying pressure, flow rate and cross-sectional area at the five locations marked with a blue circle in Fig. 7.8.

Each figure contains four plots, which have been obtained with either an elastic or viscoelastic wall and with a WK or ST boundary condition. Since the total resistance introduced by either of the two types of boundary conditions is similar, the average quantities are approximately equal at all locations inside the arterial tree. Referring first to the computations with elastic walls, the pressure values obtained with the ST boundary condition decrease at a later time inside one heart cycle, indicating that the reflected wave arrives later (an aspect which is more pronounced for the proximal parts of the arterial tree). This can be explained as follows: the ST boundary condition simulates the propagation of the waves down to the arteriolar

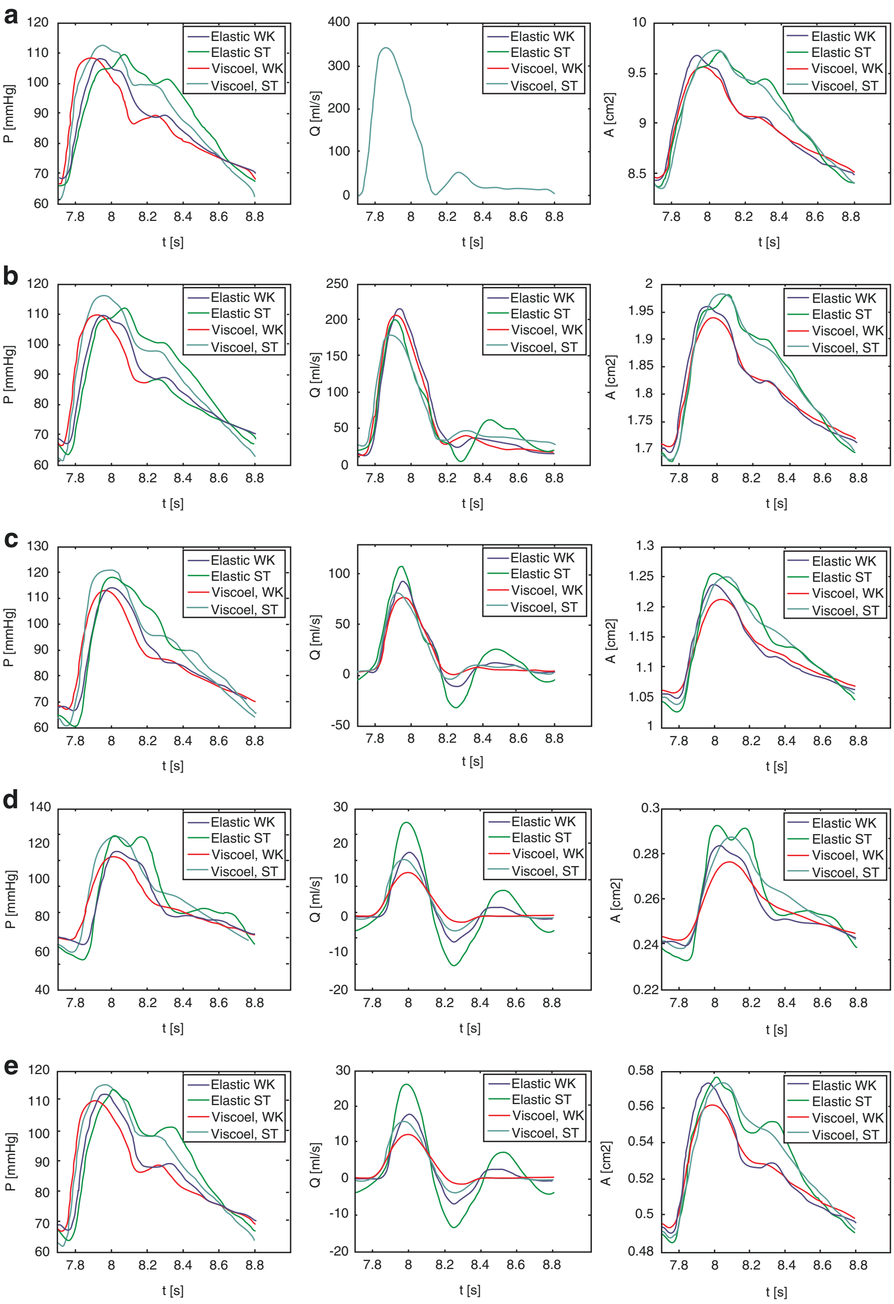


Fig. 7.13 Time-varying pressure, flow rate and cross-sectional area at (**a**) aortic root, (**b**) descending aorta, (**c**) abdominal aorta, (**d**) femoral artery, and (**e**) subclavian artery (corresponding to locations A–E respectively in Fig. 7.6). Four plots are displayed in each figure, which have been obtained with either an elastic or viscoelastic wall and with a WK or ST boundary condition (Itu et al. 2013a, b)

level where the reflections occur primarily, whereas the WK boundary condition, as a lumped model, is not able to capture the wave propagation phenomena in the distal part of the tree and introduces the reflections at the outlet points of the proximal arteries. As a result of the later arriving pressure waves, also the maximum pressure value is reached at a later moment in time. These aspects also lead to higher oscillations inside the flow rate waveforms which are displayed in the second column of Fig. 7.13. Finally, for the cross-sectional area, generally, the variation inside one heart cycle is higher with a ST boundary condition. For the elastic wall, the pressure and the cross-sectional area waveforms are in phase and a more pronounced variation of the area values is reflected by a higher pressure pulse. The higher pressure pulse obtained for the structure tree boundary condition indicates a lower total compliance than the one enforced through the WK boundary condition. We emphasize the fact the compliance of the proximal part of the tree is identical in both cases and the difference in total compliance is given only by the outlet boundary conditions.

When a viscoelastic wall is used, the main difference is that the high-frequency oscillations in the waveforms are reduced. This can be observed in both the pressure and the flow rate waveforms and the phenomenon is more pronounced at the distal locations. These observations are consistent with results reported in literature (Reymond et al. 2011). The introduction of the viscoelastic wall does not change the overall behavior of the WK and ST boundary conditions, the observations mentioned above being still valid, as would be expected. Another important consequence of the introduction of the viscoelastic wall is the fact that pressure and area are no longer in phase, the peak cross-sectional area value being generally obtained at a later moment in time inside one heart cycle.

Furthermore, Fig. 7.14 displays the pressure-area relationships at three different locations. A hysteresis loop can be observed when a viscoelastic wall laws is used, as opposed to the linear variation for an elastic wall law. The area of the hysteresis loop is proportional to the energy dissipation given by the viscoelastic properties of the wall.

7.2.3 Discussion and Conclusions

To test the speed-up potential of novel hybrid CPU-GPU and GPU only based implementations of the one-dimensional blood flow model, we have used a full body arterial model and have applied two physiologically motivated outlet boundary conditions, 3-element windkessel circuits as non-periodic boundary condition and structured trees as periodic boundary condition, and two different types of constitutive wall laws. The speed-up values over a multi-threaded CPU based implementation range from 1.84× to 4.02×, and over a single-threaded CPU based implementation range from 5.26× to 8.55×, thus significantly improving on previously reported parallel implementations and confirming the excellent speed-up potential of the GPU-based implementation.

The results showed that, for an elastic wall, if a non-periodic boundary condition is used, the PHCGCC algorithm performs best, while for a periodic boundary condition, the PGO algorithm performs best. This is motivated by various aspects. First of all, the PHCGCC algorithm decreases the execution time not only through data

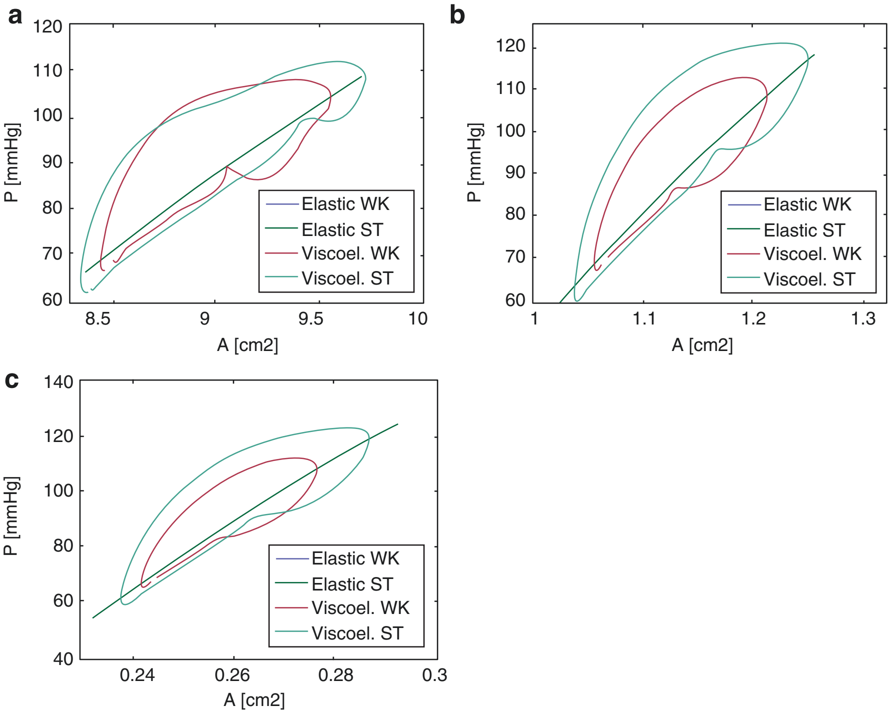


Fig. 7.14 Pressure-area relationships at (**a**) aortic root, (**b**) abdominal aorta, and (**c**) femoral artery (corresponding to locations A, C and respectively D in Fig. 7.8). Four plots are displayed in each figure, which have been obtained with either an elastic or viscoelastic wall and with a WK or ST boundary condition (Itu et al. 2013a, b)

parallelization on the GPU, but also through task level parallelization between CPU and GPU. For periodic boundary conditions, the execution time required specifically for the outlet points is more than one order of magnitude higher than for non-periodic boundary conditions (Fig. 7.9). Since for the PHCGCC algorithm the outlet points are computed on the CPU, the speed-up obtained for a periodic boundary condition is limited. In this case, the PGO algorithm performs better, although it does not employ task level parallelization, because the multiply-sum scan operation required for the convolution integral is more efficient on the GPU (especially when the number of time steps per cycle is high, as is here the case). In case a viscoelastic wall law is used, the PHCGCC algorithm performs best for both types of boundary conditions. Although the PGO reduces the execution time dedicated to the outlet points, it is slower than the PHCGCC algorithm because the computation of all grid points increases significantly due to the viscoelastic wall law.

As has been shown in Fig. 7.12, the speed-up potential of the GPU-based algorithms is even higher when the number of grid points increases. This is given by the fact that, for the standard configuration with Δx = 0.1 cm, only 6858 threads are generated at the kernel execution, whereas the GPU is able to run grids with tens of thousands of threads. On the other side, even when the number of threads is decreased 2.5 times (Δx = 0.25 cm), still significant speed-up values are obtained. This is obtained as a result of the activities performed for the kernel optimization, which optimized the kernel towards high computational intensity and few global memory operations. Whereas the grid space is a crucial factor for the final speed-up value, the number of time steps does not influence the results, since different iterations are executed sequentially both on the CPU and on the GPU.

Furthermore, when the interior points of the arteries are solved on the GPU, the maximum number of grid points for a single segment is limited by the resources of the GPU (number of registers, amount of shared memory and maximum number of threads per block). Consequently, if an arterial vessel contains too many grid points, it is split into separate segments and junction points are introduced which are treated similarly to bifurcation points (conservation of flow rate and pressure). For the arterial tree used herein, this approach has been required only when a grid space with Δx = 0.05 cm was used and only for the vessels longer than 28.2 cm (the main limiting factor has been the shared memory). Furthermore, we emphasize that a generic arterial model was used for demonstrating the algorithm. For patient-specific models, the geometry will have slight variations in terms of lengths and radiuses. This, however, does not affect the proposed algorithm and the memory requirements.

Next, we focus on a theoretical analysis of the performance gain obtained through the GPU, based on the standard computational setup with Δx = 0.1 cm. CUDA introduced the single instruction multiple threads (SIMT) architecture whereas an operation can be executed in parallel on p processors. By analogy with a SIMD system (Jordan and Alaghband 2003), p represents in this case the number of cores of the GPU. The total number of interior grid points is:

$$N = \sum_{i=1}^{m} \left(N_i - 2\right), \tag{7.21}$$

where m is the total number of arteries and N_i is the number of grid points of each artery.

Referring first to the computation of the interior grid points, if n_I different operations are required to compute the new cross-sectional area and flow-rate values at a single interior grid point, the theoretical computing time is:

$$T_p = \frac{n_I \cdot N}{p} \tag{7.22}$$

for the PHCG algorithm and:

$$T_1 = n_I \cdot N \tag{7.23}$$

for the SCO algorithm. Hence, the theoretical parallel speed-up and the efficiency are respectively:

$$S_p = \frac{T_1}{T_p} = p, \qquad E_p = \frac{S_p}{p} = 100\%. \tag{7.24}$$

The performance of GPU kernels is limited by either the global memory bandwidth or by the computational intensity. The kernel which computes the interior grid points has an overall global memory throughput of only 11.3 GB/s, as opposed to the peak theoretical of 192.2 GB/s of the GTX680 card. The instruction throughput on the other side is of 69.91 GFLOPS (Giga Floating Point Operations per Second), with $5.71 \bullet 10^6$ warps executed per second. The technical specifications of the GTX680 only contain the theoretical single precision GFLOP value (3090 GFLOPS), whereas the computations for the one-dimensional model are performed in double precision. As is specified though in, the instruction throughput for double precision computations is significantly lower than the instruction throughput for single precision computations for the GTX680 architecture.

Since for the GTX680 $p = 1536$, the ideal execution time of the interior points should be 1/1536 of the sequential execution time. Taking case 1 in Tables 7.5 and 7.6 as a representative case, the execution time of the interior points for the PHCGCC algorithm is of 20.23 s, while for the SCO algorithm it is of 240.45 s. Thus, the speed-up is of only 11.88× and not 1536×. The great difference between the theoretical speedup and the empirical speedup is given by (1) not all thread blocks can be executed simultaneously because of the shared memory limitation—occupancy is limited (only one thread block can be executed at a time on each of the 8 streaming multiprocessors of the GTX680, whereas there are 51 thread blocks in total), (2) the necessity to perform synchronization at thread level, (3) the parallelism is limited for a grid space of 0.1 cm, and (4) the number of double precision floating point units is significantly smaller than the number of cores.

Secondly, we refer to the computation of the viscoelastic component of the flow rate. The size of the tridiagonal system varies between the different vessel segments, leading to an average length of 14 cm for the arteries of the tree displayed in Fig. 7.8. Hence the average size of the tridiagonal system is 141. This value is rounded up to a power of 2, leading to a tridiagonal system of size $n_T = 256$. For the Thomas algorithm $8nT$ operations are executed, while for the CR-PCR parallel algorithm $17(n_T - m_T) + 12m_T\log_2 m_T$ operations are executed, where m_T is the size of the system solved with the PCR algorithm ($m_T = 128$ on average). Hence:

$$S_p = p \cdot \frac{8n_T}{17\left(n_T - m_T\right) + 12m_T \log_2 m_T} = 243.33, \qquad E_p = \frac{S_p}{p} = 15.84\%. \quad (7.25)$$

Similarly to the first kernel, the performance of the kernel which computes the viscoelastic component of the flow rate is also limited by the computational intensity. The overall global memory throughput is of only 1.3 GB/s, whereas the instruction throughput is of 77.94 GFLOPS, with $3.12 \bullet 10^6$ warps executed per second.

Taking case 5 in Tables 7.5 and 7.6 as a representative case, the execution time of the viscoelastic flow rate for the PHCGCC/PGO algorithms is of 44.17 s, while for the SCO algorithm it is of 376.49 s. Thus the speed-up is of only 8.52× and not 243.33×. Additionally to the reasons enumerated above (since shared memory usage for this kernel is similar to the first kernel, occupancy is limited: only one thread block can be executed at a time on a streaming multiprocessor), also the fact that the $17(n_T - m_T) + 12m_T\log_2 mT$ operations cannot be executed in parallel plays an important role (the CR-PCR algorithm is executed on average in $2\log_2 n_T - \log_2 m_T - 1 = 8$ sequential steps).

Thirdly, we refer to the computation of the convolution integral (for the PGO algorithm). Each multiplication/addition pair of this operation is combined into a single floating point operation, leading to a total of n_C operations, where n_C is the number of time steps for one heart cycle. Thus,

$$S_p = p = 336, \qquad E_p = \frac{S_p}{p} = 100\%. \tag{7.26}$$

Unlike the first two kernels, the performance of the kernel which computes the convolution integral is limited by the global memory bandwidth. The overall global memory throughput is of 184 GB/s, which is close to the peak theoretical value, whereas the instruction throughput is of only 21.8 GFLOPS, with $7.98 \bullet 10^6$ warps executed per second.

Taking case 2 in Tables 7.5 and 7.6 as a representative case, the execution time of the convolution integral for the PGO algorithm is of 14.14 s, while for the SCO/PHCGCC algorithms it is of 407.13 s. Thus the speed-up is of only 28.79× and not 1536×. Additionally to reasons (2) and (4) enumerated during the analysis performed for the interior grid points, also the fact that the n_C operations cannot be executed in parallel, and the fact that the kernel is limited by the global memory throughput, play an important role (the optimized scan algorithm uses $\log_2 n_C = 11$ sequential steps). In this case the number of simultaneously active threads on a streaming multiprocessor is close to maximum, leading to an occupancy of 81% (the shared memory does not limit the occupancy since each block requires only 8 kB of shared memory).

As expected, the overall speed-up of the application is lower than the speed-up of the individual components because of the initialization activities, control operations, task-level synchronization, writing of results to files, data exchange operations between the host arrays and the arrays used for the copy operations (PHCGCC algorithm), limited parallelism of junction points operations (PGO algorithm), etc.

Our past work analyzed the speed-up potential only for one computational setup (elastic wall, windkessel outlet boundary condition as non-periodic BC and the Lax-Wendroff numerical scheme) of the six scenarios considered in the present study and only compared with the SCO algorithm (Itu et al. 2012a, b). Additionally, a different GPU (GTX460) was used, and only the PHCG approach was considered, for which, from the three different variants introduced herein, only the PHCGCS and PHCGCA algorithms were investigated. As can be seen in the results section (Tables 7.5 and 7.6 and Fig. 7.11), the third variant (the PHCGCC algorithm), outperforms the previous two algorithms. Additionally, a new PHCG workflow, based on the operator splitting scheme, was developed to be able to employ any of the PHCG variants in case a viscoelastic wall law is applied (Fig. 7.3b).

The consideration of different computational setups has been a crucial aspect, since it demonstrated the limited speed-up obtained for the PHCG variants with periodic boundary condition and an elastic wall law. The PGO algorithm proposed in the current study leads to considerably improved execution times for these cases (case 2, and 4 in Tables 7.5 and 7.6). When comparing the results against (Itu et al. 2012a, b), one has to take into consideration that previously a grid space of 0.05 cm was used, for which a speed-up of 4.7× was obtained for the PHCGCA algorithm with the GTX460 (this speed-up changes to 4.01× for a grid space of 0.1 cm with the GTX680).

To our knowledge, the only other previous research focused on the acceleration of the one-dimensional blood flow model has been reported in (Kumar et al. 2003). An Origin 2000 SGI machine with eight processing elements was used and the Message Parsing Interface (MPI) libraries have been applied for the communication between the processing nodes. An elastic wall law was used together with a Taylor-Galerkin numerical scheme and the results were mainly reported for up to four processing nodes. An arterial model composed of 55 arteries, very similar to the one adopted herein, has been used and the speed-up factor did not exceed around 3.5× even if the grid space was decreased to 0.05 cm (~3.25× for $\Delta x = 0.1$ cm). We have seen that in our case the speed-up is of 6.45× for a grid space of 0.1 cm. In (Kumar et al. 2003) non-reflecting and resistance-based boundary condition, which are both non-periodic boundary conditions, have been used. Herein we obtained for the MCO algorithm, compared to the SCO algorithm, a similar speed-up as in (Kumar et al. 2003) (3.37× for the computational setup with an elastic wall law and windkessel boundary condition). These results show that a GPU is better suited for the acceleration of the one-dimensional blood flow model than a multi-threaded CPU based configuration.

Overall, we think that the advantages of a GPU-based implementation of the one-dimensional blood flow model outweigh the costs. Most importantly, the acceleration of the execution time is crucial when the blood flow model is applied in a clinical setting (Itu et al. 2012a, b, 2013a, b). This is given not only by the fact that results are required in a timely manner, but also by the necessity of applying tuning procedures which increase the computational intensity. Such accelerated approaches (when validated) are ideal for interventional settings, where near real-time information is needed to make the clinical decision. Under such settings, the measurement data is often acquired when the patient is undergoing an intervention. As a result, the tuning and simulation process for the hemodynamic simulation should be fast enough to generate pressure and flow information that can be used during the procedure.

Secondly, as pointed out in the introduction, research activities can also benefit from the acceleration if results depend on running hundreds or even thousands of computations with different configurations.

On the other side, costs are limited, both financially (the results reported herein have been obtained on a regular desktop computer equipped with an NVIDIA graphics card) and from the development time point of view. Especially, when the PHCGCC algorithm is employed, only the computation of the interior points needs to be ported to the GPU. Furthermore, even when the PGO algorithm is required, i.e. when periodic outlet boundary conditions are imposed, the additional development time does not exceed a couple of weeks.

7.3 GPU Accelerated Voxelizer

Solid voxelization represents the process of transforming a polygonal mesh into a voxel representation by associating each polygon of a mesh with the cells in the voxel grid. Voxel representation of solids are currently used in many applications such as physics simulations, collision detection, volume rendering, and many others. The main advantage of the voxel representation of a solid is that each voxel in the grid can be accessed directly by knowing its position in space or its position relative to another voxel, without performing a search operation, whereas in a mesh representation information is described sparsely as a set of polygons, by providing the position of each point explicitly.

Although there are many studies on this topic, solid voxelization remains a difficult problem, mostly because of computational complexity and issues related to robustness

To perform a Fluid-Structure interaction (FSI) using the Lattice-Boltzmann method (LBM) the moving geometry has to be embedded in a Cartesian grid of uniformly distributed points using a signed distance field $\phi(\mathbf{x})$. However, the geometry is typically given as a sequence of non-uniform polygonal meshes. A surface voxelization operation is required to compute the distance field. The main challenge of voxelization consists in associating each vertex $\mathbf{v}i$ of a polygonal mesh to each node $\mathbf{x}i$ of the Cartesian grid. Typically, the size of the grid is between 500,000 and 50,000,000 nodes while the size of the mesh is between 50,000 and 300,000. This makes the voxelization a computationally expensive operation.

For the FSI computations, since the surface is moving, the voxelization operation is required at each solver iteration to update the position of the surface. With the classical method (CPU based implementation), the surface voxelization operation is the performance bottleneck as it occupies around 50% of the total computation time (Nita et al. 2015). Therefore, it is crucial that an efficient implementation is developed.

7.3.1 The Classical Method

The Cartesian grid is defined from a three-dimensional image by its dimensions (N_x, N_y, N_z), an origin $\mathbf{o}$ and a grid spacing δx (the grid nodes are uniformly distributed hence $\delta_x = \delta_y = \delta_z$). The grid size is chosen to satisfy the flow solver stability constraints. The origin $\mathbf{o}$ and the grid spacing δx is used to transform from physical coordinates to grid coordinates (and vice versa) i.e. to find the voxel (i, j, k) that corresponds to a point $\mathbf{p} = (p_x, p_y, p_z)$. The transformation is defined as follows:

$$i = \frac{p_x}{\delta_x} - o_x, j = \frac{p_y}{\delta_x} - o_y, j = \frac{p_z}{\delta_x} - o_z, \tag{7.27}$$

where $\lfloor x \rfloor$ denotes the floor function. And the inverse transformation:

$$p_x = i\delta x + o_x, p_y = i\delta x + o_y, p_z = k\delta x + o_z, \tag{7.28}$$

A mesh is defined as a set of triangles $T = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, for each triange we compute an axis-aligned bounding-box (AABB) as $AABB = (\mathbf{v}min, \mathbf{v}max)$ so that the triangle will completely fit inside it, furthermore the AABB is enlarged in all directions using a small value $(2-3\delta x)$ so that the triangle vertices will never be located exactly on the AABB wall.

The classic method for surface voxelization consists in simply looping over each grid node $\mathbf{x}_i$ in each AABB and computing the signed distance $\phi(\mathbf{x}_i)$. To find all the grid nodes inside the AABB, one needs to transform $(\mathbf{v}min, \mathbf{v}max)$ to grid coordinates to get $(imin, jmin, kmin)$ and $(imax, jmax, kmax)$ and then loop over all i, j and k values located inside the bounds.

The signed distance function is defined as follows:

$$\phi(\mathbf{x}) = d(\mathbf{x},\mathbf{x}_\perp)\,\mathrm{sgn}\left[(\mathbf{x}-\mathbf{x}_\perp)^\circ \cdot \mathbf{n}\right], \tag{7.29}$$

where $\mathbf{n}$ is the triangle normal and $\mathbf{x}_\perp$ is the closest point to $\mathbf{x}$ on the triangle. The second factor in the above expression represents the sign, i.e. it will be negative or positive depending on which side of the triangle, the point $\mathbf{x}$ is located. For adjacent triangles the AABBs will intersect and will result in multiple ϕ values for the same grid point $\mathbf{x}$, one value for each AABB that point $\mathbf{x}$ is included in (Fig. 7.15). In this case the absolute minimum value of ϕ will be chosen.

For the GPU implementation, the loop that processes the mesh triangles is parallelized so that one GPU thread will process one triangle. However there are several downsides that causes very poor GPU utilization in this case. The main problem arises at the adjacent triangles where the AABBs intersect. In the intersection

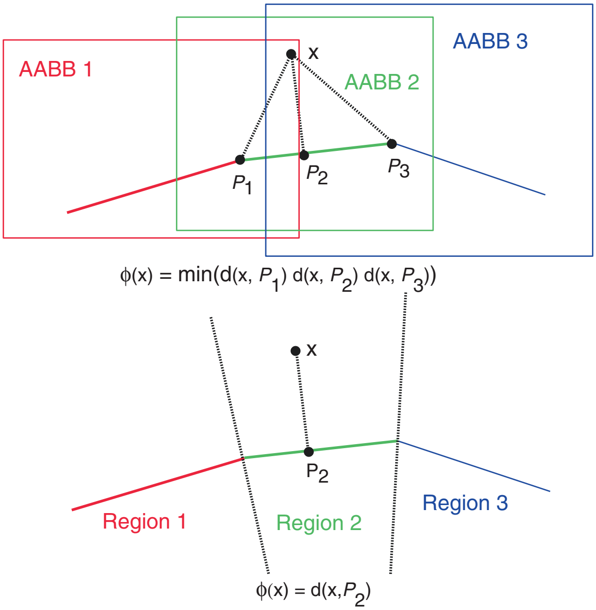


Fig. 7.15 Two-dimensional analogy of the surface voxelization algorithm. The classic approach (up): ϕ is computed for all the nodes inside an AABB. And the separating planes technique (down): nodes that correspond exclusively to a facet are identified using separating planes

regions, there will be multiple threads that need to update the ϕ value at the same grid node $\mathbf{x}$. In this case a synchronization operation is required to ensure that only one thread will update one location at the same time. The synchronization operation drastically reduces parallelism and GPU performance.

The other limitation is given by the fact that each GPU thread will process a different number of grid nodes because of the different AABB sizes. More specifically, the number of the grid nodes in an AABB is influenced by the size and orientation of the corresponding triangle. To achieve maximum performance with a GPU based implementation, all the threads should execute the same operations.

7.3.2 The Separating Plane Technique

The classical method can be improved by redefining the way grid nodes are associated with mesh triangles. Instead of computing the ϕ value for all the nodes in an AABB it is possible to identify a priori the nodes for which each mesh triangle will give the minimum ϕ. Hence, there will no longer be threads that will need to update ϕ at the same location $\mathbf{x}$. This method was initially presented in Janßen et al. Janßen et al. (2015)).

For each triangle we define a region so that each point $\mathbf{x}$ in that region has the closest point $\mathbf{x}_\perp$ located on that triangle. To define such a region for a triangle, three planes are required, one for each edge. More specifically, if a node is located on the negative side of all three planes then that node is considered to belong exclusively to that triangle.

We check if a point $\mathbf{x}$ is located in a triangle region in the following way (Fig. 7.16):

1. For each vertex $\mathbf{v}i$ on the mesh, the vertex normal is computed as an angle weighted average of the normals of adjacent triangles:

$$\mathbf{n}_\Sigma = \frac{\sum \alpha_i \mathbf{n}_i}{\left|\sum \alpha_i \mathbf{n}_i\right|} \tag{7.30}$$

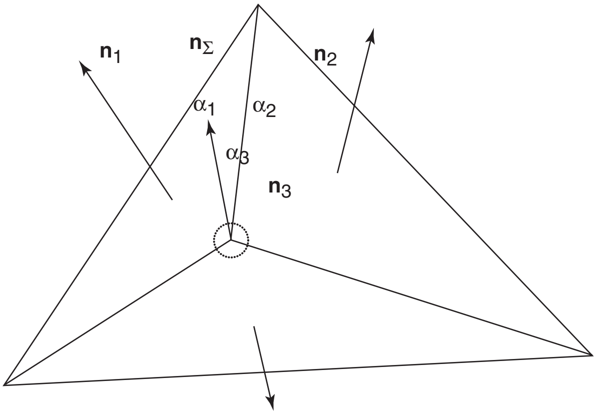


Fig. 7.16 Defining a vertex normal as an angle weighted average of the normals from adjacent triangles

2. For each edge $(\mathbf{v}_i, \mathbf{v}_j)$, with the associated vertex normals $(\mathbf{n}_i, \mathbf{n}_j)$ a separating plane is defined:

$$\mathbf{n}_s \cdot (\mathbf{x} - \mathbf{v}_i) = 0 \tag{7.31}$$

   Where $\mathbf{n}_s$ is the separating plane normal and is computed as an edge bi-normal:

$$\mathbf{n}_s = \frac{1}{2}\left[(\mathbf{v}_j - \mathbf{v}_i) \times (\mathbf{n}_i + \mathbf{n}_j)\right] \tag{7.32}$$

3. A point $\mathbf{x}$ is considered to be located inside a region of a triangle $T = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ if it is located on the negative side of all three separating planes:

$$\begin{cases} \mathbf{n}_{s_1} \cdot (\mathbf{x} - \mathbf{v}_1) \le 0 \\ \mathbf{n}_{s_2} \cdot (\mathbf{x} - \mathbf{v}_2) \le 0 \\ \mathbf{n}_{s_3} \cdot (\mathbf{x} - \mathbf{v}_3) \le 0 \end{cases} \tag{7.33}$$

For any two adjacent, non-intersecting triangles, the regions defined by Eq. (7.33) will not intersect. If each GPU thread processes the nodes in separated regions then there will never be any concurrency hence the synchronization is no longer required. This drastically improves the GPU parallelism and performance.

7.3.3 Results

To test our implementation we considered a known CFD benchmark case consisting of a large brain aneurysm (Steinman et al. 2013). Figure 7.17 displays the mesh along with the voxelized surface.

The mesh contains 318,000 triangular elements and the size of the grid in which the surface is embedded is 171 × 180 × 142. We performed the computations for this case using the CPU and GPU implementations for both the classic and the separating planes method. The hardware we used consists of an Intel i7 (8-cores) CPU and a GTX Titan Black GPU.

The execution times were:

- for the classic method on the CPU the execution time was 23.5 s and on the GPU it was 234 ms which gives a speedup of around 100 times. The GPU execution time does not contain the CPU-GPU memory copy as in an FSI simulation the memory copy should only be done once in the pre-processing stage.
- for the separating planes method, the GPU execution time was 21.4 ms. Compared to the current implementation that we use for FSI computations, the new GPU-accelerated one is around 1000 times faster. Using this approach, the performance of the FSI computations can be taken to an unprecedented level.

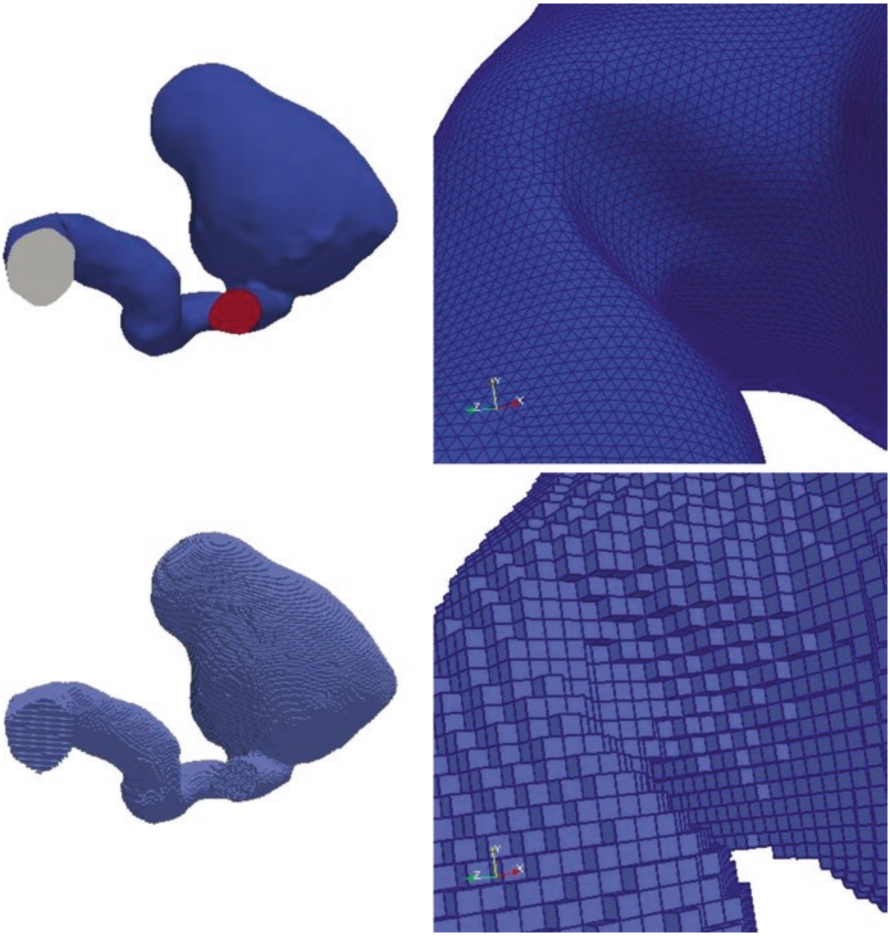


Fig. 7.17 Test case: a large brain aneurysm mesh of 318,000 triangular elements (up) and the voxelized surface (down)

7.4 GPU Accelerated Solution of Large Linear Systems of Equations Using the Preconditioned Conjugate Gradient Method

The focus for this section has been on the development of fast solutions for very large sparse linear systems of equations of the type $A \bullet x = b$, using parallel methods, where A is an N × N symmetric positive definite matrix. Such systems routinely appear when computing numerical solutions to PDEs, such as but not limited to the Finite Element Method. A widely-used iterative approach for solving such linear systems is the Conjugate Gradient (CG) method (Hestenes and Stiefel 1952).

In each iteration, the CG method performs a Sparse-Matrix Vector multiplication, a process that converges in at most N iterations to the exact solution. While current GPU technology excels in fast processing of a large number of parallel computational threads, its global memory size can still create a bottleneck in solving large linear systems. We have developed a methodology for overcoming this limitation using a streaming based algorithm. Parallel algorithms for iterative solutions to the above system have been proposed already, e.g. Ortega (1988), and using the CUDA framework (Verschoor and Jalba 2012). However, such solutions do not take into account RAM memory limitations on the size of the solution.

7.4.1 Single-GPU Solution

We propose a streaming based algorithm in order to overcome the global memory size limitation of the GPU, so as to be able to solve large systems arising in numerical solutions of various biomechanical PDEs. These may include but are not limited to fluid flow, bone or soft tissue deformation, etc. GPU cards have limited RAM memory (currently up to 12 GB), which limits the size of the system of equations (currently to around 12 million equations). To alleviate this limitation we introduce a streaming based solution whose core idea is to store the matrix A on the CPU RAM, and to transfer it slice by slice to the GPU during the matrix-vector multiplication step of the PCG method (Fig. 7.18). Our streaming based strategy can be used either in the context outlined in the next section, or in the more general context of iterative methods that need to handle during each iteration an operation involving data that exceeds the GPU memory.

When applying the PCG method, the majority of the memory required for the solution is occupied by matrix A and by the matrix used for the preconditioning (Saad 2003). To reduce the memory requirements, here we apply a Jacobi (diagonal) pre-conditioner which is stored as vector, while A is stored in a sparse matrix format (e.g. ELLPACK) (Bell and Garland 2008). The other operations of the PCG method are either vector-vector or scalar-vector operations. The seven vectors are stored throughout the entire execution on the GPU due to their limited memory size. Figure 7.18a displays the slicing strategy of A. To ensure coalesced memory accesses by the threads of the same warp, A is stored in column major order. The slicing however is performed on a row basis. To limit the number of copy operations to two for one slice, we still store data in column major order, but only at slice level (all values of one slice are stored in consecutive locations). To reduce the execution time, the memory transfer operation of one slice is overlapped with the processing of another slice. To implement the overlapping behavior, two memory slices are allocated on the GPU (marked 'A' and 'B' in Fig. 7.18a: while data is copied into one slice, the other one is processed.

The matrix-vector multiplication is performed by using two different streams: one for the memory transfer operations—host (CPU) to device (GPU)—and one for the processing of the slices—kernel execution (Fig. 7.18b). The operations of different streams are executed out of order and need to be synchronized. Therefore we use

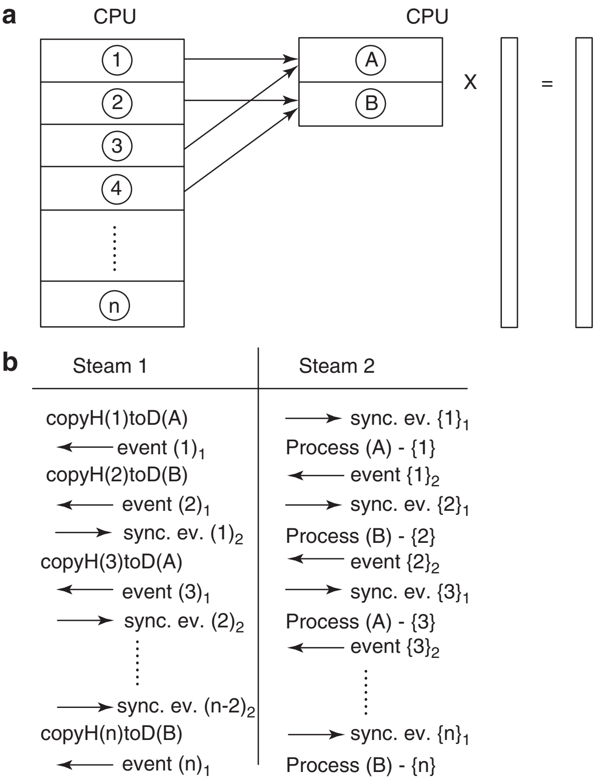


Fig. 7.18 (**a**) Slicing strategy for matrix A and successive memory transfer to the GPU; (**b**) Streaming-based execution of the matrix-vector multiplication

CUDA events: an event $(i)_1$ is recorded after copying slice *i* from host to device, while an event $(i)_2$ is recorded after processing slice *i*. The synchronization at events $(i)_1$ is used to enable the processing of one slice only after the corresponding memory transfer operation is finished. The synchronization at events $(i)_2$ is used to enable the overwriting of one slice only after it was processed. This ensures a correct synchronization between the streams, irrespective of the relative duration of the memory transfer and processing operations.

Next, we present specific results for linear systems of equations as they typically arise in solid mechanics applications. To test the method described abobe, we used four different FE models. For each model, a volume of cube was meshed with 8-node hexahedral elements, with each node having three DOFs (translation in *x*, *y* and *z* dimensions). The four linear systems were first solved with the commercial software ANSYS (Release 14.0.3, ANSYS, Inc) on a six-core processor (Intel (R) Xeon (R) E-5-2670. 2.60 GHz) with 256 GB of RAM. Next, we solved the systems with the streaming based GPU algorithm on an eight-core i7 processor, 3.4 GHz, with 8 GB of RAM and a NVIDIA Kepler GTX680 graphics card, with 2 GB of RAM. Table 7.7 compares the execution times of the CPU and GPU based algorithms. The speed-up varies between **44.2**× and **181.2**× for the largest system of equations.

Table 7.7 Comparison of CPU and GPU based algorithms

		CPU–Ansys	Streaming based GPU algorithm			
Config.	Nr. of equations	Exec. time (s)	Exec. time (s)	Nr. iter.	Number of slices in *A*	Speed-up
Test 1	2.260.713	3441	77.8	525	30	**44.2**
Test 2	4.102.893	21,334	175.9	647	30	**121.3**
Test 3	5.582.601	35,125	268.2	721	18	**131.0**
Test 4	7.057.911	66,046	363.1	785	44	**181.9**

7.4.2 Multi-GPU Solutions

The method described in the previous section was also used to evaluate multi-GPU based implementation of the preconditioned conjugate gradient method.

The strategy we implemented for running the PCG method on multiple GPUs is:

- One node is considered to be the master and performs all operations which do not involve the matrix-vector multiplication (initialization, copy operations, vector-vector operations, scalar-vector operations).
- All nodes, including the master node, perform the matrix-vector multiplication step:
  - Each GPU stores a section of matrix A, which includes multiple slices (data transfer to the various GPUs is performed during initialization). All sections are approximately equal.
  - At each iteration, before starting the matrix-vector multiplication, each GPU receives the vector values used during the multiplication.
  - Each GPU performs in parallel the matrix-vector multiplication for the section of the matrix which was assigned to it.
  - Each GPU sends the resulting vector section back to the master node.

The Message Passing Interface (MPI) is used for transferring data between the nodes. The implementation described in this section is motivated by two goals:

- Reduce the execution time of a single-GPU based implementation
- Handle cases when matrix A does not fit into the RAM memory of the CPU

Table 7.8 displays the results for a two-node configuration for the *Test 2* and *Test 3* configurations used in the previous section. As one can observe, if a 1 Gbit/s node-to-node transfer speed is used the execution time increases compared to the single-node implementation. This is given by the time required to transfer the vectors at each iteration. Conversely, if a 10 Gbit/s node-to-node transfer speed is used, the execution time decreases by a factor of around 1.5, which means that the time spent during data transfer is overcompensated by the time saved through the parallel computation of the matrix-vector product.

Table 7.8 Comparison of single-node and multi-node GPU based implementation of the PCG method.

Configuration	1 Node (s)	2 Nodes (1 Gbit/s) (s)	2 Nodes (10 Gbit/s)
Test 2	176	450	116 s (1.51×)
Test 3	268	704	179 s (1.50×)

Table 7.9 Comparison of single-node single-GPU and multi-GPU based implementation of the PCG method

			Speed-up	
Configuration	Slice distribution	Exec. time (s)	Measured	Theoretical
Titan (×16)	55	164.1 ± 0.87	–	–
Titan(×16), 680(×16)	35/20	120.0 ± 0.64	1.37	1.57
Titan(×16), 680(×8), 680(×8)	35/10/10	137.1 ± 0.69	1.20	1.57

The strategy described above was also used for a hardware configuration containing multiple GPUs in a single node. The major difference is that instead of using MPI to transfer data from one node to another, all data transfers are performed through the PCI Express bus. The specific hardware configuration used for testing is: E6989 Rampage IV Extreme Main Board which has 4 PCI slots (2 at ×16, 2 at ×8), 3 GPUs: 1× GTX Titan Black, 2× GTX680. Due to the different transfer speeds over the PCI bus, the ideal partitioning of slices to different GPUs depends on the PCI bus to which each GPU is physically connected. Hence, we implemented a methodology for automatically determining the number of slices for each GPU so as to obtain an execution time for the matrix-vector multiplication which requires approximately the same time on each GPU (so as to avoid idle times for the processors). Table 7.9 displays the results for three different configurations. One can observe that the multi-GPU implementation leads to smaller execution times. However, the measured speed-up is smaller than the theoretical value due to transfer of data between GPUs.

7.5 GPU Accelerated Random Forest Classification

Machine Learning algorithms have been proven to be useful in a variety of application domains (Zhang 2000). Herein we focus on one of the most common machine learning applications: classification. We study how to effectively implement a random forests (RF) algorithm for data classification on GPUs by evaluating the performance of the algorithm in terms of execution time, compared to the CPU-based version. The random forest consists of multiple decision trees which can be generated and evaluated independently and can classify large amounts of data, described by a large number of attributes. Therefore, the random forest algorithm is very well suited for a massively parallel approach (implemented on GPUs).

Random forest is an ensemble classifier consisting of decision trees that combines two sources of randomness to generate base decision trees: bootstrapping instances for each tree and considering a random subset of features at each node (Breiman 2001). It is a supervised learning method: the training data consists of a set of training examples; each example is a pair consisting of an input object (typically a vector of features) and the corresponding desired output value. A supervised learning algorithm analyzes the training data and produces a model, which is then used to predict the output for new examples.

During the learning phase, the data that has reached a given leaf is used to model the posterior distribution. During the test phase, these posterior distributions enable the prediction for new unseen observations reaching a given leaf.

Because the training phase is done offline, the time required by this phase is not critical. Therefore, we only focus on the acceleration of random forest classification since in most of the cases this phase is done online and the execution time may be critical. The algorithm behind the testing phase is based on the fact that each internal node of a tree contains a simple test that splits the space of data to be classified and each leaf contains an estimate based on training data of the posterior distribution over the classes. The input data transformed into a feature vector is classified by propagating the information through all the trees and performing an elementary test at each node that directs it to one of the child nodes. Each decision node contains a test function that compares a feature response with a threshold to generate a binary decision. Once the sample reaches the leaf in each three in the forest, the posterior probabilities are combined (voting or averaging) to compute the final posterior probability. Traversing a large number of decision trees sequentially is ineffective when they are built independently of each other. Since the trees in the forest are independently built and the only interaction is the final counting of the votes, the voting part (classification) of the RF algorithm can be efficiently parallelized (Grahn et al. 2011).

The first step of the GPU based implementation of the RF classifier is to load all decision tree data structures of a RF into the GPU memory. Prediction is performed in a loop over all pixels of the input image. We determine the feature response for each pixel which will become the input vector for each tree. Each decision tree is then traversed in parallel to retrieve the probability distribution over all classes for the given pixel. The probability distributions from every tree in the forest are averaged and, finally, the result is copied back into the CPU memory.

To accelerate prediction on the GPU, we use multiple threads to process each image and multiple threads to process the RF trees in parallel. After images are loaded, we calculate the integral image in a pre-processing step. Calculating image integrals is expensive with respect to processing time. We accelerate it by calculating the integral for each of the five image channels in parallel with separate threads on the GPU. Prediction time depends on the complexity of the features but scales linearly with the number of trees, depth of the trees and the number of pixels in the input image. To take advantage of the massively parallel computing power of the GPU, instead of pre-computing all values for all possible features, we sampled the feature space at runtime and calculated the feature responses on demand.

Fig. 7.19 Algorithm for the binary decision tree evaluation

```
Algorithm- Random Forest prediction on the GPU
for all trees in the forest do
    while curNode has valid children do
        if children_found then
            float right = evaluate_Harr_feature(boxCenter, curNode);
                if (right < 0)
                    curNode <- leftChildPosition
                else
                    curNode <- rightChildPosition
            else if leaf_node_reached then
            probability + = leaf_Node_Histogram
            total ++
            end if
    end while
end for
return probability / total
```

Our strategy for storing the RF decision trees involves the mapping of the data structure describing the RF to a 2-D texture array which is stored in the GPU texture memory. These texture arrays are read only, and, since they are cached, this improves the performance of reading operations (Grahn et al. 2011). The data associated with a tree is laid out in a four-component float texture, whereas the data of each node is stored on three separate columns in each channel of the texture array. We store the data of each node of a tree in the forest in sequential horizontal positions and different trees on separate rows. Data stored in the texture memory contains the position of the left and right child nodes, threshold values and all feature parameters required to evaluate the test for a node. If a node cached in the texture memory is a leaf node, then we add the probability distribution learned during training and the index of the leaf. To navigate through the tree during the evaluation, we use the tex2D function which performs a texture lookup in a given 2-D sampler based on 2-D node coordinates and channel information (Fig. 7.19). Our strategy involves launching a kernel which evaluates the probability of the random forest with the number of threads equal to the number of candidates. As the number of feature candidates can exceed the maximum number of threads per block with a maximum of 1024, 1536 and 2048 for compute capability 1.2/1.3, 2.x and 3.x, respectively, several thread blocks are launched.

In the following we present results for a spine structure detection application, whereas the 3D input image was obtained from microCT (computed tomography) and where the random forest classifier is used to detect lesions. After the system was trained, tests were run with three different implementations of the classification algorithm. The implementation was tested with two available data sets and performance benchmarks for our implementation have been compared with two CPU based implementations. The experimental results indicate that our GPU-based implementation of the Random Forest algorithm outperforms the two CPU based algorithms (CPU single-core and CPU multi-core).

To test the method for the bone lesion detection, we used a Random Forest detector consisting of 100 trees, 27K normalized candidates (voxels) for each vertebra and 12K features. The execution configuration of the classification kernel specifies

Table 7.10 Comparison of execution times of CPU and GPU based algorithms for the RF classifier

Implementation	Patient 1 (s)	Patient 2 (s)
CPU Multi-core	27.62 ± 0.39	48.60 ± 0.41
CPU Multi-core + GPU	17.85 ± 0.32	22.57 ± 0.34

Table 7.11 Total execution time for the detection system

Implementation	Patient 1 (s)	Patient 2 (s)
CPU Single-core	73.01 ± 0.58	182.33 ± 0.64
CPU Multi-core	10.19 ± 0.15	27.00 ± 0.25
CPU Multi-core + GPU	0.421 ± 0.02	0.979 ± 0.03

that the number of threads is equal to the number of candidates (27K), each block contains 128 threads (27K/128 blocks are used) and each thread traverses a candidate through all trees of the random forest.

Prediction time was compared on a machine equipped with Intel Core i7 CPU and a NVIDIA GPU GeForce GTX Titan Black. Table 7.10 compares the execution times of the CPU and GPU based algorithms. The speed-up varies between **170**× and **190**× compared to the single-core implementation and between **24**× and **28**×. Table 7.11 displays the total execution time which includes various other non-parallelized steps like the loading of the input image, etc. The RF based position detector classifier occupies the vast majority of the total execution time (~60%).

Acknowledgments The authors would like to thank Puneet Sharma, Ali Kamen and Dorin Comaniciu for their input.

References

Alastruey J, Khir A, Matthys K, Segers P, Sherwin S, Verdonck P, Parker K, Peiro J (2011) Pulse wave propagation in a model human arterial network: assessment of 1-D visco-elastic simulations against in vitro measurements. J Biomech 44:2250–2258. doi:10.1016/j.jbiomech.2011.05.041

Alerstam E, Svensson T, Andersson-Engels S (2008) Parallel computing with graphics processing units for high speed Monte Carlo simulation of photon migration (PDF). J Biomed Opt 13:060504. doi:10.1117/1.3041496

Bell N, Garland M (2008) Efficient sparse matrix-vector multiplication on CUDA. NVIDIA Technical Report NVR-2008-004

Bessems D (2008) On the propagation of pressure and flow waves through the patient-specific arterial system. PhD Thesis, Technical University of Eindhoven, Netherlands

Bessems D, Giannopapa C, Rutten M, van de Vosse F (2008) Experimental validation of a time-domain-based wave propagation model of blood flow in viscoelastic vessels. J Biomech 41:284–291. doi:10.1016/j.jbiomech.2007.09.014

Breiman L (2001) Random forests. Mach Learn 45:5–32

Chen G, Li G, Pei S, Wu B (2009) High performance computing via a GPU. In: Proceedings of the first international conference on information science and engineering, Shanghai, China, pp 238–241

Cococcioni M, Grasso R, Rixen M (2011) Rapid prototyping of high performance fuzzy computing applications using high level GPU programming for maritime operations support.

In: Proceedings of the 2011 IEEE symposium on computational intelligence for security and defense applications (CISDA), Paris, 11–15 April 2011

Courant R, Friedrichs K, Lewy H (1928) Uber die partiellen Differenzengleichungen der mathematischen Physik. Mathematishe Annalen 100:32–74

Cousins W, Gremaud PA, Tartakovsky DM (2013) A new physiological boundary condition for hemodynamics. SIAM J Appl Math 73(3):1203–1223

Formaggia L, Lamponi D, Quarteroni A (2003) One dimensional models for blood flow in arteries. J Eng Math 47:251–276. doi:10.1023/B:ENGI.0000007980.01347.29

Formaggia L, Lamponi D, Tuveri M, Veneziani A (2006) Numerical modeling of 1D arterial networks coupled with a lumped parameters description of the heart. Comput Methods Biomech Biomed Engin 9:273–288. doi:10.1080/10255840600857767

Fung Y (1993) Biomechanics: mechanical properties of living tissues. Springer, New York

Garcia V, Debreuve E, Barlaud M (2008) Fast k-nearest neighbor search using GPU. In: Proceedings of the CVPR workshop on computer vision on GPU, Anchorage, Alaska, USA

Grahn H, Lavesson N, Lapajne MH, Slat D (2011) CudaRF: a CUDA-based implementation of random forests. In: Siegel HJ, El-Kadi A (ed) AICCSA, IEEE, pp 95–101

Habchia C, Russeil S, Bougeard D, Hariona J-L, Lemenand T, Ghanem A, Della Valle D, Peerhossaini H (2003) Partitioned solver for strongly coupled fluid–structure interaction. Comput Fluids 71:306–319. doi:10.1016/j.compfluid.2012.11.004

Hasan Khondker S, Chatterjee A, Radhakrishnan S, Antonio JK (2014) Performance prediction model and analysis for compute-intensive tasks on GPUs. The 11th IFIP international conference on network and parallel computing (NPC-2014), Ilan, Taiwan, September 2014, Lecture Notes in Computer Science (LNCS), pp 612–617. ISBN:978-3-662-44917-2

Hestenes M, Stiefel E (1952) Methods of conjugate gradients for solving linear systems. J Res Natl Bur Stand 49:409–436

Itu L, Sharma P, Mihalef V, Kamen A, Suciu C, Comaniciu D (2012a) A patient-specific reduced-order model for coronary circulation. IEEE international symposium on biomedical imaging, Barcelona, Spain, pp 832–835. doi:10.1109/ISBI.2012.6235677

Itu LM, Sharma P, Kamen A, Suciu C, Moldoveanu F, Postelnicu A (2012b) GPU accelerated simulation of the human arterial circulation. In: Proceedings of the 13th international conference on optimization of electrical and electronic equipment, Brasov, Romania, pp 1478–1485. doi:10.1109/OPTIM.2012.6231764

Itu L, Sharma P, Ralovich K, Mihalef V, Ionasec R, Everett A, Ringel R, Kamen A, Comaniciu D (2013a) Non-invasive hemodynamic assessment of aortic coarctation: validation with in vivo measurements. Ann Biomed Eng 41:669–681. doi:10.1007/s10439-012-0715-0

Itu LM, Sharma P, Kamen A, Suciu C, Comaniciu D (2013b) Graphics processing unit accelerated one-dimensional blood flow computation in the human arterial tree. Int J Numer Methods Biomed Eng 29(12):1428–1455, ISSN:2040-7947

Janßen CF, Koliha N, Rung T (2015) A fast and rigorously parallel surface voxelization technique for GPGPU-accelerated CFD simulations. Commun Comput Phys 17(05):1246–1270

Jiang B, Struthers A, Sun Z, Feng Z, Zhao X, Zhao K, Dai W, Zhou X, Berens ME, Zhang L (2011) Employing graphics processing unit technology, alternating direction implicit method and domain decomposition to speed up the numerical diffusion solver for the biomedical engineering research. Int J Numer Methods Biomed Eng 27:1829–1849. doi:10.1002/cnm.1444

Jordan HF, Alaghband G (2003) Fundamentals of parallel processing. Pearson Education, Upper Saddle River, NJ

Kirk D, Hwu WM (2010) Programming massively parallel processors: a hands-on approach. Elsevier, London

Klages P, Bandura K, Denman N, Recnik A, Sievers J, Vanderlinde K (2015) GPU kernels for high-speed 4-bit astrophysical data processing. In: IEEE 26th international conference on application-specific systems, architectures and processors (ASAP)

Kumar R, Quateroni A, Formaggia L, Lamponi D (2003) On parallel computation of blood flow in human arterial network based on 1-D modelling. Computing 71:321–351. doi:10.1007/s00607-003-0025-3

Malecha Z, Mirosław L, Tomczak T, Koza Z, Matyka M, Tarnawski W, Szczerba D (2011) GPU-based simulation of 3D blood flow in abdominal aorta using OpenFOAM. Arch Mech 63:137–161

Malossi C, Blanco P, Deparis S (2012) A two-level time step technique for the partitioned solution of one-dimensional arterial networks. Comput Methods Appl Mech Eng 237:212–226. doi:10.1016/j.cma.2012.05.017

Manavski SA, Valle G (2008) CUDA compatible GPU cards as efficient hardware accelerators for Smith-Waterman sequence alignment. BMC Bioinformatics 9(suppl 2):S10. doi:10.1186/1471-2105-9-s2-s10. http://www.biomedcentral.com/1471-2105/9/S2/S10

Mittal S, Vetter J (2015) A survey of CPU-GPU heterogeneous computing techniques. ACM Comput Surv. doi:10.1145/2788396

Mynard JP, Nithiarasu P (2008) A 1D arterial blood flow model incorporating ventricular pressure, aortic valve and regional coronary flow using the locally conservative Galerkin (LCG) method. Commun Numer Methods Eng 24:367–417. doi:10.1002/cnm.1117

Mynard JP, Davidson MR, Penny DJ, Smolich JJ (2012a) A simple, versatile valve model for use in lumped parameter and one-dimensional cardiovascular models. Int J Numer Methods Biomed Eng 28:626–641. doi:10.1002/cnm.1466

Mynard JP, Penny DJ, Davidson MR, Smolich JJ (2012b) The reservoir-wave paradigm introduces error into arterial wave analysis: a computer modelling and in vivo study. J Hypertens 30:734–743. doi:10.1097/HJH.0b013e32834f9793

Nita C, Itu LM, Suciu C (2013) GPU accelerated blood flow computation using the lattice Boltzmann Method. In: 17th IEEE high performance extreme computing conference, Waltham, MA, USA

Nita C et al (2015) GPU-accelerated model for fast, three-dimensional fluid-structure interaction computations. Engineering in Medicine and Biology Society (EMBC), 2015 37th annual international conference of the IEEE. IEEE

Olufsen M, Peskin C, Kim WY, Pedersen EM, Nadim A, Larsen J (2000) Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions. Ann Biomed Eng 28:1281–1299. doi:10.1114/1.1326031

Ortega J (1988) Introduction to parallel and vector solution of linear systems. Plenum Press, New York

Owens JD, Houston M, Luebke D, Green S, Stone JE, Phillips JC (2008) GPU computing. Proc IEEE 96:879–884

Passerini T (2009) Computational hemodynamics of the cerebral circulation: multiscale modeling from the circle of Willis to cerebral aneurysms. PhD Thesis, Politecnico di Milano, Italy

Raghu R, Vignon-Clementel I, Figueroa CA, Taylor CA (2011) Comparative study of viscoelastic arterial wall models in nonlinear one-dimensional finite element simulations of blood flow. J Biomech Eng 133:081003. doi:10.1115/1.4004532

Rahimian A, Lashuk I, Veerapaneni S, Chandramowlishwaran A, Malhotra D, Moon L, Sampath R, Shringarpure A, Vetter J, Vuduc R, Zorin D, Biros G (2010) Petascale direct numerical simulation of blood flow on 200K cores and heterogeneous architectures. In: Proceedings of the ACM/IEEE international conference for high performance computing, networking, storage and analysis, New Orleans, USA, pp 1–11

Reymond P, Bohraus Y, Perren F, Lazeyras F, Stergiopulos N (2011) Validation of a patient-specific one-dimensional model of the systemic arterial tree. Am J Physiol Heart Circ Physiol 301:1173–1182. doi:10.1152/ajpheart.00821.2010

Ryoo S, Rodrigues CI, Stone SS, Stratton JA, Ueng S-Z, Baghsorkhi SS, Wen-mei WH (2008) Program optimization carving for gpu computing. J Parallel Distrib Comput 68(10):1389–1401

Saad Y (ed) (2003) Iterative methods for sparse linear systems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia

Sato D, Xie Y, Weiss JN, Qu Z, Garfinkel A, Sanderson AR (2009) Acceleration of cardiac tissue simulation with graphic processing units. Med Biol Eng Comput 47(9):1011–1015

Schalkwijk J, Harmen JJ, Jonker A, Siebesma P, Van Meijgaard E (2015) Weather forecasting using GPU-based Large-Eddy simulations. Bull Am Meteorol Soc 96:715–723

Schatz MC, Trapnell C, Delcher AL, Varshney A (2007) High-throughput sequence alignment using Graphics Processing Units. BMC Bioinformatics 8:474

Sengupta S, Harris M, Garland M (2008) Efficient parallel scan algorithms for GPUs. NVIDIA Technical Report NVR-2008-003

Senzaki H, Chen CH, Kass DA (1996) Valvular heart disease/heart failure/hypertension: single-beat estimation of end-systolic pressure-volume relation in humans: a new method with the potential for noninvasive application. Circulation 94:2497–2506

Shams R, Sadeghi P, Kennedy RA, Hartley RI (2010) A survey of medical image registration on multicore and the GPU. IEEE Signal Process Mag 27(2):50–60

Shen W, Wei D, Xu W, Zhu X, Yuan S (2009) GPU-based parallelization for computer simulation of electrocardiogram, CIT '09. In: Ninth IEEE international conference on computer and information technology, October, Xiamen, China

Steinman DA et al (2013) Variability of computational fluid dynamics solutions for pressure and flow in a giant aneurysm: the ASME 2012 Summer Bioengineering Conference CFD Challenge. J Biomech Eng 135.2:021016

Stergiopulos N, Young DF, Rogge TR (1992) Computer simulation of arterial flow with applications to arterial and aortic stenosis. J Biomech 25:1477–1488. doi:10.1016/0021-9290(92)90060-E

Tanno I, Morinishi K, Satofuka N, Watanabe Y (2011) Calculation by artificial compressibility method and virtual flux method on GPU. Comput Fluids 45:162–167. doi:10.1016/j.compfluid.2011.02.005

Vardoulis O, Papaioannou T, Stergiopulos N (2012) On the estimation of total arterial compliance from aortic pulse wave velocity. Ann Biomed Eng 40:2619–2626. doi:10.1007/s10439-012-0600-x

Verschoor M, Jalba A (2012) Analysis and performance estimation of the conjugate gradient method on multiple GPUs. Parallel Comput 38:552–575

Westerhof N, Elzinga G, Sipkema P (1971) An artificial arterial system for pumping hearts. J Appl Physiol 31:776–781

Willemet M, Lacroix V, Marchandise E (2011) Inlet boundary conditions for blood flow simulations in truncated arterial networks. J Biomech 44:897–903. doi:10.1016/j.jbiomech.2010.11.036

Yu R, Zhang S, Chiang P, Cai Y, Zheng J 2010 Real-time and realistic simulation for cardiac intervention with GPU. In: Second international conference on computer modeling and simulation, Sanya, China, pp 68–76

Zaspel P, Griebel M (2013) Solving incompressible two-phase flows on multi-GPU clusters. Comput Fluids. doi:10.1016/j.compfluid.2012.01.021

Zhang D (2000). Applying machine learning algorithms in software development. In: Proceedings of Monterey workshop on modeling software system structures, Santa Margherita Ligure, Italy, pp 275–285

Zhang Y, Cohen J, Owens JD (2010) Fast tridiagonal solvers on the GPU. In: Proceedings of the 15th ACM SIGPLAN symposium on principles and practice of parallel programming, Bangalore, India, pp 127–136

Zou C, Xia C, Zhao G (2009) Numerical parallel processing based on GPU with CUDA architecture. In: International conference on wireless networks and information systems, WNIS'09, IEEE, pp 93–96