

Clinical Physiology

A muscle centered approach

– third draft –

Sean Collins
Professor of Physical Therapy
Plymouth State University

May 22, 2024

A book by PT Centered Press

Copyright ©2024 Sean Collins

ISBN (eBook): 979-8-9907667-0-9

ISBN (Paperback): 979-8-9907667-1-6

PT Centered Press
Boscawen, NH, 03303

Permission is granted to copy, distribute, and/or modify the text of this document under the terms of Creative Commons BY-NC-SA 4.0 DEED Attribution-NonCommercial-ShareAlike 4.0 International, which is available at <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en>. This copyright only refers to the text of the book. Restrictions exist for figures rendered with BioRender and for permission to use any figure created with BioRender please contact the author for details. Permission to recreate a figure, capturing its essential concepts for example, is given through the CC BY-NC-SA 4.0 DEED license, but use of an exact figure which has been created in BioRender is restricted by the terms of the authors BioRender Academic License. In such instances this will require the requester to secure their own BioRender Account and follow the terms of that account.

The original form of this book is $\text{\LaTeX}$ source code. Compiling this $\text{\LaTeX}$ source has the effect of generating a device-independent representation of a textbook, which can be converted to other formats and printed. The code that developed this book is stored in a GitHub repository and can be made available upon request to the author at: scollinspt@tuta.com

To God. For sustaining grace and the ultimate questions.

To my parents, Nancy and Joe. For teaching me pragmatism.

To my sons, Tim & Andrew. For answering the ultimate question about unconditional love.

To my friend, Gary. For elevating my thinking to the ultimate questions.

To my wife, Traci. For grounding me in the reality in which ultimate questions matter.

Preface

Clinical Physiology: A muscle centered approach is a manuscript (book) in progress. The draft you are reading was released May 22, 2024.

This project emerged late in my sabbatical year, early April 2022. I had been working on a book for my course sequence titled *Clinical Inquiry* and was unable to break through some theory to practice barriers. I had been meeting for weekly discussions with my friend Bog Sniezek who is a systems scientist that was at Draper Labs in Cambridge and is now a consultant for the Air Force. During our conversation about the nature and function of unifying systems (Bog's theory), he inspired some thoughts about how physical therapists should learn clinical physiology. It is foundational knowledge for all areas of practice. The idea that struck me based on unifying systems theory was a muscle centered approach. From the perspective of a physical therapist, physiology is the generative (muscle cell) and supportive (other systems) role in movement. Tension developed by muscles, which is developed by muscle cells, is at the core of movement. Muscle cells developing tension is a necessary, but not sufficient, condition for movement. Muscles have always been included in the course I teach on clinical physiology. In fact, there's no new material in the book from the perspective of what's being taught. The only thing new is that all things are considered from the perspective of the muscle fiber. Since my sabbatical was nearing completion, and I did not have a draft of my original plan (a book on clinical inquiry), and I had a *Clinical Physiology* course to teach in a couple of months; I made the decision to start working on this text. In my 25 year history of teaching this topic I was never satisfied with the textbook options. It was a Goldilocks situation. I either spent class time explaining to students what they did not need to know and adding topics related to movement to a medical physiology textbook; or I spent time adding clinical topics to an exercise physiology textbook. Either way, it discouraged students from reading the book and focusing on the lectures. However, the lectures were dense and students would often say such things as: "*I wish I had a transcript of the lecture.*" This book is a transcript of those lectures. A secondary goal of this book is that students will learn how to learn through graduate level academic reading.

A muscle centered approach puts the muscle fiber (cell) at the center of inquiry. **Part I** establishes preliminaries to put the muscle fiber in context. **Part II** focuses on what a muscle fiber does as a system. Muscle fibers create tension, a specific type of force. The process of creating tension is an inward looking process. We consider the cellular, mechanical and chemical mechanisms of the muscle fiber. **Part III** focuses on the necessary supporting systems for the muscle fiber. Each muscle fiber receives and releases material from its surrounding environment (extra cellular fluid). Several physiological systems integrate to ensure the stability of that environment. Part I through III are covered in the author's DPT course, *Clinical Physiology*, at Plymouth State University.

Part IV (in progress) will focus on muscle molecular biology. Specifically, how the muscle continues to fulfill its role of creating tension without being too large of a burden on the physiological support systems. Topics include myostasis, atrophy, hypertrophy, and repair. **Part V** (in progress) will include a variety of integrated exercises. These topics have been selected to highlight the need for physiological coordination.

Clinical Physiology in a DPT curriculum

A DPT program has the challenge of educating individuals from diverse backgrounds in the core set of knowledge to become physical therapists. Clinical Physiology is foundational knowledge. The book fits best in a first term course, or a

first course on human physiology. The book assumes the reader has taken pre requisite courses in anatomy & physiology, physics, and chemistry. It is not an exercise physiology text, though it has a lot of content that would be found in an exercise physiology text. It is not a medical physiology text, though it has a lot of content that would be found in a medical physiology text. The book includes foundational knowledge of physiology for a physical therapist. That knowledge falls in a previously unmet textbook space that combines exercise physiology and medical physiology.

Tips for the Reader & Student

Read actively. Question what you read as you read. Engage in what you read while you read. Every chapter begins with a brief introduction to orient your mind towards what you'll be reading about; a mini table of contents so you can see the general layout; and objectives for you to start to think about what you'll be learning. It is recommended that you start by reading these and then skim through the rest of the chapter as a second step. This process helps you to read actively. During your pre-read skim you should make mental note of the headings, subheadings, figures, and equations. The purpose of skimming is to get a sense of what is coming up as you start to read the chapter. When you start reading, read as if you are listening to the author speak. If you get a question in your head as your listening to me, write it down. Proceed to read even if you're not completely sure you understand what you're reading. Sometimes a concept becomes more clear as you proceed. If not, make a note of what you're not understanding and bring it to class. Identifying and clearing confusion is one of the purposes of class time, the second purpose of class time is application, depth and breadth.

Every chapter ends with a summary that includes the *Next Steps* to help you understand what has been covered and where we are headed. At the end of each chapter you will also find sample questions. You should be comfortable answering these questions. But you should also be comfortable manipulating them, turning them into other versions, making them easy or hard multiple choice questions based on the response options. Often the difficulty of a multiple choice question is not the question, the neighbors of the correct response option. Use the review questions to drive discussions during group study time. Yes, study in groups. Group study is the only thing you can do beyond class time to make sure you're not stuck in your own biased trap of understanding. You do not know what you do not know. You need others to bring you from unconscious incompetence to conscious incompetence, the first step towards conscious competence.

A Note about Chapter Objectives

The chapter objectives include specific behavioral terms. There is an ordered relationship between the behavioral terms that represents increasing and nested expectations.

To **define** is to know the words being used, to know what they represent. Whether you can define something can be tested with a straightforward question such as identifying the correct definition from a set of definitions; whether you can write the definition; respond true or false to a proposed definition for a term; or identify the correct term from a set of terms when presented with a definition.

To **explain** goes beyond defining, but it also implies that you can define the required terms. Explaining is the act associated with understanding. If you understand something you can explain it. Objectives are typically written as acts - what the person learning can do if they have achieved the objective. Since understanding is observed as an act of explaining, we will use the objective "explain" when there is something you must "understand". How do you demonstrate that you understand? You explain whatever it is that you understand. To explain you simultaneously hold a set of definitions together including their relationship to one another. You can explain a model by demonstrating that you know all the definitions involved, and describe how they relate to one another, and can communicate that explanation. Whether you can explain something can be tested by asking about several aspects of what you are to explain or asking you to infer which part is missing; or by asking about one part and asking you to infer the other part. The point is, to explain goes beyond defining and has an expectation of definition. But to define does not require you to explain.

To **evaluate** requires you to make judgements and to infer beyond the thing you are asked to evaluate. To evaluate implies that you can define and explain the thing (or concept) you are asked to evaluate. Being tested for whether you can evaluate can include asking you to infer to or from something that has not been directly covered or discussed, going beyond the thing (or concept) you are asked to evaluate to consider its implications in different situations.

Providing an example is an objective based on a specific task and is considered a form of asking you to evaluate since it goes beyond a particular thing (or concept) to a totally new thing (or concept). Asking you to provide an example of a model requires you to know the concept of a model well enough that you can use your background knowledge to come up with something you're already familiar with that fits what you know about models. That level of knowledge requires you to evaluate a model because you evaluated your background knowledge and determined that it is a model. If a chapter objective asks you to provide an example of a model that is useful to physical therapy practice it assumes you will be able to define, explain and evaluate the concept of a model, and that you have background knowledge about physical therapy practice that you can use to evaluate models used in practice to consider what those models are and whether they are useful. Providing an example is similar to the objective **Apply**. With both of these objectives you are being asked to use your knowledge of the general concept (or thing) to a specific instance. If the concept was fruit, then an example is an apple. And you can apply the the concept of fruit by listing several fruits including what makes them all fruits.

Compare and contrast is also an objective based on a specific task that requires evaluation of two (or more) things (or concepts) to identify what is alike and different about them. This can occur with the general concept (e.g., fruit) by comparing fruits with vegetables, or with specifics (apples) by comparing different specific fruits (apples and oranges), or by comparing different specifics from each concept (apples and broccoli). Comparing and contrasting is an important higher level evaluative process that helps you to develop a deeper understanding.

Updates to the Third Draft

The third draft of this manuscript is now released as a single PDF as opposed to separate chapters. Therefore it comes complete with front matter (though no back matter). Each chapter now has a summary as well as a few (at least 5) sample questions at the end of the chapter. This draft is still not a completed manuscript, but it is ready to get feedback from colleagues for the purpose of peer review and possibly sharing with publishers. I had hoped to complete a glossary, index and increased between chapter contextual links, but alas those are delayed for another draft. As with each draft, I must take complete ownership of all faults found herein. All I can say to the reader is *mea culpa, veniam peto*.

Boscawen, New Hampshire
May 2024

Sean Collins

Contributors

About the Author

Sean Collins is a Professor of Physical Therapy and practicing physical therapist with 30 years of clinical, teaching, research and writing experience in the area of clinical physiology. He earned a Doctor of Science and won the Lorin Kerr Award for Excellence in Ergonomics Research from the Department of Work Environment in the College of Engineering at the University of Massachusetts Lowell for his dissertation on Job Strain & Electrocardiograph Assessed Pathophysiologic Correlates in 2003. He is Editor Emeritus of the Cardiopulmonary Physical Therapy Journal and in 2018 was awarded the Linda Crane Lecture Award from the American Physical Therapy Association for making outstanding and enduring contributions to the practice of physical therapy for his use of causal models and reasoning in PT education. He has over 50 publications and has presented at numerous national and international conferences. He was a presenting panel member on the use of heart rate variability at the International Conference of Occupational Health in Los Angeles in 2005; gave the Keynote lecture on the biopsychosocial model of health in Athens Greece at Euro prevent 2010 for the European Cardiovascular Society; was presenting panel member for the use of electrocardiography in the measurement of stress in Moscow Russia at the Noninvasive Electrocardiography Society in 2011; and was an invited lecturer for the Danish National Institute of Occupational Health on the measurement of stress biomarkers in 2012. For six years he directed a lab at UMass that secured over 100k in funding from the Department of Defense and industry partners to support the research and development of ambulatory electrocardiography and vital sign monitoring. He has taught physiology at the undergraduate, professional and PhD level for over 20 years.

Acknowledgements

First, I must acknowledge the role that Bog Sniezek played to inspire the thinking that led to the muscle centered approach and the continued conversations during the first draft of this manuscript in 2022. Second, I must acknowledge all of the students that I've taught over the years. It is the rigor that they bring to their studies that inspires me to be better. Third, I must acknowledge the Plymouth State University DPT Classes of 2025 and 2026 for their patience, tolerance and encouragement during the first and second drafts of the text. Finally, I'd like to acknowledge my colleagues and the administration at Plymouth State University for providing me with an environment that both supports and inspires creative and critical inquiry for such projects to flourish. Without the sabbatical in the 2021-2022 academic year I would not have recovered from the burnout associated with founding the program and had the time to start this project. And without the support of the faculty, particularly Dr. Barbara Crane for assuming the role of Program Director to free me from the shackles of administrative burdens, this project would be impossible.

Contents

Part I Preliminaries Overview

1	Introduction	3
1.1	Clinical Physiology	3
1.2	Basic Concepts of Physiology	6
1.3	Summary	12
2	Fundamentals	15
2.1	Muscle Fibers and Muscles	16
2.2	Attaining & Sustaining Tension	16
2.3	Connective Tissue Scaffold	17
2.4	Muscle Mechanics	18
2.5	Muscle Roles	23
2.6	Summary	24

Part II Muscle Function Overview

3	Tension	27
3.1	Anatomy of Tension	27
3.2	Passive Tension	29
3.3	Active Tension	31
3.4	Force & Velocity	37
3.5	Summary	40
4	Excitation	41
4.1	Muscle Fiber Excitation	42
4.2	Excitable Membranes	50
4.3	<i>Clinical Physiology Connections</i>	56
4.4	Summary	59
5	Regulation	61
5.1	Motor Units	62
5.2	Motor Unit Twitch & Tetany	67
5.3	Active Tension Regulation	69
5.4	Regulatory Feedback	72
5.5	<i>Clinical Physiology Connections</i>	73
5.6	Summary	75
6	Energetics	77
6.1	Energetic Transformation Overview	78
6.2	ATP Regeneration Pathways	78
6.3	Motor Unit & Muscle Fiber Types	85
6.4	<i>Clinical Physiology Connections</i>	86

6.5	Summary	91
Part III Muscle Support		
7	Micro-Circulation	97
7.1	Micro Circulation	98
7.2	Filtration	104
7.3	<i>Muscle Connections: Smooth Muscle</i>	110
7.4	<i>Clinical Connections: Compression</i>	112
7.5	Summary	112
8	Renal Clearance	115
8.1	Mass Balance	116
8.2	Renal Function Overview	119
8.3	Filtration-Reabsorption-Secretion	123
8.4	Renal Regulation	126
8.5	Disturbances in Fluid Volume & Osmolarity	129
8.6	<i>Muscle Connections</i>	138
8.7	Summary	138
9	Circulation	141
9.1	Structures of Circulation	142
9.2	Cardiac Output	143
9.3	Regulation of Cardiac Output	145
9.4	Blood Pressure Regulation	147
9.5	Cardiovascular Vital Signs	152
9.6	Summary	154
10	Cardiac Pump	157
10.1	Cardiac Muscle	157
10.2	Cardiac Cycle	167
10.3	ECG Rhythm Analysis	169
10.4	Summary	171
11	Respiration	173
11.1	Respiration Overview	174
11.2	Cellular & Internal Respiration	174
11.3	Blood Transport	179
11.4	Pulmonary Circulation	184
11.5	External (Alveolar) Respiration	186
11.6	Practice Connections	188
11.7	Summary	190
12	Ventilation	193
12.1	Ventilation Overview	194
12.2	Airways	195
12.3	Volumes & Capacities	200
12.4	Mechanics: The Ventilatory Pump	201
12.5	Ventilation / Perfusion Matching	210
12.6	Regulation	213
12.7	Practice Connections	216
12.8	Summary	218
Next Steps		219

Part I

Preliminaries Overview

These first two chapters cover context for the rest of the book. Chapter 1 is an introduction to clinical physiology including what it means to take a muscle centered approach. It then reviews the core concepts of physiology. Chapter 2 includes the broader context for muscle fibers. Specifically, how they connect together and some of the characteristics they display when working together as muscles. Physical therapists are interested in muscle fibers because when connected together they form muscles which create movement.

Chapter 1

Introduction

Updated on May 22, 2024

Contents

1.1	Clinical Physiology	3
1.2	Basic Concepts of Physiology	6
1.3	Summary	12

This chapter introduces basic concepts for knowing and applying clinical physiology in physical therapy practice. First, it covers what is meant by clinical physiology. Second, it reviews the basic concepts of physiology.

Objectives:

1. Explain a muscle centered approach to clinical physiology for the practice of physical therapy
2. Provide an example of a model that is useful in physical therapy practice.
3. Explain the basic concepts of physiology.
4. Explain how whole system function and capacity relates to multi system function and capacity.
5. Provide an example of how each basic concept of physiology applies to the analysis of patient/client problems.
6. Explain the hierarchy of adaptation and adaptability from genetic, epigenetic, anatomic, physiologic, behavioral and cultural and its relevance in physiological adaptation.
7. Explain how the International Classification of Function (ICF) clinical physiology relates to the hierarchy of adaptation and muscle centered approach..

1.1 Clinical Physiology

Clinical physiology focuses on physiology relevant to understand health and "un" health (injury, illness, disorder, dysfunction, disequilibrium). This is a clinical physiology text for physical therapists. Since physical therapists are movement specialists, and since the physiology of muscles is integral to the movement system, the approach to clinical physiology is a muscle centered approach. A muscle fiber is a system, and muscle fibers are the structural unit of muscle.

Muscle Fibers are Muscle Cells

Muscle fibers are muscle cells, and muscle cells are muscle fibers. We do our best throughout the book to refer to muscle fibers, but if we refer to a muscle cell please know that this equates to saying muscle fiber. Similarly, when we talk about the general biological properties of cells, please know that these properties are true for muscle fibers.

A muscle fiber creates tension. It interacts with other systems and is supported in order to create tension. Clinical physiology for a physical therapist considers muscle tensioning, what and how a muscle fiber interacts with other systems, and what support is required for the muscle fiber to create tension.

Muscle fibers create tension for the purpose of generating movement. To create tension muscle fibers interact with other fibers and are supported by other systems. Interactions and support are an important part of the book. Muscles get a lot of their support from the extra cellular fluid (ECF), and maintain their ability to create tension through cellular and structural integrity.

Clinical physiology for physical therapy considers the three interactions depicted in Figure 1.1. These interactions are supporting relationships. The interaction between the muscle fiber and movement is an undeniable part of the movement specialist's domain of knowledge and physical therapy students have little problem understanding its importance. The relationship between the ECF and the muscle fiber, and between movement and the ECF are less obvious and deserve additional explanation. The following sections explain these supporting interactions.

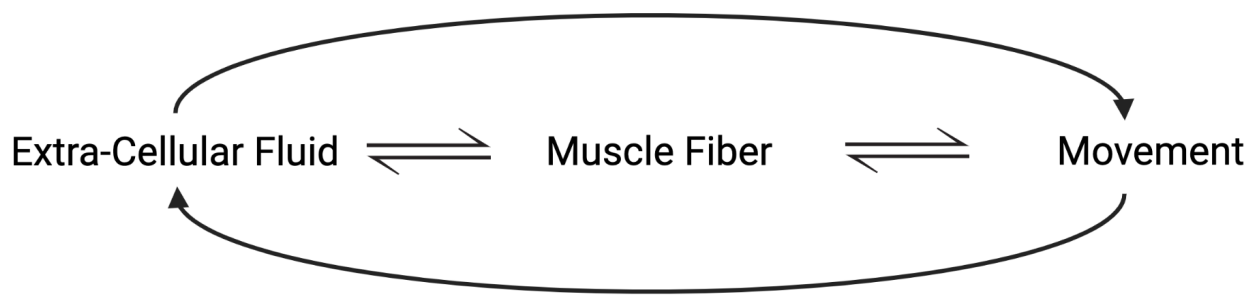


Fig. 1.1: A Muscle Centered Approach (Created with BioRender.com)

1.1.1 Muscle Fibers & Movement

Muscle fibers are necessary but not sufficient causes for human movements. Movements include an intricate combination of active and passive tension that is created by muscle fibers. However, muscle fibers cannot create tension for movement in isolation. First, muscle fibers must come together as a muscle and combine their ability to create tension in order to create movement. They create tension with interaction to attachments, mostly, to bones¹ that form joints that are capable of movement. The process of coming together, and creating tension and transmitting tension to bones is covered in the next two chapters (Chapter 2 on Fundamentals & Chapter 3 on Tension). Other aspects of movement such as the mechanics and kinematics of the musculoskeletal system, or the coordination, control and learning of the neuromotor and behavioral systems, are not covered in this book. To create tension muscle fibers need to possess the capabilities of being excited (Chapter 4 on Excitation), being regulated to create the right amount of tension (Chapter 5 on Regulation), and converting energy (Chapter 6 on Energetics). It is important to keep in mind that these capabilities are characteristics of muscle fibers. They exist in muscles because they exist in muscle fibers.

1.1.2 Extra-Cellular Fluid & Muscle Fibers

The capability to create tension requires support. Support is received from the surroundings of the muscle fiber, the extra cellular fluid (ECF). ECF is supported and tightly regulated through the activity of several physiological systems. In this

¹ Mostly to bones because sometimes muscles attach to something other than bones such as the eye muscles or facial muscles that attach to skin to create facial expressions, or muscles that attach to themselves.

sense, as depicted in Figure 1.1, the ECF supports muscle tension.² When a muscle fiber creates tension it utilizes and exchanges resources with the ECF, thus muscle cells thus rely on and contribute to ECF. These systems that refresh and regulate the ECF are the focus of several chapters in Part III on Muscle Support.

? Questions

What are some substances that enter a muscle fiber from the ECF for it to do what it needs to do as a muscle fiber? What are some substances that leave the muscle fiber and enters the ECF while it is doing what it does?

1.1.3 Movement and Extra-Cellular Fluid

The ECF contains and sustains the correct materials for supporting and supplying muscle fibers. The ultimate source of materials in the ECF is movement (Figure 1.1), meaning, movement supports the ECF. It may be the movement of the heart, the movement of the arterioles that regulate blood flow, the movement of the breathing muscles that move air into and out of the lungs, the movement of limbs to get food and water, the movement of the jaw and esophagus and gastric tube to digest and move food and water for absorption, the movement of the colon for waste products. These have all been examples of movements caused by muscle tension in support of the ECF that is ultimately supporting the muscle fibers causing the muscle tension. Conversely, the ECF directly supports movement beyond bathing muscle fibers in a supportive environment by providing the environment for all tissues to work together for coordinated movement of the entire body.

1.1.4 Models

Figure 1.1 constitutes a model. Models are simplified abstractions of reality. This model does not include all the details. It is not supposed to include all the details. If it included all the details it would be overwhelming. Hashing out just some of the detail requires the rest of this book. Hashing out more of the details includes most of your education. Hashing out all of the details is not currently possible and requires more than one lifetime.³ Hashing out the details we continue to create and consider additional models, over and over again, models of models of models of models as we successively approximate something useful to understand and knowledge to base practice.

There are numerous models throughout this book. Every figure provided is a model. With each model we consider a quote by George Box: “All models are wrong some models are useful.” The muscle centered approach is our first step. Three variables and interrelationships provide a model of clinical physiology as a guide to reasoning for physical therapists in the context of injuries, conditions, and diseases. This is a useful model. There are certainly going to be situations that don’t fit the model since not all details are contained in the model. But for the task of learning clinical physiology to practice physical therapy, it’s useful.

? Questions

What are some models that have been useful to your learning up to this point? Can you think of any useful models in physical therapy practice? Can you think of some reasons that models are useful for clinical reasoning?

² It is admittedly much more complicated than this and the use of the term causes can be challenged, but the assertion is that without the proper ECF muscle active tension cannot occur. It is true that muscle passive tension can occur with a passive stretch of the muscle but such a passive stretch is typically the product of another system putting energy into the muscle to create the passive stretch and that other system putting energy into the muscle often relies in some way on ECF.

³ The lifetime of many scientists has already been spent on hashing out the details that we currently know. Some may say that all the details is computationally intractable considering the depth of mechanisms at the molecular level.

1.1.5 Clinical Physiology Summary

Clinical physiology from a muscle centered approach starts with a simple model. Based on this model if a PT considers muscles as a cause of movement but doesn't consider the factors that influence the ECF that is critical to the excitation, regulation, and energetics⁴ of muscle fibers then they are ignoring a critical component. If a PT considers how the ECF influences muscle fibers and thus movement and does not consider how continued movement influences the ECF they may miss the opportunity to stop repeated hospital admissions.⁵ Considering these relationships is a first step to thinking about clinical physiology with a muscle centered approach for the purpose of practicing physical therapy.

1.2 Basic Concepts of Physiology

The basic concepts of physiology provide a unifying framework applied throughout the book. They summarize pre-requisites for reading this book. It is expected that physical therapy students are already familiar with the basics of physical, chemical and biological sciences. This section explains the basic concepts and how they are useful to the physical therapist. Basic concepts of physiology are applied to the analysis of patient/client problems and are a foundation for understanding altered physiological states (pathophysiology).

1.2.1 Causality

Living organisms have causal mechanisms whose functions are explainable by a description of cause-effect relationships. Identifying and specifying causal relationships and models (to understand mechanisms) are an important part of physical therapy practice [Col18]. Causal mechanisms are learned through rigorous manipulation and observation (experiment) or systematic observations (epidemiology) and subjecting these observations to statistical inference. There is also an assumption that underlying any effect we observe there is a cause, or a set of inter-related causes. This is an essential aspect of causality in physiology.

Analysis of Patient/Client Problems

Causality is applied to the analysis of patient/client problems through the very fact that if we observe an effect we infer some underlying cause. If there is an alteration in blood pressure, glucose or sodium we infer something is the cause of those alterations. Inference from effects to causes is called abduction, and when there are multiple possible causes from a set of effects, is called inference to the best explanation. The statistical inference used to calculate the probabilities of inferences from causes to effects is called Bayesian inference and it is the foundation of any diagnostic process. Muscles generate tension. Tension is an effect. Part II considers the causes of that tension. If muscles generate too much or too little tension (an effect that is noticed as muscle stiffness or muscle weakness) then we consider the causes of the tension and what may be wrong. What is the cause of muscle stiffness? What is the cause of muscle weakness? Not just any muscle stiffness or weakness, but the muscle stiffness or weakness being observed at the moment of observation. To consider the cause in the particular situation requires an understanding of potential causes.

⁴ The convention we will utilize is that word "energetic" is an adjective, something is energetic, marked by vigor or effect, related to energy. However, adding an 's' to the end of energetic, energetics changes it to a verb, the process of converting energy and should not be confused as a plural form of the adjective energetic.

⁵ Such as readmissions due to repeated breakdowns in a person that goes home, cannot move enough to sustain hydration or nutrition or ventilation and returns to the hospital dehydrated, malnourished and with pneumonia [CD15]

1.2.2 Cells

1.2.2.1 Cell Theory

Cell Theory states that all cells making up an organism have the same DNA. Cells are considered the basic unit of life. All of the specialized functions we cover in this book are based on differentiation and the subsequent specialization of cells throughout the body.

Analysis of Patient/Client Problems

Cell Theory is applied to the analysis of patient/client problems through the observation of the health and well being of cells as an indication of whether there are problems, first with the cell, and second as a possible abnormal cause. For example, any cells that divide at a rapid rate (which itself is the result of DNA expression specific to the specialization of the cell), such as skin cells, have a greater chance of DNA mutation during cell division. A DNA mutation during skin cell division can result in a variety of benign or malignant cancerous cells that appear on the surface of the skin. The presence of a mole, particularly one that is not circular or has rough edges, is a possible sign of altered DNA of those cells creating cancerous cells. A benefit to having a high rate of cell division is ability of these cells to replaced dead or damaged cells resulting in faster healing and complete repair as opposed to an adaptive repair (such as scar tissue). Rates of cell division are also related to rates of healing. Chondrocytes, cells that spawn cartilaginous material, and some nervous system cells are known to have slower or absent rates of healing and repair. Osteoarthritis is a "wear and tear" of the joints. The rate of wear and tear exceeds the rate of healing and repair. This influenced by the rate of cellular division in those specialized cells.

1.2.2.2 Cell Membrane

The boundary of a cell is delimited by a semi permeable plasma membrane called the cell membrane. In a muscle fiber it is the sarcolemma. The sarcolemma is the boundary between the inside of the cell (cytoplasm, for a muscle fiber the sarcoplasm) and the outside of the cell (ECF). It determines what substances enter or leave the muscle fiber. The sarcolemma is essential for cell signaling, transport and other processes. The proteins (receptors, channels, pumps) embedded in the sarcolemma influence the permeability of the membrane to certain substances. Therefore influences the rate those substances enter (or exit) the cell. Adaptations of the muscle fiber sarcolemma to exercise training can allow events such as excitation to occur more rapidly. To increase the rate of excitation the sarcolemma can include additional sodium/potassium pumps, more sodium and potassium channels that allow faster excitation cycling (discussed in much more detail in Chapter 4 on Muscle Excitation).

Analysis of Patient/Client Problems

The cell membrane is applied to the analysis of patient/client problems in a rather fundamental way. Whether a cell is a cell includes whether it has a boundary. If damage to the cell membrane is complete, the cell no longer exists, it is dead. The assessment of cardiac cell death (infarction) includes whether certain proteins that are found in the cardiac muscle are in the blood. The reasoning is that for those proteins to be in the blood, they had to escape the cardiac muscle and that can only happen if the cell membrane is irreversibly damaged, and if the cell membrane is irreversibly damaged then the cell is dead. Therefore, if there is Troponin I or T in the blood, or Creatine Kinase - Myocardial Band (CK-MB) in the blood, then some sort of cardiac muscle fiber infarct has likely occurred.

1.2.2.3 Cell-Cell Communication

The function of any multi-cellular organism requires the ability to coordinate individual action using the most concise means. This can be achieved by cells passing information to one another through "cell-cell communication". These communication processes include endocrine (hormones delivered system wide in the blood which alert all muscle fibers

receiving blood flow that have a receptor within their sarcolemma) and neural signaling (which tends to more specifically target particular muscle fibers).

Analysis of Patient/Client Problems

Cell-cell communication is applied to the analysis of patient/client problems in any situation where endocrine or nervous system interactions with the cells may be abnormal. The release of insulin signals to cells that they should increase their glucose uptake. If cell-cell communication is impaired, as with Type II Diabetes, and the insulin receptors on the cell do not adequately receive that communication. Then cells will not uptake more glucose in response to insulin and blood glucose values will increase. One possible cause of an elevated blood glucose is a problem with cell-cell communication of the insulin receptors not responding to insulin.

1.2.3 Genes to proteins

Genes (DNA) code for the synthesis of proteins (muscle filaments and enzymes are proteins). This is not all the genome⁶ or the epigenome⁷, does, but it is an important part of muscle cell differentiation, specialization and myostasis. Genes in a particular cell that are expressed determine the functions of that cell (e.g., make it a muscle fiber and not, say, a liver cell), and additionally can make it a particular type of muscle fiber (e.g., a fast or slow twitching muscle fiber).

Analysis of Patient/Client Problems

Genes to proteins is applied to the analysis of patient/client problems whenever we consider the impact of cancer (mutations in DNA during division that then influence the proteins being created) on cellular function. Genes to proteins is also the fundamental process by which muscle fibers are maintained (myostasis - isotrophy, staying the same) or adapt (myostasis - hypertrophy or atrophy). Mechanisms generating tension in the muscle fiber are dependent on proteins, and the proteins are produced during the translation and transcription of DNA.

1.2.4 Energy

The life of the organism requires the constant expenditure of energy. The acquisition, transformation, and transportation of energy are essential functions. Experimental approaches to understand muscle are focused on muscle fibers as energy conversion machines, taking chemical energy and converting it into mechanical energy [WCH85]. The earliest of these studies measured work and heat based on the physical law of the conservation of energy. Energy is neither created or destroyed. Energy is transformed from one form to another. When a muscle fiber breaks down adenosine triphosphate (ATP) to generate tension, as discussed in Chapter 6 on Muscle Energetics, the energy of that chemical bond is transformed to mechanical work (force x distance) or as heat [Hil38].

Analysis of Patient/Client Problems

Energy is applied to the analysis of patient/client problems since life is dependent on the constant expenditure of energy. Therefore, requires ongoing acquisition, transformation, and transportation of energy. If muscle fibers lack the ability to cause movements that enable the acquisition of energy (consumption of nutrients), then muscle subsequently deteriorates (atrophy, loss of muscle mass) due to the organisms continual need to acquire energy. Muscle fibers are sacrificed in situations of malnutrition or starvation, even when that malnutrition or starvation is caused by the inability to move.

⁶ All the genes together

⁷ Above the genome, the meta genome

1.2.5 Mass Balance

The quantity of "stuff" in any system, or in a compartment of a system, is determined by the inputs into the system and the outputs from that system or compartment. This concept is based on the law of conservation of mass. When someone consumes 50 grams of carbohydrate (CHO) those 50 grams are then part of them and will be processed (digested, metabolized and either eliminated, stored or transformed).

Analysis of Patient/Client Problems

Mass Balance is applied to the analysis of patient/client problems any time we consider the amount of some physiological variable. This is very common in practice. Blood volume of all of the components that make up blood volume must have the correct mass balance for the ECF to have the correct mass balance for the muscle fibers to have the correct mass balance. When blood work is performed it is a general check on whether mass balance is being maintained on a whole host of components of blood volume.

1.2.6 Flow Down Gradients

The transport of "stuff" (ions, molecules, nutrients, blood, and gas) is a central process at all levels of organization in the organism. A Ohm's Law is a simple model that describes such transport.

Ohm's Law describes the relationship between current, voltage and resistance:

$$Current = \frac{Voltage}{Resistance} \quad (1.1)$$

For fluid transport, such as blood through vessels and air through the airways we consider:

$$Flow = \frac{Pressure}{Resistance} \quad (1.2)$$

For diffusion across membranes we consider:

$$Diffusion = \frac{Concentration}{Resistance} \quad (1.3)$$

Important Concept

Current, flow and diffusion are proportional to the voltage, pressure or concentration gradient (numerator), and inversely proportional to resistance (denominator). You must understand these questions before proceeding. Ohm's Law can be expanded for additional context and explanation. For example, considering the factors that influence resistance of blood or air flow (expanding the denominator) provides additional context and explanation for the regulation of cardiac output. Ohm's Law is an often utilized equation throughout the book. You should practice manipulating it. Above it is solved for current. Manipulate it to solve for voltage. Then manipulate it to solve for resistance.

Tip

Equations and laws are an extremely helpful way to consolidate and summarize your understanding about concepts. Concepts have universal applicability. When you are reasoning in practice you are situated at one time, in one place, with one situation (a particular situation). The more universal concepts you can relate to a particular situation the more you understand that situation (even when you have never before been in that situation). It is recommended that write down the equations (separate from the book) so you can refer back to it easily while reading.

Analysis of Patient/Client Problems

Flow Down Gradients is applied to the analysis of patient/client problems any time we consider pressure (blood pressure, oxygen partial pressure, etc) or concentrations (sodium concentration in the ECF). Disturbance in these pressures or concentrations fundamentally disturb the flow that is necessary for normal physiological function.

1.2.7 Homeostasis

The internal environment of an organism is actively maintained "relatively constant" by the feedback function of cells, tissues, and organs. This is organized into primarily negative feedback systems.

Important points about homeostasis:

- Not everything is regulated (i.e. heart rate)
- Homeostasis is not an on/off switch
- "Relatively constant" – meaning there can be acceptable variation
- Set points can change
- There is a hierarchy of homeostatic regulation

The hierarchy of homeostatic regulation includes the fact that cells manage their internal environment (homeostasis), which includes taking and giving molecules and water to the ECF, the ECF is kept in homeostasis via exchange with the blood volume (intra-vascular ECF) and homeostasis of the blood volume (water, cells, electrolytes, nutrients, etc) is exchanging with organ systems (lungs, liver, pancreas, gut). A simple graphic model of glucose homeostasis is shown in Figure 1.2 and shows the interactive relationship between glucose, glucagon and insulin.

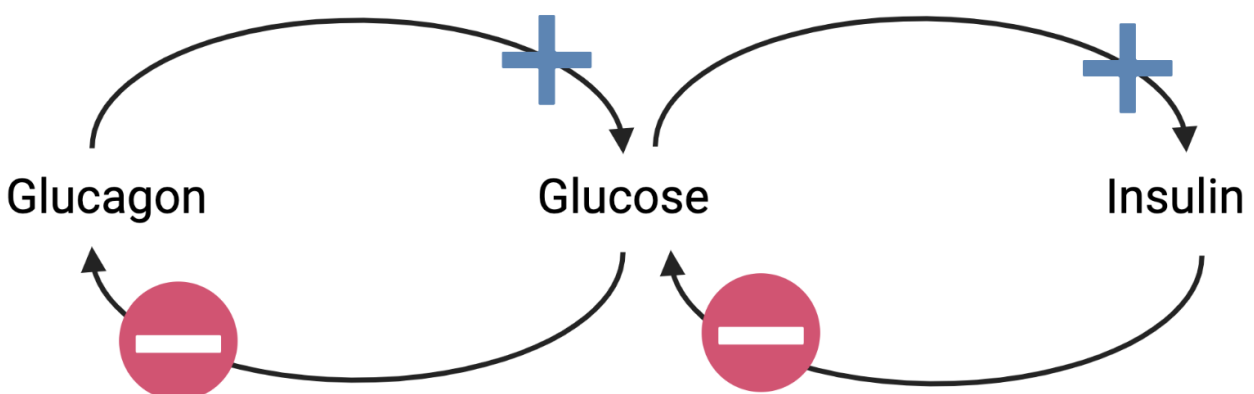


Fig. 1.2: Simple Model of Glucose Homeostasis that does not consider the organs or cellular mechanisms underlying the homeostatic relationship between the three components. In this system, glucose is being regulated. (Created with BioRender.com)

Figure 1.3 below provides a more complete model of glucose regulation. It includes multiple mechanisms that glucose can be regulated and the role of metabolic organs such as the liver and the endocrine pancreas.

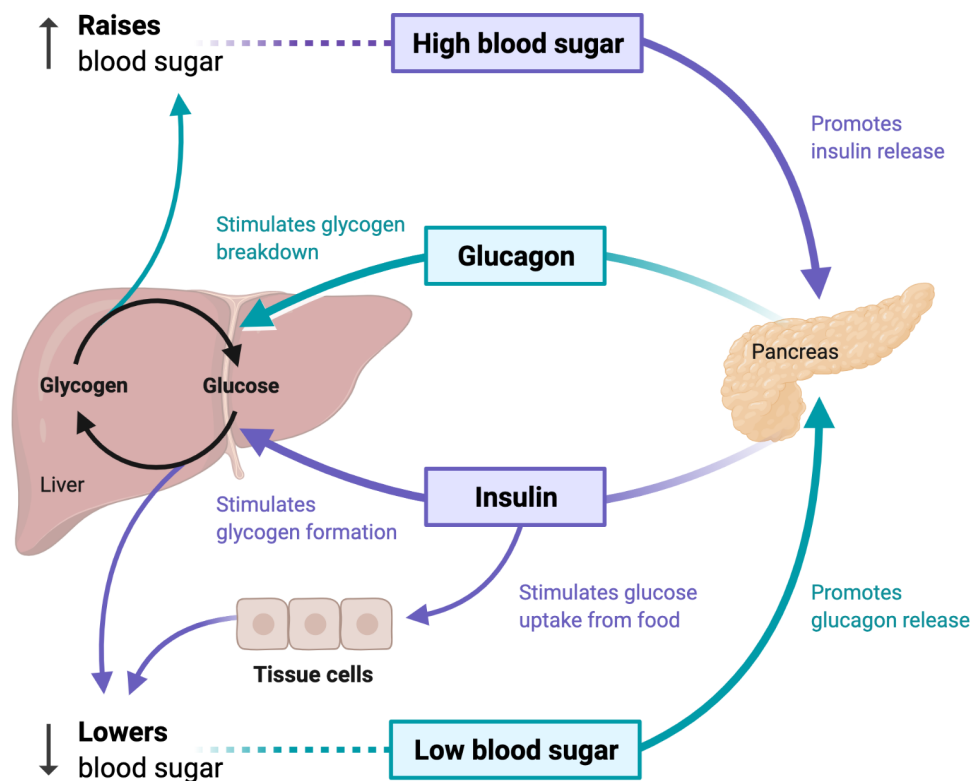


Fig. 1.3: Model of Glucose Homeostasis that includes the organs involved in this homeostatic system (Created with BioRender.com)

Analysis of Patient/Client Problems

Homeostasis is routinely applied to the analysis of patient/client problems. Many diseases, conditions and syndromes, regardless of their underlying cause, results in some sort of disturbance to homeostasis; and thus any disturbance in homeostasis is ultimately analyzed for the possible causes.

1.2.8 Interdependence

Cells, tissues, organs, organ systems interact with one another (are dependent on the function of one another) to sustain life. Interdependence includes hierarchies of homeostatic regulation. The function and capacity of the whole system is related to the interdependent relationship between the parts, capabilities and attributes of each of the multiple systems.

Analysis of Patient/Client Problems

Interdependence is applied to the analysis of patient/client problems during the process of differential diagnosis. This process includes asking the question - "What are all the things that could cause this problem?" The list of possible causes can grow quickly given the interdependence of the whole system (the whole person).

1.2.9 Structure - Function

The function of a cell, tissue, or organ is determined by its form. Structure and function (from the molecular level to the organ system level) are intrinsically related to each other. Structure influences function (right now) and function influences structure (eventually). Part II is how muscle structure influences function right now. Part IV includes how function influences structure eventually (hypertrophy and atrophy).

Analysis of Patient/Client Problems

Structure - Function is applied in the process of diagnosis of pathoanatomical conditions and also is the foundation for pathokinesiological interventions. If function is impaired, we consider the structure that may result in such an impaired function. For example, if there is an deviation in the gait pattern (function), we may consider the length of the hip flexors (structure impacts function). If there is a functional movement deviation, prolonged sitting (function), we may teach corrective exercises to prevent future changes in structure (shortened hip flexors) that become dysfunctions (function impacts structure eventually).

1.2.10 Hierarchy of Adaptation

A hierarchy of adaptation and adaptability works within and around physiological systems. Adaptation refers to changes that occur to the system so that it optimizes function. A muscle fiber creates tension. Adaptation refers to changes that occur so a muscle fiber can attain, sustain and maintain the ability to create tension. Adaptability refers to the ability to adapt. The hierarchy includes genetic, epigenetic, anatomic, physiologic, behavioral and cultural changes. Cultural changes being a mechanism of passing behavioral adaptation from one generation to another. This entire hierarchy is involved through the life span and has relevance to physiological adaptation. Physiological adaptation includes genetic, epigenetic, anatomic, physiologic mechanisms and include changes to structure and function. Behaviors are enabled by our physiology, but then also influence our physiology (our physiology allows us to eat sugar, and then our eating of sugar influences our physiology). And the culture around us gives rise to acceptable, and not acceptable, behaviors.

Analysis of Patient/Client Problems

Physical therapists incorporate the hierarchy of adaptation with the World Health Organization's International Classification of Function (ICF), see Figure 1.4. The ICF is a model that describes how physiological systems contribute (as body systems and functions) and interact with functioning (or dysfunction or impairments). The ICF model focuses on the bidirectional interactions between bodily functions (and structures), activities (movements and behaviors) and participation (behaviors and interactions). It also considers to health conditions that impact functions, activities and participation along with external and personal factors.

The ICF is useful for understanding physiological interdependence, structure - function relationships, and the hierarchy of adaptation. The middle of the ICF is *Activity*, which is primarily referring to movement, making the ICF a movement centered approach. The model considers what is necessary for movement, body structure and functions; and why we do movements, to participate. In this way the ICF overlaps with our muscle centered approach as depicted in Figure 1.5

1.3 Summary

A muscle centered approach to clinical physiology for a physical therapist is based on the relationship between muscle fibers, movement and extra-cellular fluid. Muscle fibers are necessary, but not sufficient, causes of movement. Movement is a multi-system physiological function and contributes to the ECF (Figure 1.1). Therefore, movement is critical to the health and well-being of the body since it is critical to the ECF, and the ECF is critical to all cells. The muscle centered approach is a *middle out* [Nob08] approach to understanding the complexity and hierarchy of physiology for the physical

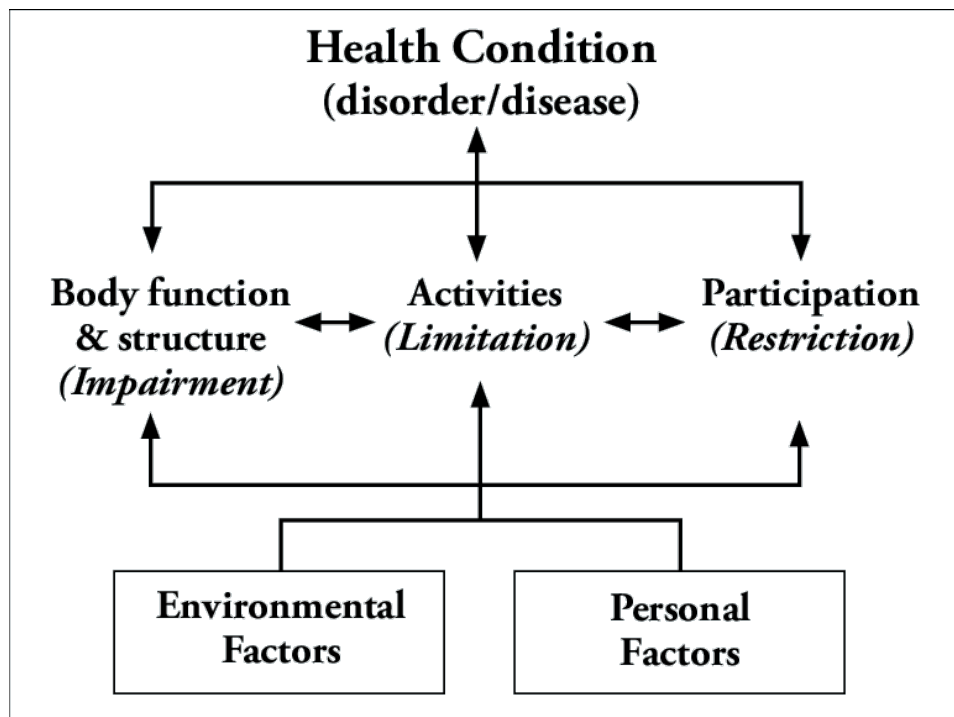


Fig. 1.4: International Classification of Function (ICF) (Created with BioRender.com)

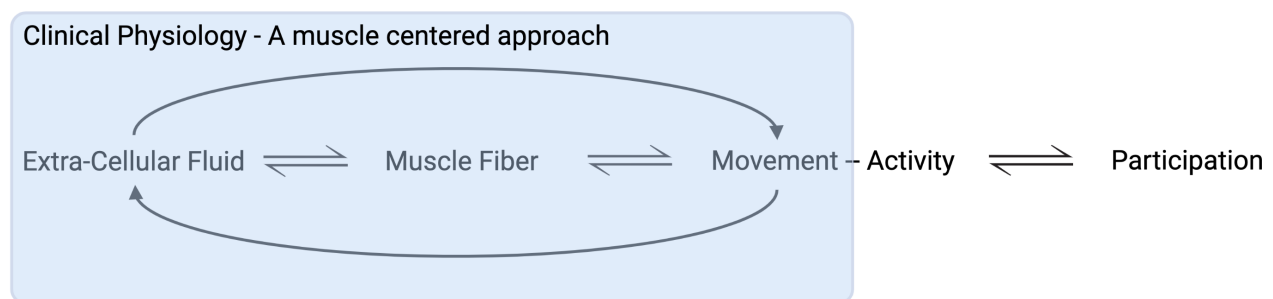


Fig. 1.5: Muscle centered approach overlap with the ICF (Created with BioRender.com)

therapist as a movement specialist. By *middle out* we refer to the entire approach as a model of clinical physiology that looks upward in scale (movement as a behavior within a culture) and looking downward in scale (to the genes).

Next Steps

In Chapter 2 on Fundamentals we review muscle *in situ* including the muscle scaffold including connective tissue, tendons and bony attachments. We then consider movement concepts such as attaining and sustaining tension, active and passive tension, and muscle organization and arrangement. It is important to have a foundation of these fundamentals for understanding the behavior of muscles *in situ* that muscle fibers contribute and the bidirectional implications of the muscle - muscle fiber relationship in causing movement.

References

- [CD15] Sean Collins and Konrad Dias. *Heart Failure: Stability, Decompensation, and Readmission*. en. Cardiopulmonary Physical Therapy Journal 26.3 (Sept. 2015), pp. 58–72. ISSN: 1541-7891. DOI: [10.1097/CPT.0000000000000111](https://doi.org/10.1097/CPT.0000000000000111). URL: <http://Insights.ovid.com/crossref?an=01823246-201509000-00003> (visited on 01/17/2020).
- [Col18] Sean M Collins. *Synthesis: Causal Models, Causal Knowledge*. en. Cardiopulmonary Physical Therapy Journal 29.3 (2018), pp. 134–143.
- [Hil38] A.V. Hill. *The heat of shortening and the dynamic constants of muscle*. en (1938), p. 60.
- [Nob08] Denis Noble. *The music of life: biology beyond genes*. Oxford University Press, 2008.
- [WCH85] Roger C Woledge, Nancy A Curtin, and Earl Homsher. *Energetic aspects of muscle contraction*. Monographs of the physiological society 41 (1985), pp. 1–357.

Sample Questions

1. The ability to create tension, be excited, be regulated and convert chemical to mechanical energy are fundamentally characteristics of?
2. A person that finds it difficult to get out of their chair and walk around their home neglects the activities of daily living and is eventually found to be dehydrated and malnourished. This is an example of a problem with?
3. A high rate of cell division is a benefit for healing. This is an example of which core physiology concept?
4. Cardiac muscle cells are attached by gap junctions that allow them to send excitation signals from one cell to the next for a faster excitation of the entire heart. This helps make sure cardiac muscles coordinate the generation of tension to be an effective pump. This is an example of which core physiology concept?
5. Strength training increases the release of growth hormone which ultimately contributes to increased muscle cross sectional area (hypertrophy). This is an example of which core physiology concept?

Chapter 2

Fundamentals

Updated on May 22, 2024

Contents

2.1	Muscle Fibers and Muscles	16
2.2	Attaining & Sustaining Tension	16
2.3	Connective Tissue Scaffold	17
2.4	Muscle Mechanics	18
2.5	Muscle Roles	23
2.6	Summary	24

Objectives:

1. Explain what is meant by muscle *in situ*.
2. Describe and explain the importance of the macroscopic muscle scaffold including connective tissue, tendons and bony attachments.
3. Explain what is meant by attaining and sustaining tension.
4. Explain the difference between active and passive tension.
5. Describe and explain the implications of muscle pennation.
6. Explain the contexts that result in Δ length during active tension (concentric, eccentric, isometric).
7. Explain roles of muscles such as agonist, antagonist, synergists and stabilizers.
8. Explain the relationship between of muscle active tension, muscle force, muscle torque and movement.
9. Provide of an example of a free body diagram including the relevant forces to generate torques and which relationships between the torques produce different activation types and movements.

Introduction

This chapter covers fundamentals of muscle structure and function. It defines terms and explains concepts used through the book. It is expected that anyone reading this book has a background that includes Anatomy & Physiology. Skim or even skip sections that you have already mastered. The book attempts to balance the depth and breadth of clinical physiology with a muscle centered approach for physical therapy. This means the book is not the deepest nor the broadest coverage of each topic. There are areas that you have gone deeper, but we may take a broader perspective. There are areas where you are comfortable with the breadth, but we may take a deeper perspective. Physical therapists balance depth and breadth (look down, look up and beyond) during examination, evaluation and interventions. The PT may consider the molecular effect of calcium for muscle activation, its role within the supportive matrix of bone, its absorption and distribution through the body, and whether someone is physically capable of independently obtaining and preparing meals that provide sufficient calcium in their diet. The terms - deeper and broader - are relative.

2.1 Muscle Fibers and Muscles

Through the book we take a muscle centered approach. Most of the time this means a muscle fiber centered approach. The emphasis is on the cellular events occurring in the muscle fiber to attain and sustain tension as needed for movement, and the need for muscle fibers to be supported by complex integrated physiological function. While muscle fibers are the basic unit of the muscle system, each muscle fiber constitutes one cell of an entire muscle. It is the muscle as a structural and functional entity that physical therapists think about more often. However, it is also important to recognize, investigate and relate what is happening in that muscle to what is happening in its fibers (cells), and how the body is supporting those fibers. The muscle includes muscle fibers and connective tissue. A single muscle fiber has a limited functional role in the moving human. It is the collective action of muscle fibers, as a muscle, that muscle fibers play an integral functional role in the moving human. To be organized as an effective collective, muscle fibers are bound together by connective tissue, and through connective tissue are attached to bones or other anatomical structures.

Muscles *in situ*

In situ is an adverb or adjective meaning "in original position." The muscles that you know (biceps femoris, deltoid, gastrocnemius, extensor digitorum longus, etc) are classified and identified (named) by shape and location. Shape is determined by how the muscle fibers are connected together with connective tissue. Variations in shape occur primarily from variations in how the muscle fibers are connected together and arranged. Similar to how you can construct many different shapes with a set of similarly shaped bricks. The location of a muscle is determined by its tendon attachments. The muscles you learn in anatomy and consider as part of a physical exam when palpating and testing are muscles that are identified in their original shape and in their original location. This is the muscle *in situ*.

Muscle fibers create tension. Creating tension with muscle *in situ* occurs in the context of the muscle fibers connected together, and by interacting with other structures through tendon connections. This book centers on the muscle fiber. However, we cannot lose contact with muscle *in situ* so we need to know how muscle fibers come together as muscles, and how they transmit tension to bones, and how the muscle *in situ* impacts the muscle fiber.

2.2 Attaining & Sustaining Tension

The muscle fiber creates tension and collectively muscle fibers transmit this tension to muscles *in situ*. Muscles can be evaluated based on their ability to, and how well, the muscle attains tension. How much tension can be attained is often measured as force. There must be internal parts and mechanisms that allow muscle fibers and muscles to attain tension. Muscles also sustain tension long enough for its intended use. To sustain tension means keeping a certain amount of tension in the muscle. Sustaining tension includes attaining tension and then integrating over time. It is measured as the time an attained amount of tension is sustained. Muscle fibers must attain to sustain tension.

In upcoming chapters we consider differentiation of muscle fibers. Some muscle fibers are optimized to attain tension at the expense of sustaining tension, and others are optimized to sustain tension at the expense of attaining tension. "At the expense of attaining tension" refers to the amount of force that could be measured from the attained tension. This may not be the only metric by which to quantify or classify attaining tension. For example, the rate of attaining tension may be more important than the maximal quantity of tension in certain contexts.

2.2.1 Tension

Muscle fibers and muscles create tension. Tension creates a pulling force between two objects. For a muscle fiber *in situ* those objects are connective tissues connected to other muscle fibers serially and in parallel. For a muscle *in situ* those objects are bones, with tendons forming a transitional bridge between muscle fibers and bone. An important feature of the physical concept of tension is that it is related to a pulling on two objects. When pulling a rope between two objects (there is a force involved) the objects continue to move closer to one another (to approximate each other). The rope is using and has tension. If pushing an object towards another object that is connected by a rope the rope may remain in

place but it will be slack and may start to compress and limit further approximation of the objects. But that is not tension, it is compression. A muscle fiber creates tension, not compression. There may be particular muscles that compress at certain joint ranges of motion (ROM). For example, passive elbow or knee flexion can be said to have a "soft end feel" because the slackened flexor muscles are compressed which limit further flexion. A soft end feel occurs when soft tissue compression limits further ROM. But this is also not the action of the muscle fibers of the biceps they are merely the inactive rope going from being slack to be compressed.

2.3 Connective Tissue Scaffold

Building a muscle from muscle fibers requires an intricate connective tissue scaffold that connects and shapes muscle fibers. This scaffold functions to transmit tension between muscle fibers and tendons. This connective tissue binds muscle fibers into what we see and define as a muscle. Three connective tissue compartments connect and organize muscle fibers so that the tension created by and transmitted to the muscle fiber interacts as required with the muscle *in situ*. Figure 2.1 shows the hierarchical relationship the endomysium, perimysium and epimysium. The **endomysium** is thin layer of connective tissue that surrounds each muscle fiber. The endomysium should not be confused with the sarcolemma which is a cell membrane and also surrounds each muscle fiber. The **perimysium** is a thicker layer of connective tissue that surrounds groups of muscle fibers and forms a **fascicle**. The muscle fibers in a fascicle are arranged in parallel (and thus influences a muscle's cross sectional area (thickness)). The arrangement of multiple fascicles influences both the muscle length (if fascicles are arranged in series) and pennation (if fascicles are arranged at an angle to the overall muscle's line of tension through its tendons. The **epimysium** surrounds all the fascicles to form a muscle belly (readily observed on visual observation). A muscle belly may or may not be distinguished from a muscle. This depends on context. For example, the deltoid muscle is known to have separate parts (anterior, medial and posterior), each of these parts has it's own epimysium and is thus a separate belly. The point is that the epimysium does not necessarily surround the muscle as you know it and identify it (i.e. if you identified the deltoid you'd be identifying three separate muscle bellies that are each encased in an epimysium connective tissue wrap). There are additional connective tissue wraps that do surround muscles as you know them (superficial to the epimysium), and there are additional connective tissues that surround ever more superficial layers that result in muscles being bound to one another (i.e. fascia and fascial trains), but these levels of connective tissue encasement are not covered in this book. A reminder of the fascia layer is easy to experience with the difference you feel between a hamstring stretch with ankle dorsiflexion vs. plantar flexion.

2.3.1 Roles of the Connective Tissue Scaffold

The endomysium, perimysium, epimysium connective tissue scaffold connect, contain and arrange muscle fibers. There are two other notable functional roles. First, they transmit tension and to and through tendons to attachments [TMS13]. Second, they secure the physical location and approximation of the neurovascular supply to the muscle fibers.

2.3.1.1 Transmit Tension

Each muscle fiber transmits tension to one another and through the connective tissue scaffold. To transmit tension the connective tissue must securely, and with minimal elasticity, connect each muscle fiber to each other and to the tendon. The endomysium and perimysium tend to be thinner and contain both collagen and reticular fibers. Whereas the epimysium tends to have a greater volume of collagen (less elastic).

The transmission of tension from muscle fiber to tendon to bone includes two zones of transition: muscle-to-tendon, and tendon-to-bone. The muscle-to-tendon zone is referred to as the **musculotendinous junction (MTJ)**. The MTJ is where muscle fibers start to diminish and muscle connective tissue, primarily epimysium starts to increase and meet the collagenous tendon tissue. From the muscle fibers toward the tendon the MTJ includes multi directional tendon projections gradually reorienting to being longitudinally oriented tendon tissue which allows a uniform transmission of tension. The multidirectional projections combine the muscle fiber tension to the tendon [Knu+15]. Tendons consist of dense connective tissue that allows them to absorb and transmit substantial tension. Tenocytes (cells that make tendons)

Structure of a Skeletal Muscle

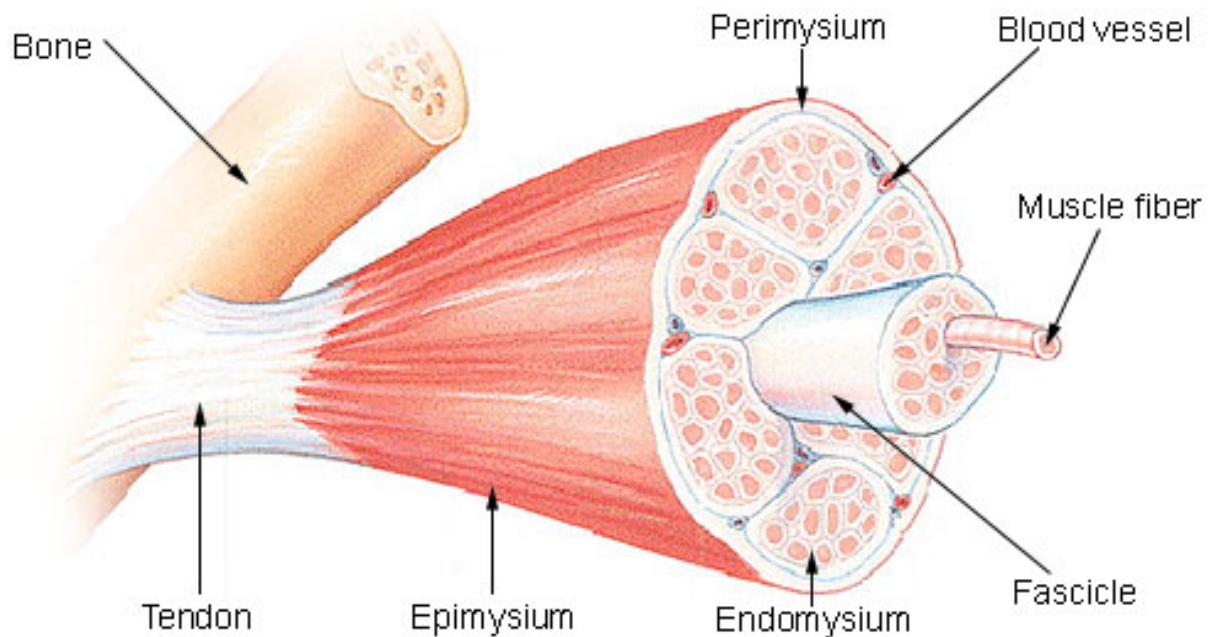


Fig. 2.1: Connective Tissue Scaffold of Muscle (Public Domain figure from [Wikimedia Commons](#))

account for about 20% of tendon volume, and the extracellular matrix (ECM) accounts for about 80% of tendon volume [Kjæ04].

The tendon-to-bone zone is referred to as the tendon-bone attachment. The formation of an attachment between tendon and bone starts during embryonic development. During musculoskeletal system assembly, the tendon–bone attachment forms a structure that is mechanically complex given the need to transfer tension between two materials that differ greatly in stiffness (tendon being non mineralized connective tissue, and bone being mineralized connective tissue).

2.3.1.2 Secure the Neurovascular Supply

The connective tissue scaffold secures the neurovascular supply to muscle to make sure that the neurovascular supply stays where it is supposed to stay. Entering through epimysium and perimysium connective tissue the neurovascular structures (nerves, arteries, veins) are secure in their path regardless of muscle length. They are then embedded in the endomysium in close proximity to the muscle fiber regardless of the length of the muscle. The vascular supply is delivered by arteries and arterioles, and removed by veins and venules. They provide blood flow that penetrates the epimysium and perimysium to reach the muscle fibers. Capillaries then surround each muscle fiber embedded and anchored in the endomysium to interact and exchange mass with the extra-cellular fluid surrounding the sarcolemma. Innervation of muscle for both excitation (Chapter 4) and regulation (Chapter 5) by the nervous system comes from branches of peripheral nerves that penetrate the epimysium and perimysium and terminate in and are anchored by the endomysium at the required location for each muscle fiber (at the neuromuscular end plate).

2.4 Muscle Mechanics

Chapter 2 on Tension focuses on the mechanisms internal to a muscle fiber for generating tension. Our inclusion of muscle fiber mechanics here is a first pass on basic topics of muscle mechanics such as passive and active tension, pennation, and the velocity of changes in length (Δ length).

2.4.1 Active & Passive Tension

Tension created by a muscle fiber is active or passive. Active tension is created when the muscle fiber is activated via excitation of its membrane which triggers processes that transform chemical into mechanical energy. Activation as a word captures the essence of active tension and is preferred to the more commonly utilized term, contraction. The mechanical energy of activation pulls the ends of a muscle fiber together. It attempts to shorten the muscle fiber and through transmission of that tension to the connective tissue scaffold it attempts to shorten the entire muscle (bring the tendon attachments together). Passive tension is generated when an "other force" lengthens the muscle and its fibers. This "other force" is always external to the muscle fiber (or muscle) itself. The other force can be another muscle, or it can be external to the musculoskeletal system completely, such as gravity. Passive tension occurs when an external force attempts to lengthen the muscle fiber. For a muscle *in situ* this is commonly when an external force pulls the tendon attachments apart, and in response the muscle and its fibers create tension that resists lengthening **without** activation.

2.4.2 Pennation - Fascicle Arrangement

Fascicles can be organized (mostly) parallel to the overall tendon line of pull, or with an angle between the fascicle and the overall tendon line of pull. Figure 2.2 shows an arbitrary muscle with a line of pull of the whole muscle along the overall tendon line in red (force of the whole muscle (F_{wm})); and the line of pull of the muscle fibers bundled in fascicles at angle α in orange (force of the fiber (F_{fiber})). When there is an angle, such as α , the muscle is pennated. Pennation exists when there is an angle between the line of pull of the fascicles and the line of pull of the tendons. In Figure 2.2 the line of pull of the tendon is the vector sum of the lines of pull of the fascicles. With pennation there is a vector component of the muscle fiber that acts in line with the tendon pull, and there is a vector component that does not work in line with the tendon pull. The vectors that act in line with the tendon pull sum together; whereas the vectors that do not act in line of the tendon pull cancel each other out. Fascicle arrangement is one of the primary features that distinguishes various classifications of muscles.

Fascicle arrangement is a muscle *in situ* factor that influences the velocity of shortening of the muscle per unit of shortening of a muscle fiber; and the force of activation. For the most part fascicle arrangement is not a modifiable aspect of a muscle. There are no training (stretching, pulling, needling, strengthening) programs that alter a particular muscle's pennation status (such as whether it is pennated or not). However, there is evidence of particular types of training having a small influence on the angle of pennation [Cut+20].

2.4.2.1 Velocity of Shortening

If two muscles, one pennated and one not pennated, had a set of homogeneous muscle fibers that all shortened at the same velocity during activation with no resistance; the muscle without pennation would shorten with a higher velocity. This occurs because the shortening of each fiber is summed to the whole muscle. If each fiber shortens to 50% of its length in one second, then the muscle as a whole would shorten to 50% of its length in one second. However, in a pennated muscle, when each fiber shortens to 50% of its length in one second, the muscle as a whole shortens at about $50\% \times \cos(\alpha)$ (recall that α = angle of pennation). For example, if $\alpha = 35^\circ$ then with all muscle fibers shortening 50% in 1 second there would be about 40% in 1 second.

2.4.2.2 Force of Active Tension

The amount of force that a muscle can develop during maximal activation is proportional to the cross sectional area (CSA) of the muscle. The force of the whole muscle is proportional to the number of parallel muscle fibers that are activated. Note the stipulation that these fibers must be parallel. Adding muscle fibers in series (which make a muscle longer but not thicker) does not contribute to the force generated by the entire muscle (at its ends). In series each muscle fiber pulls from both ends during activation and some of the force is cancelled out. In Figure 2.3 there are three fibers in series (A) and in three fibers in parallel (B). In series the forces from 1 and 2 cancel each other out and the force of 3 is

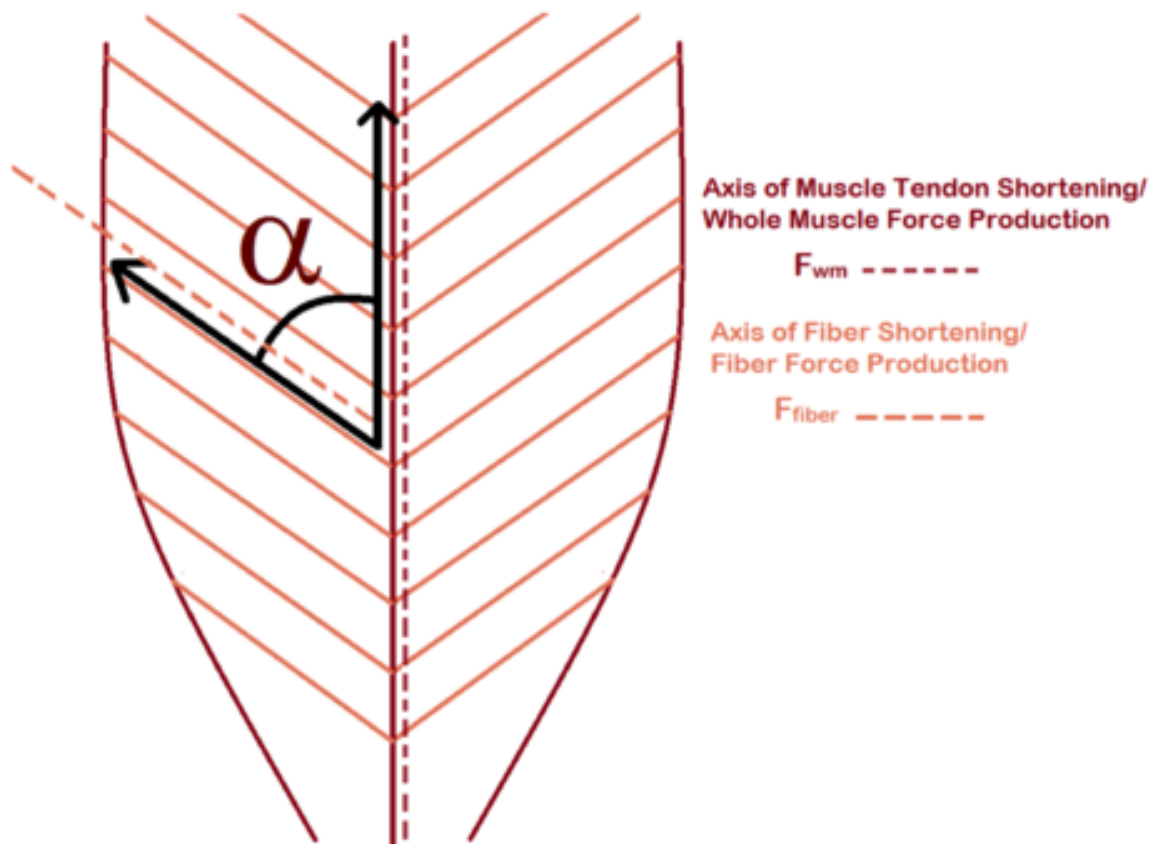


Fig. 2.2: Pennation of Muscle Fibers (Public Domain figure from [Wikimedia Commons](#))

transmitted to the the ends. When in parallel the forces of the three sum. Note that the muscle fibers in parallel increase the CSA.

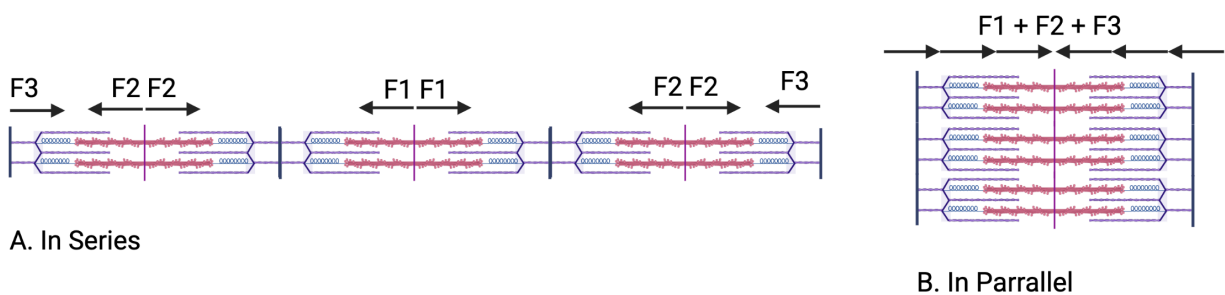


Fig. 2.3: Muscle Fibers in Series & Parallel (Created with Biorender.com)

Pennation tends to increase the CSA (the number of muscle fibers in parallel). Similar to the loss of velocity in a pennated muscles, not all of the muscle fiber tension is transmitted to the line of pull of the tendon and thus is lost. However, the

increase in CSA compensates for loss in force associated with the vector sum. For example, if a muscle fiber can generate 1 unit of force and we have 50 fibers in parallel the muscle as a whole can generate 50 units of force. If we can get 100 fibers in parallel with a pennated arrangement of those fibers with an angle of pull that transmits 70% of the force to the line of pull of the tendon then the muscle can generate 70 units of force. Despite a slight loss of efficiency at the level of the muscle fiber (less tension transmits to the line of pull of the tendon), there is a gain in tension at the level of the muscle. Muscles that are not pennated (fusiform or strap muscles) tend to have higher velocities of shortening during unresisted active tension and lower maximal forces during resisted active tension.

2.4.2.3 Overall Effect of Pennation

Pennation tends to be a decrease in velocity of shortening during unresisted activation and an increase in maximal force during resisted activation. These variations (pennation or not pennation) do not occur due to muscle adaptations to training or other modalities. There is some evidence that angles of pennation can be modified slightly, and there is evidence that all muscles can change their cross sectional area regardless of their pennation status.

2.4.3 Δ Length During Activation

Δ length is a change in length of a muscle. By convention, negative Δ length indicates a muscle shortens, positive Δ length indicates a muscle is lengthening and zero Δ length indicates a muscle does not change length. Activation types are defined based on the Δ length during activation.¹

Δ Length	Muscle	Activation Type
Negative	Shortens	Concentric
Zero	Does not change	Isometric
Positive	Lengthens	Eccentric

Table 2.1: Classification of Δ Length During Activation

Anything other than a concentric activation requires counteracting forces. A muscle not connected to anything and not under the influence of other forces will shorten (negative Δ length) with activation. To have activation of a muscle with 0 Δ length (isometric) or positive Δ length (eccentric) requires other forces such as other muscles or external forces acting on the joint(s). Forces acting on joints create joint movement based on the torque being created. Therefore, whether a muscle has a concentric, isometric or eccentric activation depends on the balance between torque it is creating (muscular or internal torque) and the torque being created by all other forces acting at the joint. If the forces are coming from outside of the body then they are called external torque.

The free body diagram in Figure 2.4 shows the biceps muscle acting as an elbow flexor in two situations. In 2.4-A the internal force of the biceps muscle is creating **MORE** flexion torque than the external force is creating extension torque. Therefore the elbow flexes. Since the biceps muscle fibers are activated with a negative δ length it is a concentric activation. Since it is creating flexion, it can be called concentric flexion. In 2.4-B the internal force of the biceps muscle is creating **LESS** flexion torque than the external force is creating extension torque. Therefore the elbow extends. Since the biceps muscle fibers are activated with a positive δ length it is an eccentric activation. The external force is creating extension and the biceps are resisting extension, it can be called eccentric extension.

The free body diagram in Figure 2.5 shows the forces and related torque during the nordic hamstring exercise (NHE). During this exercise the subject is on their knees with their ankles supported so that the lower limb (below the knee) cannot move. The subject keeps their hips straight and slowly shifts their body mass forward. Once the body mass (black arrow) is anterior to the axis of rotation of the knee it creates a knee extension torque (black circular arrow). The subject

¹ You may have learned these as contraction types. You should be comfortable exchanging the terminology muscle contraction for muscle activation and vice versa. However, activation is the preferred language for this course and avoids some of the confusion with eccentric activation.

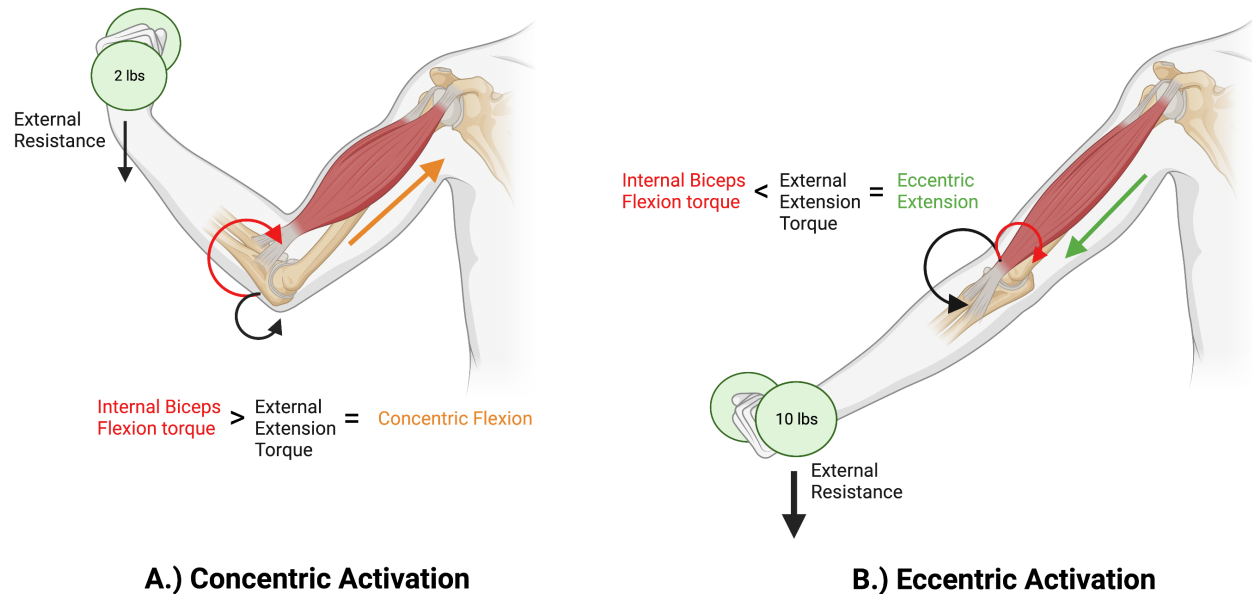


Fig. 2.4: Biceps Muscle Concentric Flexion vs. Eccentric Extension
(Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

allows themselves to not fall forward to quickly but usually reaches a point where they can no longer slow their descent and they fall to the floor. While they are attempting to slow their descent to the floor they are using their hamstring muscles which create a force through active tension (red arrow), which then creates a knee flexion torque (red circular arrow). The exercise can be done a variety of ways. In one scenario the subject can stop and hold themselves in one position. In another scenario, after holding themselves they can use the force of hamstring active tension to bring them upright again.

There is a relationship between the force of the body mass (F_{bm}) and the knee extension torque created by the body mass (T_{bm}) which is external since it is the result of gravity. There is a relationship between the force of the muscle (F_m) and the knee flexion torque created by the muscle (T_m), which is an internal or muscular force. The muscular force varies based on how much tension is created by the muscle activation. Movement depends on the relationship between the two opposing torques as seen in Table 2.2. The movement and the activation type are dependent on the relationship of the torques. The torques are related to both the force generated by the tension of the muscle, the body weight and their respective perpendicular distances.²

Torque Relationship	Δ Length	Muscle	activation Type	Movement
$T_m > T_{bm}$	Negative	Shortens	Concentric	Knee Flexion
$T_m = T_{bm}$	Zero	Does not change	Isometric	No Movement
$T_m < T_{bm}$	Positive	Lengthens	Eccentric	Knee Extension

Table 2.2: Δ Length During Active Tension Based on Torque Relationships

? Questions

² This book will not spend much time at this level of analysis. However, if you are not comfortable with torque and vectors from physics and other courses you have had it is a worthwhile to refresh your memory now.

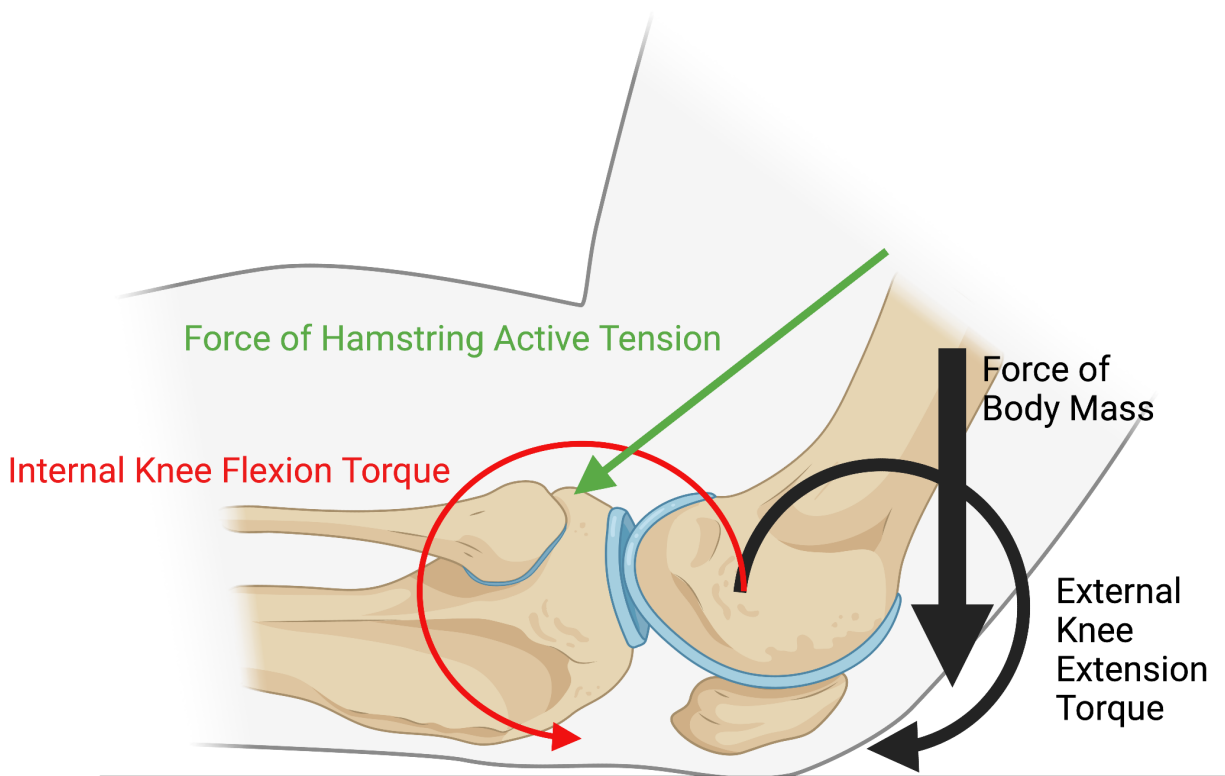


Fig. 2.5: Free Body Diagram of the Nordic Hamstring Exercise
(Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

If an external flexion torque is less than an internal (muscle tension generated) extension torque, what would be the movement and the type of muscle activation?

If an external extension torque is greater than an internal (muscle tension generated) flexion torque, what would be the movement and the type of muscle activation?

2.5 Muscle Roles

- Agonist: Muscle creating tension for the intention of the movement.
- Antagonist: Muscle that creates the opposite movement of the agonist.
- Synergists: Muscles that work together and combine their tension to create a movement. We can talk about synergist agonists and synergist antagonists, though the former is typically the intention when just referring to synergists.
- Stabilizer: Muscles that provide tension to neutralize certain motions toward the purpose of a movement

To define a muscle role we must consider context. Context for muscle roles are typically movements. Movements include all the motions that are occurring, the directions of those motions and the intention of the mover doing the movement. If the movement is the downward phase of a squat one of the motions is knee flexion. The intention of the mover is knee flexion (they intend to move downward). The knee extensor muscles are involved as agonists even though knee is flexing. Just because the motion is knee flexion, the agonists are not necessarily the knee flexors.

2.6 Summary

This chapter covered some of the fundamentals of muscle function. We discussed how the connective tissue scaffold allows muscle fiber tension to transmit to the muscle tendon and bony attachments for the purpose of whole muscle tension, including what it means to consider a muscle *in situ*, and how that same scaffold provides an anchor for the neuromuscular supply and the overall shape of a muscle, including fascicle arrangement. Fascicle arrangement influences how the muscle *in situ* force and length is related to muscle fiber force and Δ length. Not only a particular muscle's activation, but other forces, have an integral influence the Δ length of a muscle during activation and therefore activation type. Finally we classified four muscle roles, agonist, antagonist, synergist and stabilizer. The next chapter covers the muscle fiber, its structure and the mechanisms involved in attaining and sustaining both active and passive tension.

References

- [Cut+20] Matthew Cuthbert et al. *The Effect of Nordic Hamstring Exercise Intervention Volume on Eccentric Strength and Muscle Architecture Adaptations: A Systematic Review and Meta-analyses*. en. Sports Medicine 50.1 (Jan. 2020), pp. 83–99. issn: 0112-1642, 1179-2035. doi: [10.1007/s40279-019-01178-7](https://doi.org/10.1007/s40279-019-01178-7). url: <http://link.springer.com/10.1007/s40279-019-01178-7> (visited on 04/19/2022).
- [Kjæ04] Michael Kjær. *Role of Extracellular Matrix in Adaptation of Tendon and Skeletal Muscle to Mechanical Loading*. en. Physiological Reviews 84.2 (Apr. 2004), pp. 649–698. issn: 0031-9333, 1522-1210. doi: [10.1152/physrev.00031.2003](https://doi.org/10.1152/physrev.00031.2003). url: <https://www.physiology.org/doi/10.1152/physrev.00031.2003> (visited on 05/18/2022).
- [Knu+15] A. B. Knudsen et al. *The human myotendinous junction: An ultrastructural and 3D analysis study: The human myotendinous junction*. en. Scandinavian Journal of Medicine & Science in Sports 25.1 (Feb. 2015), e116–e123. issn: 09057188. doi: [10.1111/sms.12221](https://doi.org/10.1111/sms.12221). url: <https://onlinelibrary.wiley.com/doi/10.1111/sms.12221> (visited on 05/18/2022).
- [TMS13] Andrea Turrina, Miguel Antonio Martínez-González, and Carla Stecco. *The muscular force transmission system: role of the intramuscular connective tissue*. eng. Journal of bodywork and movement therapies 17.1 (Jan. 2013). Place: United States, pp. 95–102. issn: 1532-9283 1360-8592. doi: [10.1016/j.jbmt.2012.06.001](https://doi.org/10.1016/j.jbmt.2012.06.001).

Sample Questions

1. A muscle in its original position is:
2. Once tension is attained, how long does a muscle need to sustain tension?
3. A fascicle is surrounded by?
4. Two muscles are maximally activated and shorten in one second. Muscle A shortens 75% of its length. Muscle B shortens to 50% of its length. Which muscle is more likely to have fascicle pennation?
5. Delta length is Zero and the muscle is activated. What type of contraction is the muscle undergoing?
6. If an external flexion torque is greater than an internal (muscle tension generated) extension torque, then what is the joint movement and what type of contraction is occurring?

Part II

Muscle Function Overview

Part II on Muscle Function includes four short chapters directly related to how muscles generate tension. Tension generating mechanisms are the primary focus of Chapter 3 on Tension. To generate active tension a muscle fiber must be excited - Chapter 4 on Excitation introduces the general concepts of an excitable membrane and signaling processes and how these relate to excitation - contraction coupling. Muscle function depends on sequences of excitation, across time and space, that determine the rate and quantity of active muscle tension. Chapter 5 on Regulation describes the basic mechanisms of regulation of muscle active tension and tone by the central and peripheral nervous system. With all of that in place, we are ready to turn our attention in Chapter 6 on Energetics to the transfer of chemical energy as delivered to the muscle fiber as glucose or fat to mechanical energy in the form of tension.

Chapter 3

Tension

Updated on May 22, 2024

Contents

3.1	Anatomy of Tension	27
3.2	Passive Tension	29
3.3	Active Tension	31
3.4	Force & Velocity	37
3.5	Summary	40

This chapter describes how a muscle fiber generates active and passive tension. Passive tension occurs when a muscle fiber resists lengthening and can be referred to as stretch. Active tension is when the muscle fiber is activated. The activated muscle fiber creates tension that shortens (or attempts to shorten) the fiber. The difference between passive and active tension is whether the muscle is converting chemical energy from the form adenosine triphosphate (ATP) to mechanical energy at the crossbridge site between the proteins actin and myosin. When activated the muscle is actively converting chemical energy to mechanical energy at crossbridge sites. During passive tension the muscle is not actively converting chemical energy to mechanical energy at crossbridge sites, but is resisting an increase in length. The muscle cannot generate passive tension to resist shortening. The important point is that the difference is not whether there is positive or negative Δ length. The difference is whether the muscle fiber is converting chemical energy to mechanical energy.

Objectives include:

1. Describe the structures that generate tension.
2. Explain the role of passive tension in muscle function.
3. Describe the sliding filament theory of active tension.
4. Explain the role of active tension in muscle function.
5. Explain the underlying mechanisms and implications of the length-tension and force-velocity relationships.
6. Apply the concept of muscle tension to the analysis of patient/client problems related to the generation of muscle force.

3.1 Anatomy of Tension

The anatomy of tension focuses on the structures that are responsible for generating tension within the muscle fiber. The focus is on the muscle fiber not the muscle *in situ*. Even if we take the muscle *in situ* and consider the passive tension of the connective tissues (endomysium, perimysium, epimysium) and the tendon attachments, most of the tension from a muscle is thought to still emerge from the muscle fiber structures. Muscle fibers are cells, and cells are the basic

unit of life. There are three basic principles of physiology on the topic of cells (cell theory, cell membrane, cell-to-cell communication).

The parts of a muscle fiber that generate tension are located and organized into the sarcomere. It is important to understand the sarcomere within a hierarchical organization of the full muscle tension structure. Below we proceed through this hierarchical organization top down and then bottom up. The reader should be comfortable going in either direction and should pay particular attention to the middle position of muscle fibers, and how there are combinations of parallel and series combinations of structures.

Hierarchical organization of tension structure, proceeding top down:

- Muscles are built by fascicles.
- Fascicles are built by muscle fibers.
- Muscle fibers are built by myofibrils in parallel.
- Myofibrils are built by sarcomeres in series.
- Sarcomeres are built by myofilaments and structural proteins.

Hierarchical organization of tension structure, proceeding bottom up:

- Myofilaments and structural proteins build a sarcomere.
- Sarcomeres in series build a myofibril.
- Myofibrils in parallel build a muscle fiber.
- Muscle fibers in parallel build a fascicle.
- Fascicles in parallel and series build a muscle.

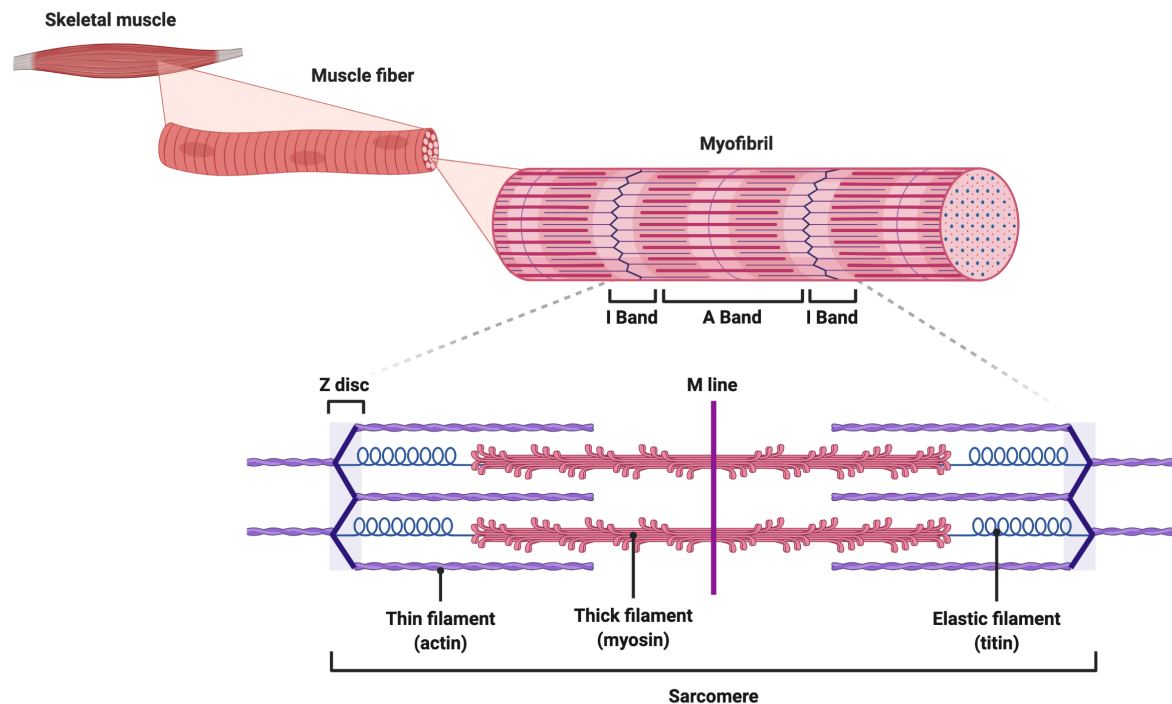


Fig. 3.1: Myofibril Structure (Created with Biorender.com)

Muscle Fibers

In the hierarchical organization depicted in Figure 3.1 muscle fibers include a set of myofibrils arranged in parallel. The boundary of a muscle fiber includes the endomysium and sarcolemma (cell membrane of the muscle fiber). Myofibrils are surrounded by cellular components with critical roles to play when creating tension. Components such as extensions of the sarcolemma terminating at sarcoplasmic reticula are critical for excitation and regulation; the mitochondria are critical for energetics; and nuclei and associated protein manufacturing components such as the Golgi apparatus and ribosomes are critical for muscle maintenance and adaptation. For now we are focused on the muscle fiber components that generate tension. For the process of generating tension, the myofibrils are the functional building block of muscle fibers, and sarcomeres are the functional building blocks of a myofibril.

3.1.1 Anatomy of a Sarcomere

Figure 3.2 provides a view of the tensioning proteins of the sarcomere. The sarcomere includes all of the proteins between two Z-discs. These proteins are all actors that combine to create passive tension (resistance to positive Δ length) and active tension through the cross-bridge cycle (sliding filament model).

Primary and Secondary Roles of Primary Protein "Actors"

- Actin - primary role in active tension; secondary role in passive tension
- Myosin - primary role in active tension; secondary role in passive tension
- Troponin - Tropomyosin - primary role in active tension; secondary role in passive tension
- Titin - primary role in passive tension; secondary role in active tension

Passive tension is created when the Z-discs are pulled apart. Active tension is created to pull the Z-discs closer together. Knowing the precise structure of each of these proteins is not necessary. Just keep in mind that their structure is of great importance. As with all proteins the structure is the foundation of function. Proteins form molecular machines that have very precise functions based on their very precise structure. Anything that alters the structure of a protein alters its function. Variations in pH and temperature that exceed certain boundary conditions, for example, have the capability of changing protein structure. Here we focus on the general structure of the proteins that allow the proteins to play their role in creating tension. The arrangement of these proteins with one another within the sarcomere is important. Everything you need to know about the structure and arrangement is based on the roles they play in the upcoming sections.

3.1.2 Structure and Arrangement of the Sarcomere During Changes in Length

The structural change to the sarcomere at different lengths is the distance between the Z-discs (Figure 3.3). Increasing the distance between Z-discs reduces the space between the actin and myosin and, once at a certain length, starts to pull on the elastic structure of titin (unwinding of titin). Decreasing the distance between Z-discs results in increased space between actin and myosin, and recoil of titin to a point. Further decreases in distance between Z-discs results in crowding of potential actin myosin crossbridges and compression of titin. The implications of these structural changes are discussed in the sections on Active and Passive Tension.

3.2 Passive Tension

Passive tension is created when the Z-discs are pulled apart from one another. During positive Δ length Z-discs of sarcomeres are pulled apart and muscle fibers increase in length. The amount of passive tension developed (force that resists lengthening) is a function of the relative length of the muscle. This is due to the relative length of titin and is the predominant component of the passive component of the length - tension relationship (Figure 3.4).

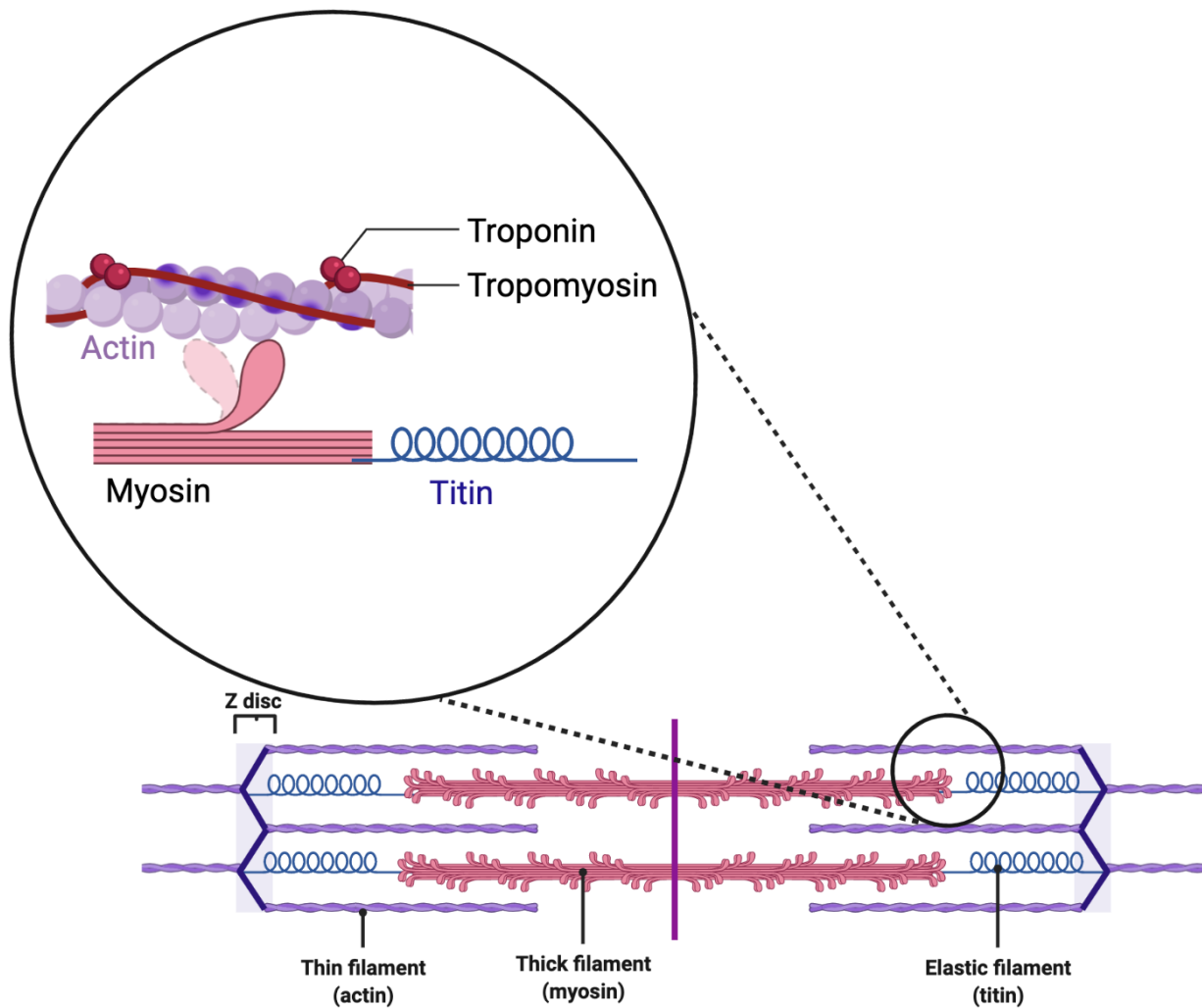


Fig. 3.2: Sarcomere Structure: Primary Protein "Actors" (Created with Biorender.com)

Muscle Tone

In an "inactivated" muscle there are some attachments between actin and myosin. The number of actin and myosin attachments influences the passive tension developed at various lengths. This is referred to as muscle tone. Muscle tone is the amount of tension in a muscle when it is not activated. Muscle tone is dependent on the length of the sarcomere since that length influences the number of possible actin myosin attachments. The presence of these actin myosin attachments in an inactivated muscle is primarily related to the underlying reasons for the attachments and is covered in later chapters.

3.2.1 Passive Tension Role In Movement

Passive tension can have a fundamental role in movement. Alterations in passive tension can therefore have a fundamental role in altered movement (kinesiopathology). When passive tension is developed the force by the muscle undergoing passive tension creates a torque across the joint. For example, when the hamstrings are lengthened and generate passive tension, that passive tension creates a flexion torque at the knee. If the hamstrings are then activated (start to generate active tension) then the total tension created is the sum of the active and the passive tension, which together create the

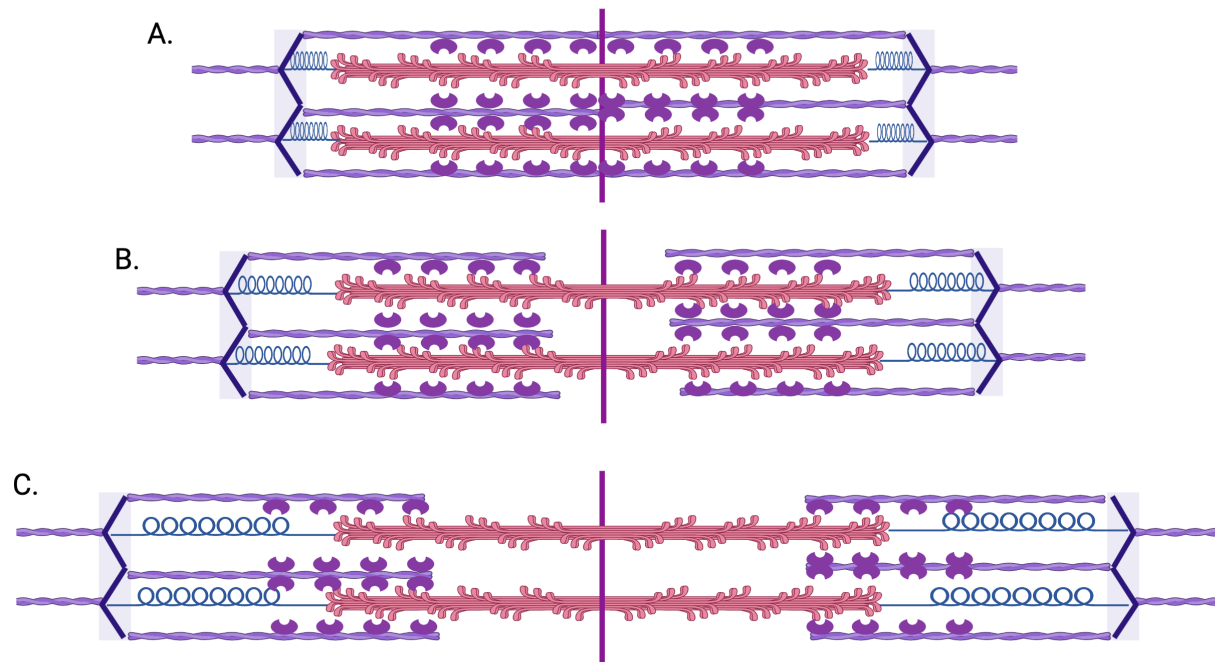


Fig. 3.3: Sarcomere Lengths. A. Shortened; B. Relaxed / resting; C. Lengthened (Created with Biorender.com)

force that creates the flexion torque at the knee. Muscles involved in gait (walking) can cycle through periods of passive tension and then use that passive tension to supplement active tension. For example, during running the calf muscles are lengthened while the foot is on the ground which increases passive tension. When it is time for that foot to propel the body the torque produced by the calf muscles is a combination of passive and active tension. In this situation, additional passive tension reduces the need for active tension and thus increases muscle efficiency (less energy is required, and therefore the muscle can sustain tension for a longer period of time).

3.3 Active Tension

To create active tension the muscle fiber transforms the chemical energy stored in adenosine tri-phosphate (ATP) into mechanical energy allowing myosin to actively slide actin from one actin-myosin crossbridge to another. This process is activated at the level of the muscle fiber and all of the events for energy transformation occur in the sarcomere. The result is tension that pulls (or attempts to pull) the Z-discs of the sarcomere closer together.

3.3.1 Sliding Filament Model

The sliding filament model of active tension is fairly well understood. There remains active investigation into the sliding filament process, the role of titin and the details of crossbridge kinetics. Figure 3.5 depicts the steps in the process of creating active tension. The steps are also enumerated below.

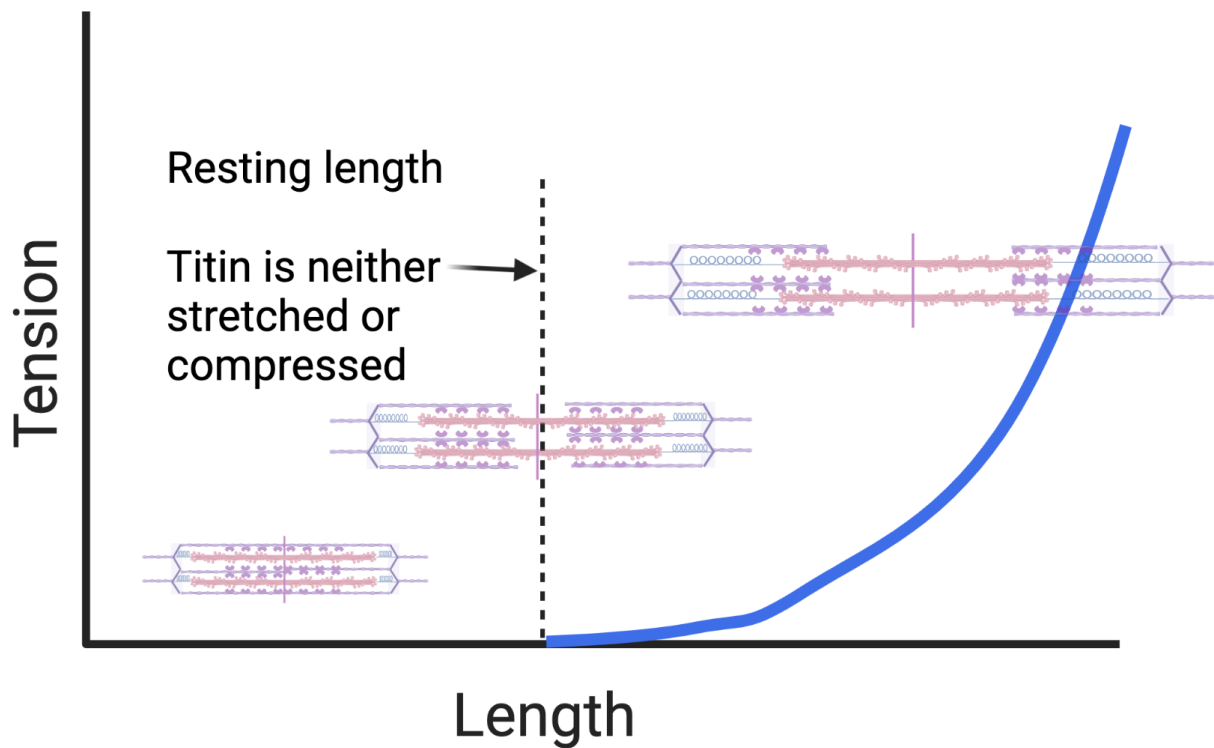


Fig. 3.4: Passive Component of the Length Tension Curve (Created with Biorender.com)

1. ADP-bound myosin head is energized with potential energy (from the release of a phosphate from ATP) and is ready to bind to actin. While the myosin head is energized, it is inactivated.
2. In the presence of calcium, calcium binds to troponin, moving tropomyosin and exposing actin binding sites so that myosin can bind to actin. In crossbridge kinetics terminology this is the change in state of the crossbridge from inactivated to activated.
3. The bound myosin rotates its head (releasing potential mechanical energy), producing a 'power stroke'. The crossbridge remains activated.
4. ATP molecule binds to the myosin head and changes its shape allowing it to detach from actin (state change from activated to inactivated).
5. Actin and myosin are detached and not energized (inactivated)

3.3.1.1 Crossbridge Kinetics

The current understanding of the molecular basis for muscle contraction comes from Huxley's (1957) kinetic model of the cyclic interaction between actin myosin crossbridges [Hux57]. Active tension behavior of a sarcomere (such as the force - velocity relationship described below) does not originate from the behavior of individual crossbridges, but from the collective action of all activated crossbridges in the sarcomere as they asynchronously go through the energy (ATP) dependent cycle described in the sliding filament model. In the crossbridge kinetics model the sarcomere has, at any given length, a number of possible crossbridges. The number of possible crossbridges is the crossbridge interaction space. Each possible crossbridge has two states, activated and inactivated. The active tension of a sarcomere is directly proportional to the number of crossbridges in the activated state. The number of activated crossbridges is influenced by both the concentration of calcium in the sarcomere and the length of the sarcomere. The sliding filament model describes one cycle of crossbridge kinetics (one cycle of inactivation to activation to inactivation).

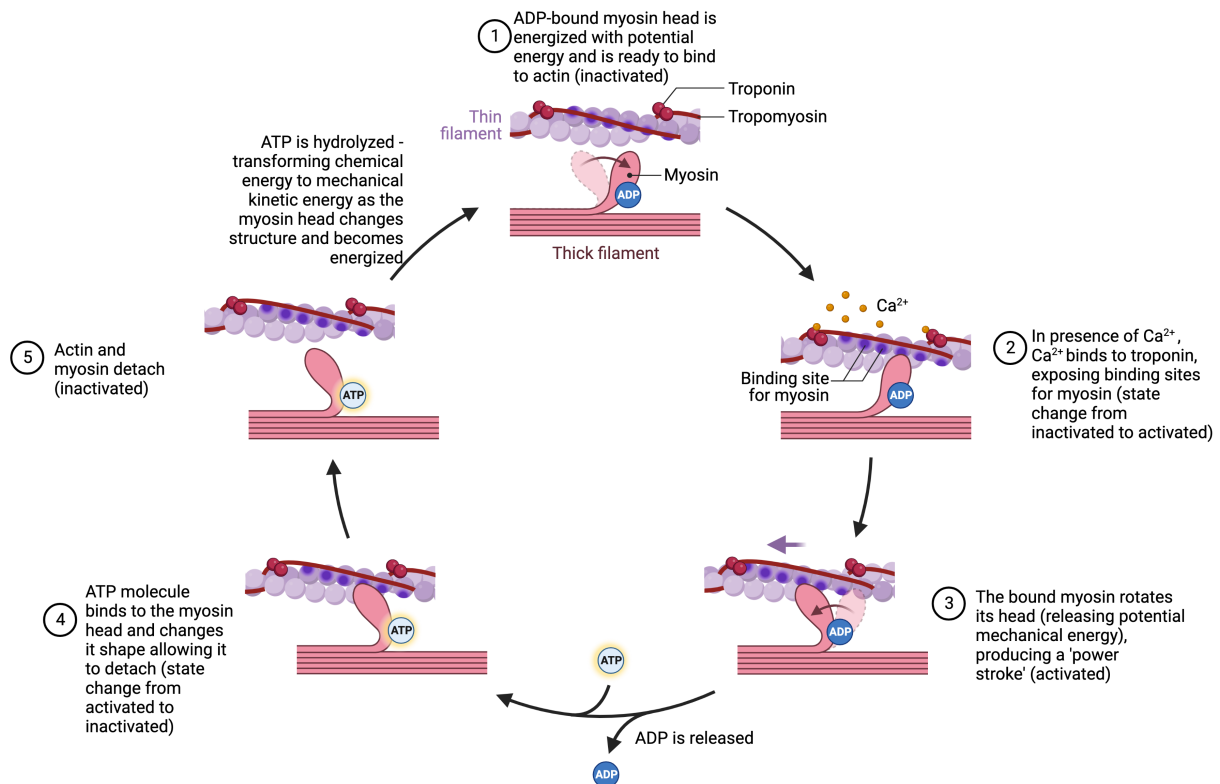


Fig. 3.5: Sliding Filament Model of Active Tension (Created with Biorender.com)

3.3.1.2 Introduction to Twitch

A twitch is the fundamental unit of active tension. It is a spike in active tension generated when crossbridges are activated from one excitation of a muscle fiber. A twitch can be considered at the level of the sarcomere, myofibril, fiber and muscle. A twitch *in situ* is perceptible but not functional. If many muscle fibers twitch at the same time it may be experienced as a spasm, though not long lasting. Sustained spasms (myotonia) involve more than a twitch. A twitch can be thought to equate to one cycle of crossbridge kinetics. The force of a twitch is directly proportional to the number of crossbridges in the activated state, meaning if there are more crossbridges cycling into the activated state there is more tension. Experimentally, all sarcomere twitches will tend to produce the same tension when performed at the same length and when separated by enough time. Myofibril twitches can have small variations in tension depending on how many of its sarcomeres have crossbridge cycling from the stimulus provided. Muscle fiber twitches can vary greatly in the amount of tension during a twitch because there is greater potential for variation in the number of sarcomeres involved in crossbridge cycling. Experimentally the number of sarcomeres involved in cross bridge cycling depends on the amount of excitation stimulus provided.

3.3.1.3 Introduction to Tetany

When a sarcomere is repeatedly excited the resultant twitches from each excitation start to fuse and the spikes in tension from each twitch get smooth while the overall tension increases. This transition to a smooth rise and even maintenance of tension from a sarcomere is tetany. The greater the number of excitations the greater the tension developed in tetany because the result is a greater number of crossbridges in the activated state at any given moment. The amount of tension developed reflects the ability to attain tension, and the period of time tetany can last during repeated excitations is the

ability to sustain tension. In Chapter 4 the concepts of twitch and tetany are the bridge between muscle excitation and crossbridge kinetics. In Chapter 5 they are used to connect muscle regulation to excitation.

3.3.2 Active Component of the Length Tension Relationship

Since the sarcomere crossbridge interaction space varies with its length¹ the number of activated crossbridges also varies with its length. Therefore, under conditions of maximal activation (high calcium concentrations), the active tension varies with sarcomere length. The crossbridge interaction space decreases as the distance between the Z-discs increases or decreases from an ideal distance. The optimal distance between the Z-discs (optimal sarcomere length) is defined as the length that maximizes the crossbridge interaction space.

The active component of the length tension relationship was originally measured using isolated muscle fibers under isometric conditions at different lengths. However, we tend to consider the impact of the underlying phenomenon during activation that includes dynamic changes in length. As the distance between the Z-discs decreases, such as during negative Δ length (concentric) or when muscle activation begins while in a shortened length, the actin-myosin interaction space is decreased because of the actin binding sites increase in concentration with a limited number of myosin heads (near the M-line) which reduces the number of potential crossbridge sites. This reduces the active tension. As the distance between the Z-discs increases, such as during positive Δ length (eccentric) or when muscle activation begins while in a lengthened length, the actin-myosin interaction space is decreased. This occurs because parts of actin are not in proximity with myosin which reduces the number of crossbridge sites. This result is a reduction in active tension. The active component of the length tension relationship is depicted in Figure 3.6.²

3.3.3 Length-Tension Relationship (Active & Passive Components)

The full length tension relationship of a sarcomere, and a myofibril, is depicted in Figure 3.7. This relationship is based on experiments with myofibrils removed from their anatomical locations and maximally activated under isometric (zero Δ length) conditions. The length tension relationship of any particular muscle *in situ* may vary from this general relationship based on factors such as the normal resting position of the muscle with anatomical attachments, the pennation arrangement, whether the conditions of maximal activation are met (rarely), whether the conditions of zero Δ velocity is met (rarely during functional movement), or whether the muscle crosses more than one joint. Two situations deserve special mention.

3.3.3.1 Manual Muscle Testing

During manual muscle testing (MMT) physical therapists are taught about the ideal joint range of motion (ROM) for testing. These ROMs are typically positions that put the muscle at a length intended to maximize tension based on the active component of the length tension relationship.

3.3.3.2 Active & Passive Insufficiency & Sufficiency

A special situation arises in muscles that have attachments crossing more than one joint. The ROM of both joints influences the muscle length. Since muscle length influences both active and passive tension the position or movement of both joints influences active and passive tension. **Active insufficiency** occurs when the active tension a muscle creates is impacted by the position or movement of the joints (at least 2) that it crosses. For example, if you attempt to extend the hip and flex the knee at the same time using active tension your ability to generate active tension declines quickly due to the fact that the hamstring sarcomere Z-discs are getting closer and actin-myosin interaction space is decreasing.

¹ Distance between the Z-discs

² The depiction of the actin myosin crossbridges in this particular rendering includes overlap of the actin during the shortened range. Whether there is such overlap vs. crowding of sites near the M-line is still an open discussion, or at least as far as the author is aware.

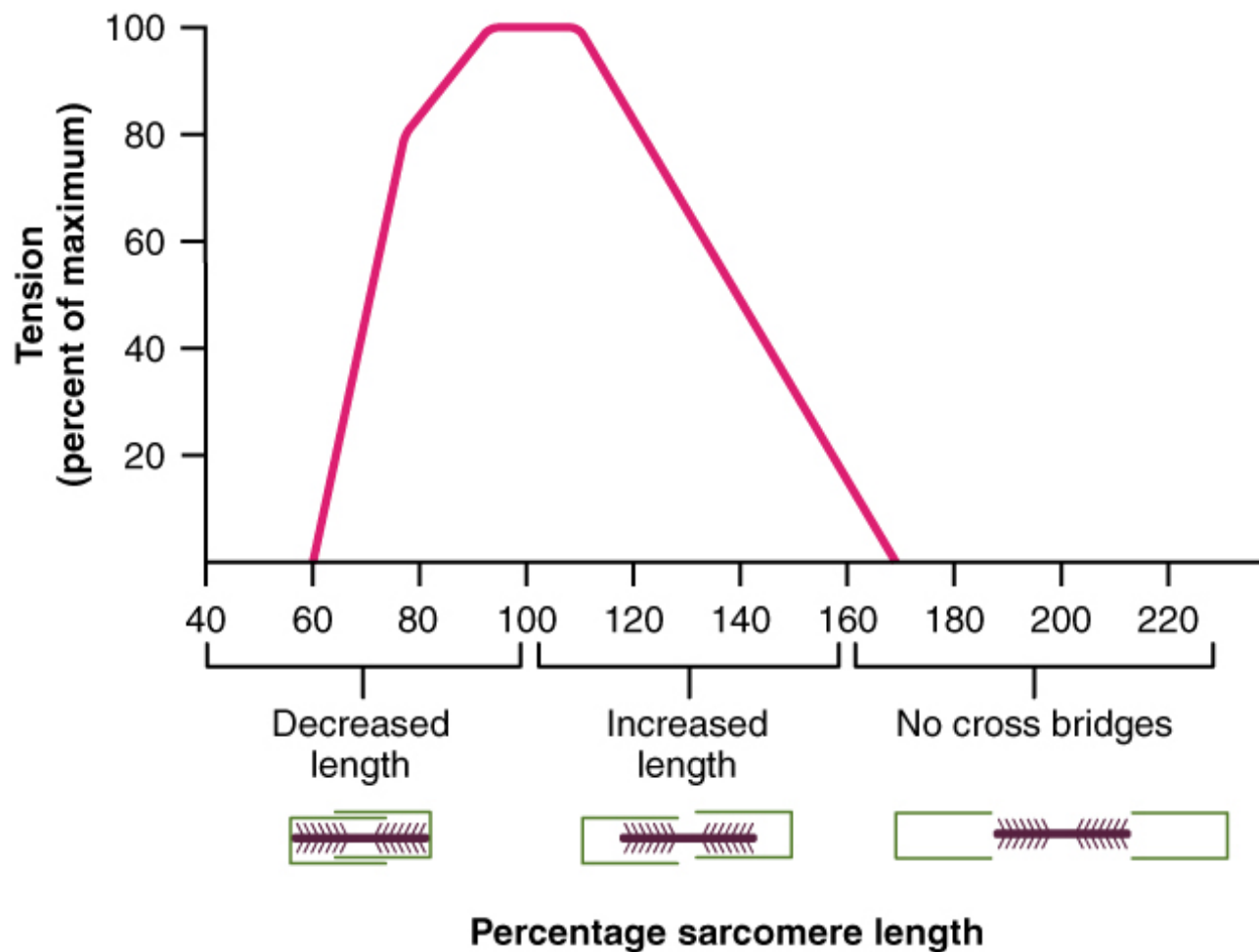


Fig. 3.6: Active Component of the Length Tension Curve (OpenStax, CC BY 4.0, [via Wikimedia Commons](#))

Passive insufficiency occurs when the passive tension a muscle creates is impacted by the position or movement of the joints (at least 2) that it crosses. For example, if you attempt to flex the hip and extend the knee at the same time, passive tension develops due to the hamstring sarcomere Z-discs getting further apart and titin is resisting further changes to its length.³ **Active and passive sufficiency** is the principle that most functional movements avoid active and passive insufficiency. Most functional movements tend to use multi-joint muscles in a way that optimize their tension generating capabilities based on the length tension relationship. For example, during squatting and climbing, the hamstring sarcomere length does not change much since these movements couple hip flexion with knee flexion (one direction of movement), and hip extension with knee extension (the other direction of movement).

3.3.4 Necessary Molecules

Most of the molecules involved in the sliding filament model are proteins in the sarcomere. There are two molecules that make important appearances and deserve special mention as a preface to where we are heading as we progress toward connecting muscle physiology to clinical physiology.

³ Please note that the focus here on titin is for simplicity. Other connective tissues within the muscle contribute to this passive tension. It is also possible that extra-muscular connective tissue such as fascia can contribute to the passive tension developed during passive insufficiency since these movements tend to not serve critical movement functions. They tend to be movements contrary to what would be considered critical movement functions.

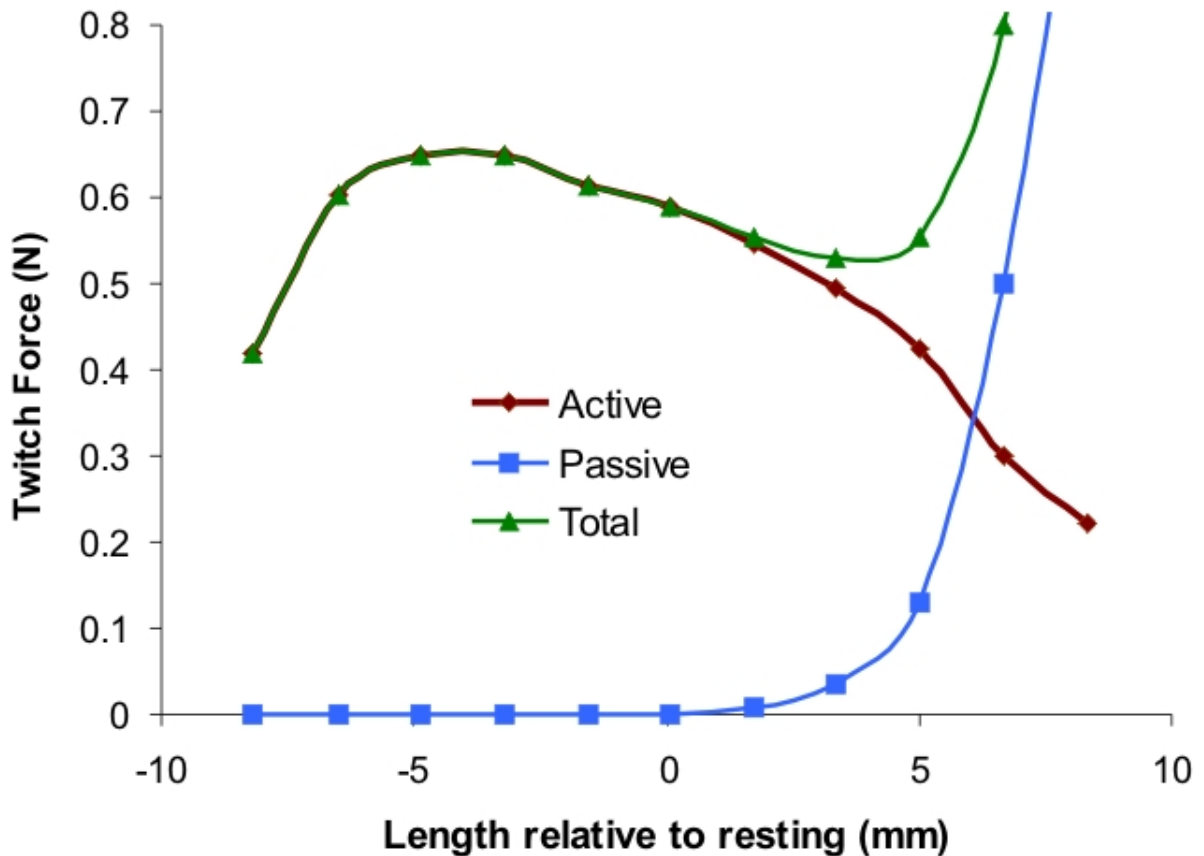


Fig. 3.7: Full Length Tension Curve with Total Tension (Public Domain figure from [Wikimedia Commons](#))

3.3.4.1 Calcium

Calcium(Ca^{2+}) enters the scene to get things started. It triggers events that lead to the crossbridge state change from inactivated to activated. It is the signaling molecule utilized to activate the process of the sliding filaments for active tension. It is stored in an organelle in close proximity to the sarcolemma called the sarcoplasmic reticulum. Ca^{2+} enters the cell once the sarcolemma and sarcoplasmic reticulum is excited and is an important part of Chapter 4 on Excitation. Ca^{2+} is the connecting molecule in the process called “excitation - contraction coupling” which we call “excitation - activation coupling”. Having sufficient Ca^{2+} for the muscle fiber is so essential that the endocrine system helps ensure it is always available, even at the expense of Ca^{2+} in the bones.

3.3.4.2 Adenosine Triphosphate (ATP)

Adenosine Triphosphate (ATP) provides the chemical energy that is transformed into mechanical energy by the molecular motors within the sarcomere. ATP is kept in limited quantities inside the muscle cell. It is used to energize cellular processes that require energy (do not confuse cellular processes with chemical reactions). Examples of cellular process include transcription and translation or a pump that moves ions across a membrane to form a gradient. Chemical reactions are transformations that proceed bidirectionally with one direction usually favored by mass balance and enzyme catalysts. Since ATP is stored in limited quantities (enough for seconds of myosin energizing), the muscle fiber must be able to convert other forms of energy (usually chemical energy) into ATP. That is the primary topic of Chapter 6 on Muscle Energetics.

3.4 Force & Velocity

There is a relationship between force and velocity. The relationship between them is considered from two perspectives: the Force - Velocity (FV) and the Velocity - Force (VF) Relationships. This chapter introduces the concept and these two perspectives (FV & VF), and considers the how sarcomeres contribute to this understanding. By force we are referring to force being measured that is the result of muscle tension (total tension). By velocity we are referring to changes in length over time.

Δ length and velocity are related based on Equation 3.1:

$$velocity = \frac{\Delta length}{time} \quad (3.1)$$

Velocity involves changes in length which makes the force velocity relationships different from length-tension relationships. Length tension relationships are experimentally observed under conditions of isometric (zero Δ length) conditions and is theoretically extended to dynamic (non zero Δ length) conditions. Force velocity relationships must be experimentally observed under dynamic (non zero Δ length) conditions. Since velocity involves changes in length, the force velocity relationships must include considerations related to the implications of the length tension relationship. Even though the length tension relationship is observed under conditions of isometric activation, what the length tension relationship represents (changes in sarcomere structure that result in changes in passive elastic and active tension) is occurring during a change in sarcomere length.

Movement Energetic Perspective

Before proceeding with the FV and VF relationships it is important to point out that these relationships can be differentiated from a third perspective - the movement energetic perspective. Movements are the sum of torques acting across joints with a larger imbalance between torques producing a “faster” movement, and keeping in mind that torques are generated by force which is generated by active muscle tension, and that active muscle tension is proportional to the number of activated crossbridges, and that the number of activated crossbridges is proportional to the amount of ATP being utilized, we can conclude that “faster” movements utilize more ATP. Therefore they are more energetic. The movement energetic perspective is discussed in Chapter 6 on Muscle Energetics. In summary, it is why running at a faster pace makes you breath heavier than a slower pace; and why certain running paces (sprinting) are not possible for very long.⁴

3.4.1 Force - Velocity & Velocity - Force Relationships

Velocity - Force Relationship

We consider the VF Relationship under the condition of no resistance. In the VF relationship force is a function of velocity. The muscle is not resisted⁵ and we consider the force that can be achieved with varying levels of velocity. The independent (manipulated) variable is on the X axis, and as the dependent (observed) variable is on the Y axis. We go into more depth for the VF relationship in Chapter 5 on Muscle Regulation since the phenomenon involves variations in muscle fiber types (heterogeneity between muscle fibers).

Force - Velocity Relationship

We consider the FV Relationship under the condition of resistance. In the FV relationship velocity is a function of force (as in the force required to move a load). The relationship between force and shortening velocity is visualized by plotting

⁴ The movement energetic perspective is not traditionally considered as part of the FV and VF relationships.

⁵ By anything of any significance beyond the resistance of joint friction, limb mass and air resistance. The point is that, in this situation, velocity is not impacted by the necessity for the muscle to generate enough active tension to overcome the force of the resistance.

the velocity of a shortening muscle as a function of the load (or force) pulling on the muscle [SS22]. The muscle is resisted and we consider the velocity that can be achieved with varying levels of resisted muscle force. The independent (manipulated) variable is on the X axis, and the dependent (observed) variable velocity is on the Y axis. FV has an inversely proportional hyperbolic relationship and was first described by Hill [Seo13]. Because power is the product of force and velocity it has a parabolic relationship which achieves its peak at neither the extreme of velocity or force [Seo13].

3.4.1.1 Force Velocity Relationship during Shortening (Concentric)

Figure 3.8 shows the FV relationship during concentric (shortening) activation with power. It is a consistent finding that peak power occurs between 20 and 30% of peak force.

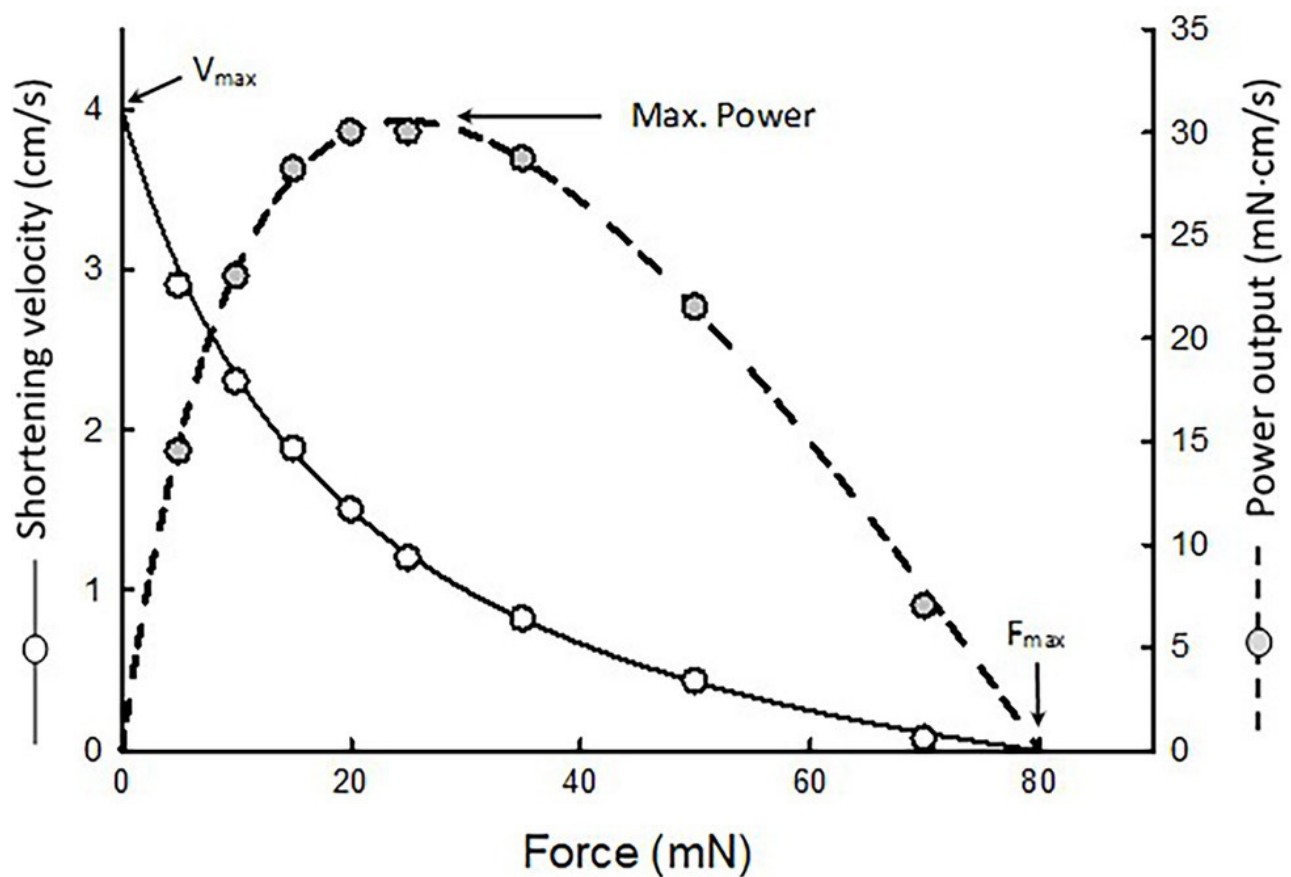


Fig. 3.8: Force Velocity Curve and Power Relationship during Shortening (Creative Commons Attribution License (CC BY) Figure from [Seo13])

There are two factors contributing to the FV relationship during concentric activation. First, is the fact that a concentric activation requires the muscle generate more torque than the resistance (load). When someone generates just enough torque, they generate a minimal (non zero) velocity. As load is reduced, the maximal active tension of the muscle creates an abundance of torque and this results in greater velocity.

Second, under the condition of attempting to move as fast as possible, as the load on and the force generated by a muscle increases, there is an increase in the velocity. The drop in force is due to the drop in load, which then allows for a higher velocity. However, it is also true that at the higher velocity there is a lower capability for generating force (the VF relationship). This is due to two factors that can be traced back to crossbridge kinetics. First, with increased velocity of shortening the crossbridge interaction space decreases rapidly which brings the sarcomeres into a shorter length and

reduces the tension they can create (i.e. length tension relationship). Second, with increased velocity of changes in length there is less time for crossbridges to be in their activated state, which means at any moment there is a lower number of possible crossbridges in the activated state.

Experience of the FV relationship during concentric activation

The FV relationship during concentric activation is something people can easily relate to and have experienced. There is a minimal velocity associated with any activity people perform with maximal resistance (force). Note that minimal velocity varies between activities. For example, the minimal velocity of a dead lift is lower than the minimal velocity of a power clean.

3.4.1.2 Force Velocity Relationship during Lengthening (Eccentric)

The FV relationship for eccentric (lengthening) activation is depicted by plotting velocity on the X-axis and force on the Y-axis since it requires a change in the direction (sign) of velocity that occurs as you go from concentric to eccentric activation. We will see this same axis flip when we consider the VF relationship in Chapter 5.

Figure 3.9 shows the relationship between force and velocity during both shortening (right side of Y axis, positive velocity) and lengthening (left side of the Y axis, negative velocity). There is a rapid increase in force at the initial transition point. The force (load, resistance) is now overwhelming, that is exceeding, the capacity of the sarcomere to generate force. But it is just exceeding which engages the passive elements of the sarcomere (titin) to contribute to force and has not yet resulted in rapid loss of the crossbridge interaction space or time for crossbridge activation. The force that can be generated by the muscle (in opposition to the external force (load, resistance)) quickly plateaus once titin has been fully engaged and the increased velocity diminishes time for crossbridge activation and the lengthening of the muscle results in less crossbridge interaction space (again, due to the length tension relationship).

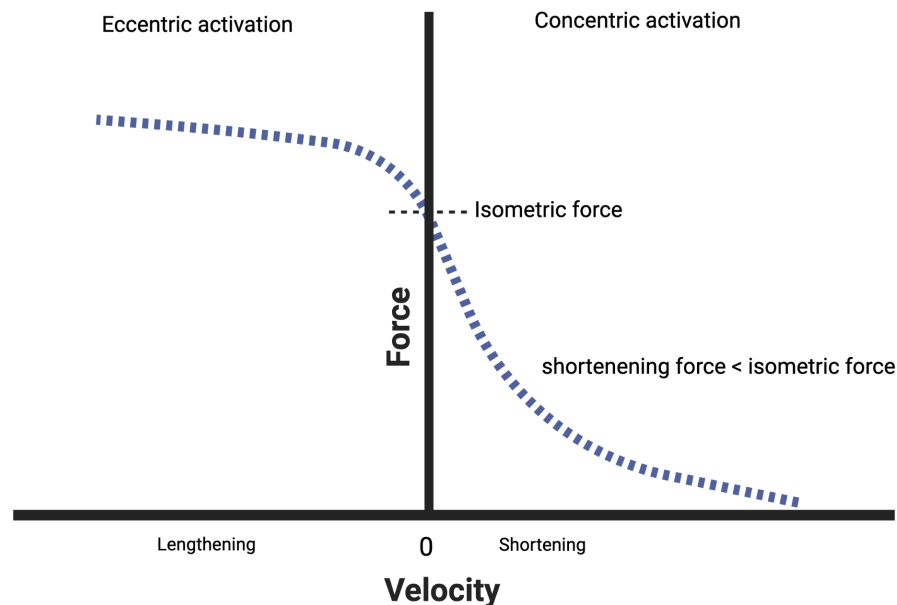


Fig. 3.9: Force Velocity Curve during Shortening (Created with Biorender.com)

Experience of the FV relationship during eccentric activation

The FV relationship during eccentric activation is something people can easily relate to and have experienced. Very simply, as the load someone attempts to hold (but cannot actually hold) increases, the velocity of movement also increases.

3.5 Summary

In this chapter we have covered the micro anatomy of muscle fibers at sufficient depth to understand the mechanisms behind active and passive tension as well as several concepts relevant for physical therapy practice such as the length tension and force velocity relationships. The sarcomere is the functional unit of the muscle fiber for active tension. Within the muscle fiber, the parts and mechanisms within a sarcomere can generate active tension by converting chemical energy into mechanical energy, or passive tension by resisting change to length. Active tension involves a sequence of activation events repeated over and over. The sequence of activation begins with excitation. Therefore, our next step is to consider excitation.

References

- [Hux57] A.F. Huxley. *Muscle Structure and Theories of Contraction*. en. Progress in Biophysics and Biophysical Chemistry 7 (1957), pp. 255–318. ISSN: 00964174. DOI: [10.1016/S0096-4174\(18\)30128-8](https://doi.org/10.1016/S0096-4174(18)30128-8). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0096417418301288> (visited on 05/25/2022).
- [Seo13] Chun Y. Seow. *Hill's equation of muscle performance and its hidden insight on molecular mechanisms*. en. Journal of General Physiology 142.6 (Dec. 2013), pp. 561–573. ISSN: 1540-7748, 0022-1295. DOI: [10.1085/jgp.201311107](https://doi.org/10.1085/jgp.201311107). URL: <https://rupress.org/jgp/article/142/6/561/43161/Hills-equation-of-muscle-performance-and-its> (visited on 05/25/2022).
- [SS22] Kathryn N. Seow and Chun Y. Seow. *Molecular Events of the Crossbridge Cycle Reflected in the Force–Velocity Relationship of Activated Muscle*. en. Frontiers in Physiology 13 (Mar. 2022), p. 846284. ISSN: 1664-042X. DOI: [10.3389/fphys.2022.846284](https://doi.org/10.3389/fphys.2022.846284). URL: <https://www.frontiersin.org/articles/10.3389/fphys.2022.846284/full> (visited on 05/25/2022).

Sample Questions

1. Which structures do not have a primary role in active tension?
2. Which structures do have a primary role in active tension?
3. What is the correct sequence for generating active tension?
4. Maximal quadriceps shortening velocity during knee extension at an 80% maximal tension load would be a (higher or lower) velocity than during knee extension at 90% maximal tension load.
5. Maximal quadriceps tension during knee extension at 360 degrees per second would be (higher or lower) than during knee extension at 30 degrees per second.

Chapter 4

Excitation

Updated on May 22, 2024

Contents

4.1	Muscle Fiber Excitation	42
4.2	Excitable Membranes	50
4.3	Clinical Physiology Connections	56
4.4	Summary	59

This chapter covers muscle excitation, the molecular mechanisms that allow excitation, and related clinical physiology connections.

Muscle fiber excitation starts with an excited motor axon (α -motor neuron) and ends with the binding of calcium to troponin. Binding calcium to troponin allows a crossbridge to go from an inactivated to an activated state. The final step in muscle excitation is the coupling between excitation and activation (Excitation - Activation Coupling).

The underlying molecular mechanisms of the muscle excitation pathway include the characteristics and actions of excitable membranes. Excitable membrane characteristics including ion channels, pumps and receptors are fundamental to many areas of clinical physiology. Actions include using transport to establish a resting membrane potential and the ability to have an action potential (excitation).

Understanding excitable membranes creates Clinical Physiology Connections opportunities in three areas: 1. systemic regulatory function of the neuroendocrine system; 2. function of sensory receptors; and 3. pharmacodynamics.

Objectives include:

1. Explain the events of muscle excitation and excitation - activation coupling.
2. Relate the events of excitation - activation coupling to the creation of active tension.
3. Explain the components and capabilities of an excitable membrane.
4. Explain the events that result in sarcolemma depolarization.
5. Explain the events that result in neuromuscular endplate depolarization.
6. Explain the possible result of hypothetical ("what if") situations on the excitation and activation of a muscle fiber (for example, What if calcium was depleted and less was being released from the SR?; What if ACh was not broken down quickly at the NMJ?).
7. Explain the function and role of the neuroendocrine system for homeostasis.
8. Explain the basic function of sensory receptors.
9. Explain pharmacodynamics based on actions occurring at end organ receptors
10. Demonstrate the ability to apply basic physiology concepts such as the cell membrane to the analysis of patient/client problems related to the generation of muscle excitation, the effectiveness of the neuroendocrine system, sensory receptors and pharmaceuticals.

4.1 Muscle Fiber Excitation

Excitation

Muscle fiber excitation is the signal that initiates crossbridge activation and results in a twitch. Excitation is a state of the membrane. Based on the concept of excitation a muscle fiber exists in one of two states, resting or excitation. Excitation occurs when a membrane depolarizes and the process of depolarizing is called an action potential (AP). The resting state of the membrane is polarized and is the resting membrane potential (RMP). Cellular membranes that can undergo excitation are excitable membranes, and include nerve and muscle cell membranes.

Excitation as Signal Transduction

Excitation is a mechanism for signal transduction (sending and receiving signals). One example is when membrane excitation results in cellular actions that influence cellular processes, such as exocytosis¹ at the terminal axon at the neuromuscular junction. A second example is when a membrane is excited through binding of a ligand² to a receptor on its surface which can then excite the surface, direct another signal into the cell (or both).

4.1.1 Muscle Fiber Excitation Overview

Skeletal muscle fiber excitation starts with excitation of an α -motor neuron (Step 1 in Figure 4.1) which crosses the neuromuscular junction (NMJ) and excites the motor end plate (Step 2 in Figure 4.1). Excitation of the motor end plate causes excitation of the sarcolemma, which causes excitation of the transverse tubule system (T-tubules). Excitation of T-tubules cause excitation of the sarcoplasmic reticulum (SR), which causes the release of calcium into the muscle fiber and the activation of available crossbridges (Step 3 in Figure 4.1). Excitation of the SR and release of Ca^{2+} is referred to as Excitation-Activation Coupling (or Excitation-Contraction Coupling). A single excitation results in a twitch (the fundamental unit of active tension).

4.1.2 Step 1 - The α -Motor Neuron

Excitation of an α -motor neuron is regulated by the central nervous system and occurs at the membrane of its dendrites in the spinal cord. A wave of excitation travels along the axon to its terminal branches at neuromuscular junctions (NMJs) by sequential excitation facilitated by voltage gated ion channels. Each muscle fiber has one NMJ and receives one axon terminal. However, each α -motor neuron has a variable number of terminal branches (Chapter 5). Excitation travels from the spinal cord to the NMJ by means of saltatory conduction which allows excitation to take big steps as it travels along the axon membrane. Saltatory conduction increases the nerve conduction velocity and is made possible by the myelin sheathes (Figure 4.2). The membrane is exposed along the axon at nodes of Ranvier. Excitation travels by jumping from one node of Ranvier to the next.

¹ Exocytosis is a form of active transport and bulk transport in which a cell transports molecules (e.g., neurotransmitters and proteins) out of the cell). From <https://en.wikipedia.org/wiki/Ligand>

² A ligand is any molecule or atom which binds reversibly to a protein. A ligand can be an individual atom or ion. It can also be a larger and more complex molecule made from many atoms. A ligand can be natural, as an organic or inorganic molecule. A ligand can also be made synthetically, in the laboratory. From <https://biologydictionary.net/ligand/>

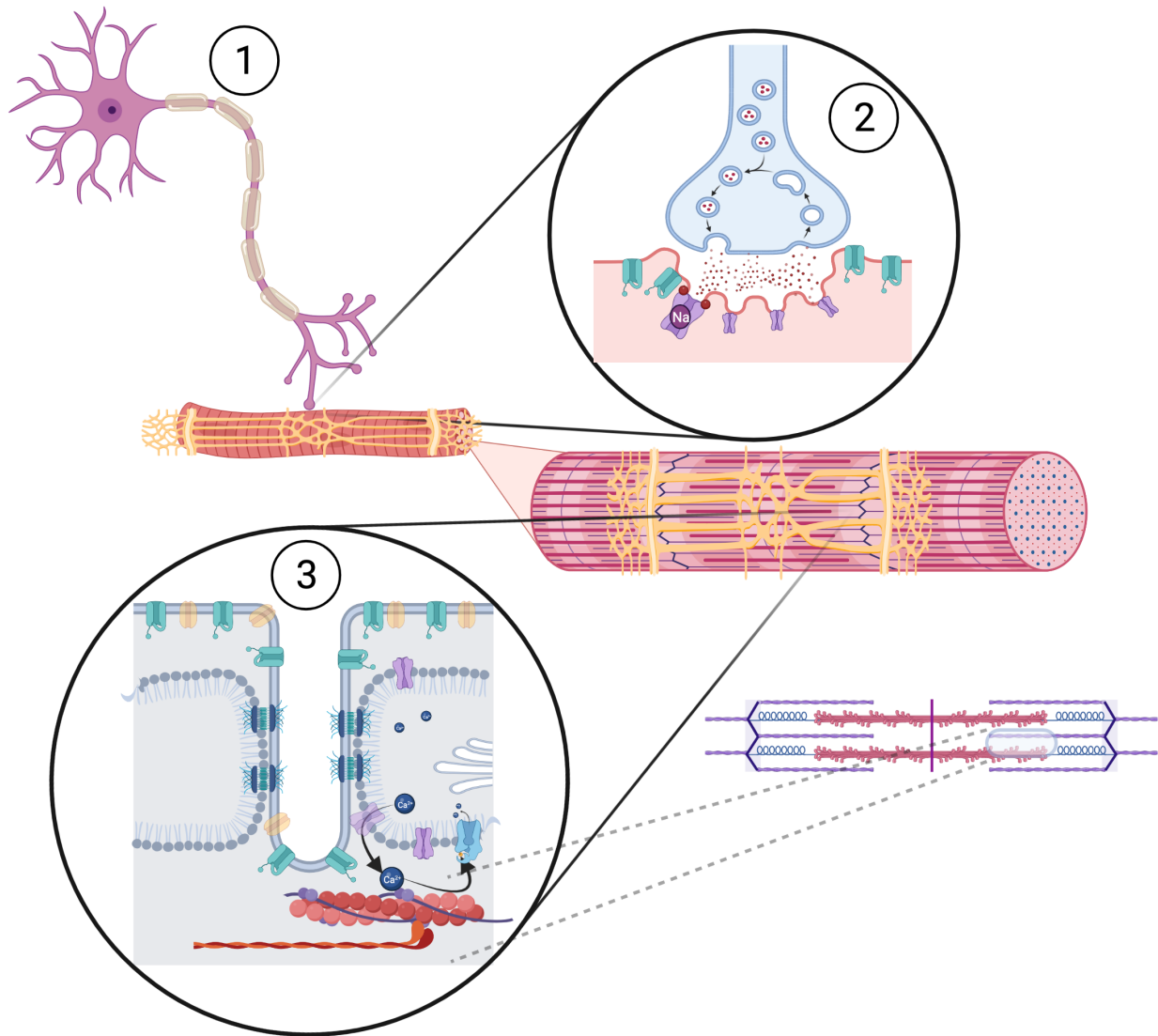


Fig. 4.1: Overview of excitation from the α motor neuron to crossbridge activation (Created with BioRender.com)

4.1.3 Step 2 - Neuromuscular Junction - Motor End Plate - Sarcolemma Excitation

Membrane excitation at the axon terminal releases the neurotransmitter acetylcholine (ACh) due to the action of voltage gated Ca^{2+} channels signaling exocytosis of vesicles filled with ACh. ACh crosses the NMJ synapse (Figure 4.3) and excites the motor end plate by binding to receptors and opening ligand gated Na^+ ion channels. ACh is quickly broken down by cholinesterase, an enzyme synthesized by the muscle fiber that hydrolyzes ACh. The free choline molecules are transported back into the axon terminal (endocytosis) to be resynthesized. Rapid uptake and reuse of ACh allows for regulation of muscle activation at the NMJ. When high concentrations of ACh accumulate at the NMJ (due to a high frequency of axon excitations) there are two processes that can result in faster excitation of the motor end plate. Faster excitation allows for a higher frequency of excitation. **Temporal summation** of receptors refers to repeated stimulation of one receptor by ACh. **Spatial summation** of receptors refers to stimulation of a larger number of motor end plate receptors. Temporal summation is limited by the amount of ACh, which is influenced by the release of ACh from the axon terminal and the breakdown of ACh in the NMJ. Spatial summation is also limited by the amount of ACh, but is also limited by the number of receptors on the motor end plate. Therefore spatial summation is adaptable over time

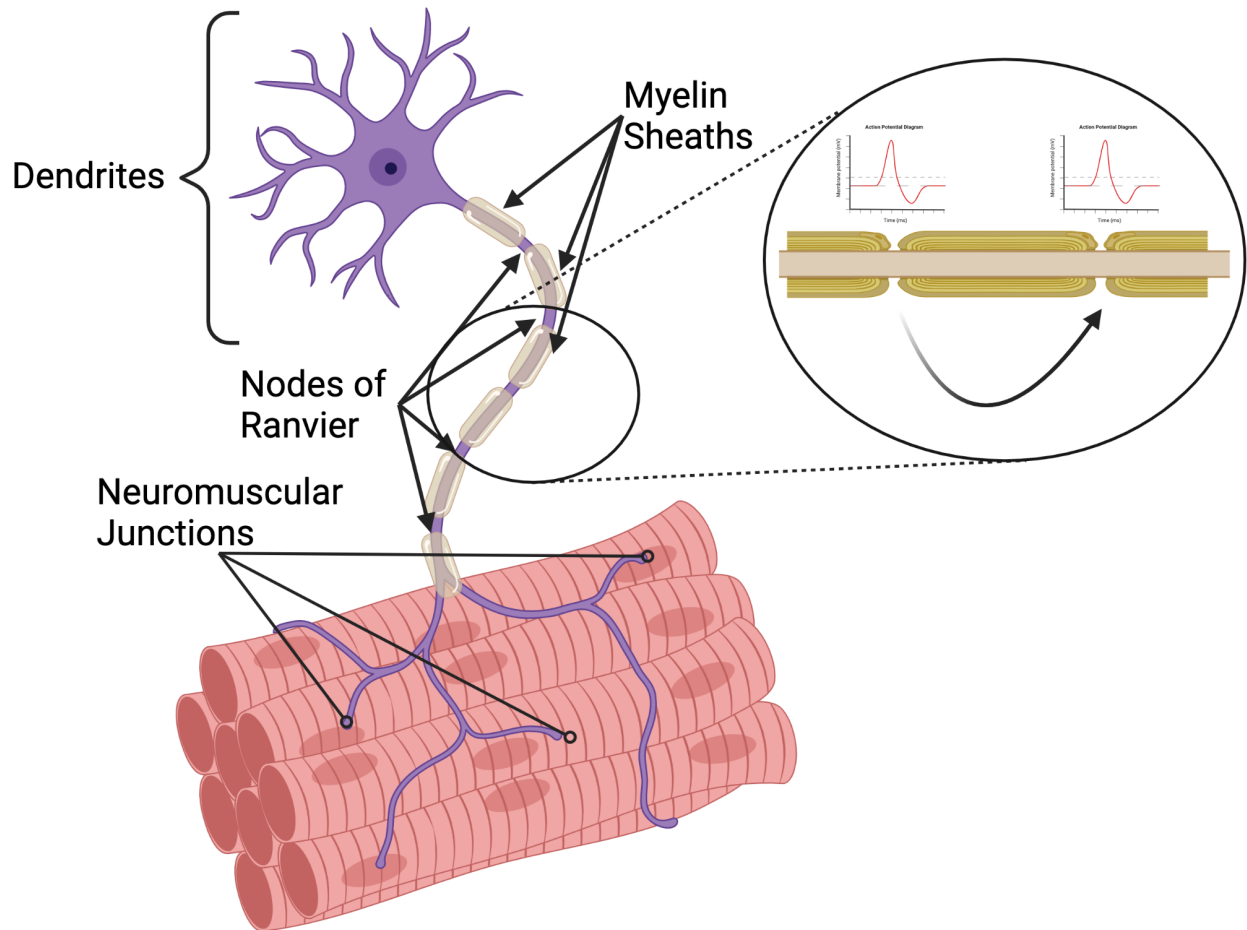


Fig. 4.2: α -motor neuron terminating at several muscle fiber motor end plates (Created with BioRender.com)

through cellular regulatory processes. Up regulation of receptors (more receptors) or down regulation of receptors (less receptors) changes the response of the motor end plate to the release of ACh.

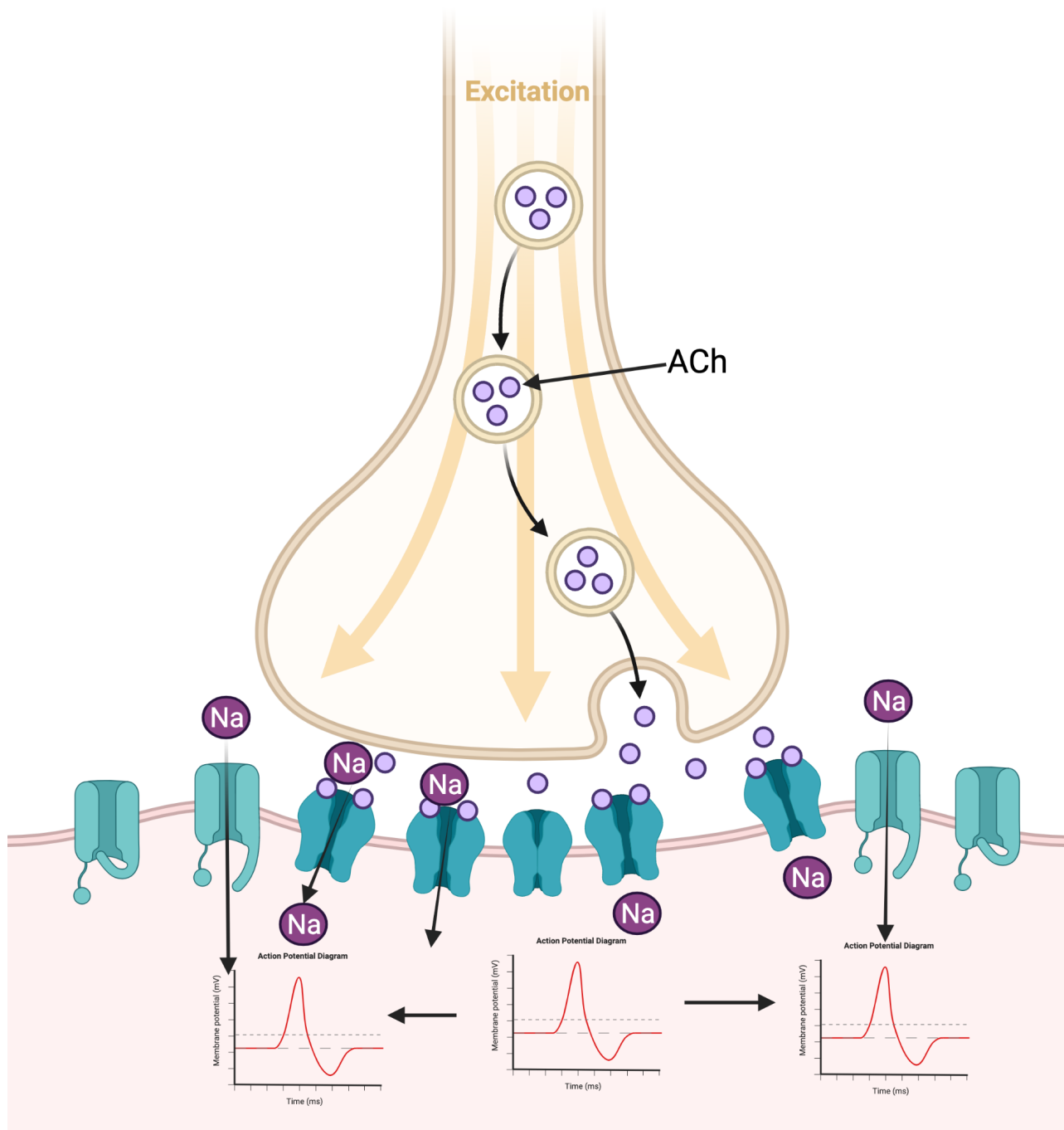


Fig. 4.3: NMJ Excitation of the Motor End Plate (Created with BioRender.com)

Excitation of the motor end plate spreads to the sarcolemma due to the opening of voltage gated ion channels on the sarcolemma. Sarcolemma excitation creates a wave similar to that in the axon (but without myelin) that spreads due to opening of voltage gated ion channels. The sarcolemma surrounds and penetrates the fiber and circles myofibrils through the intricate T-tubule system. T-tubules that encircle myofibrils form a triad with a SR on each side (See Figure 4.4).

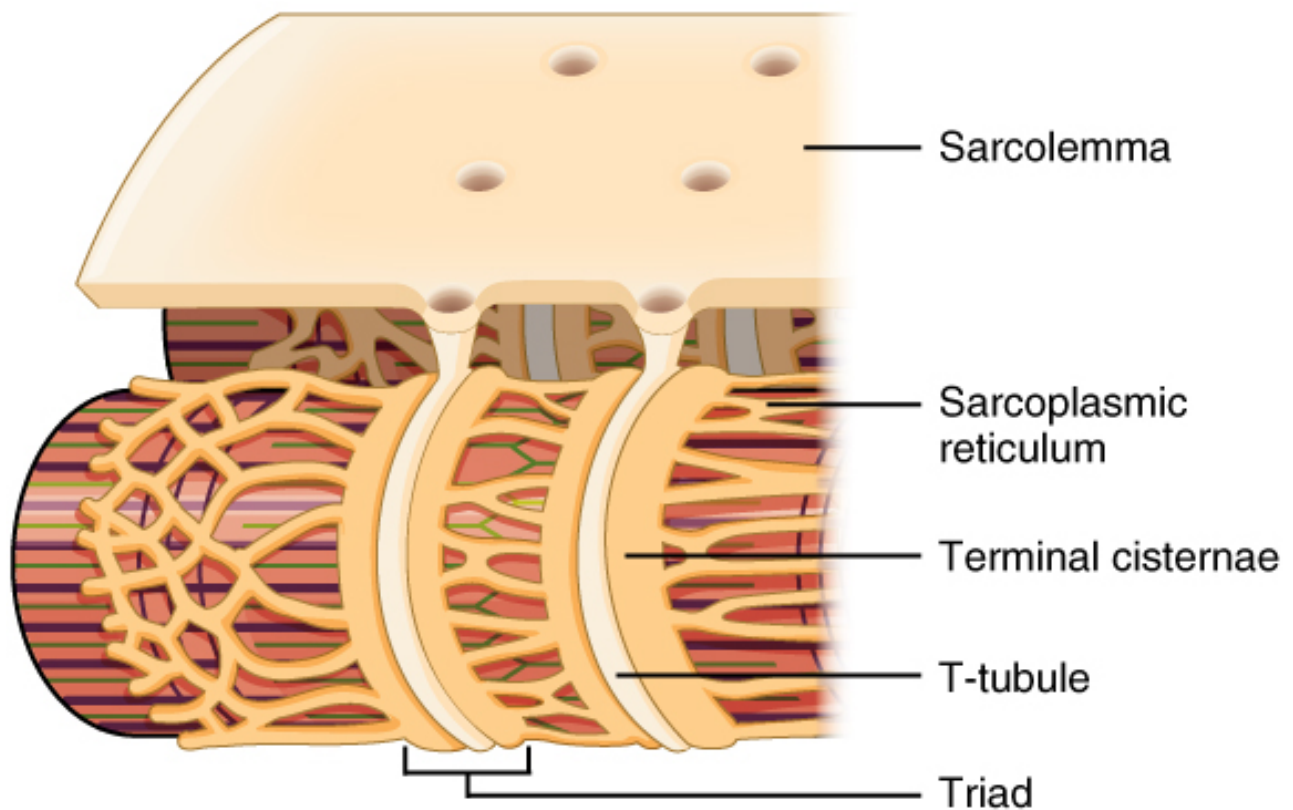


Fig. 4.4: T-Tubule Encircling a Myofibril & Forming a Triad with Sarcoplasmic Reticulum (OpenStax, CC BY 4.0, [via Wikimedia Commons](#))

4.1.4 Step 3 - Excitation - Activation Coupling (EAC)

Excitation of the T-tubules excites the sarcoplasmic reticulum (SR). There are unique junctions between T tubules and the SR connected by Ca^{2+} channels. The junctions act as a voltage sensor in the T tubules, and a Ca^{2+} release channel in the SR. The functional implication of this unique T-tubule - SR junction is that excitation of the T-tubule quickly results in a combined excitation of the SR along with the release of Ca^{2+} into the sarcoplasm. The release of Ca^{2+} in close proximity of troponin binding sites results in rapid crossbridge activation (Figure 4.5).

Sarcoplasm

Sarcoplasm is the muscle fiber equivalent of the cytoplasm, the intracellular region that is in the cell but outside of any organelles (such as the SR, nucleus, mitochondria).

Excitation & EAC Latency

Figure 4.6 represents each major phase of excitation and EAC from the α -motor neuron to development of a twitch. The latency period time scale is in milliseconds (ms, 1/1000 of a second). The latency period of each excitation and EAC forms an upper boundary on the number of possible twitches per second. However, this boundary is determined by the latency of the excitation portion of Figure 4.6 which amounts to a few milliseconds. The fact that the twitch lasts for an estimated 100 ms means that excitation (a few milliseconds) can occur at a frequency that results in twitch summation. Since each twitch takes about 100 milliseconds, they can sum together to produce tetany.

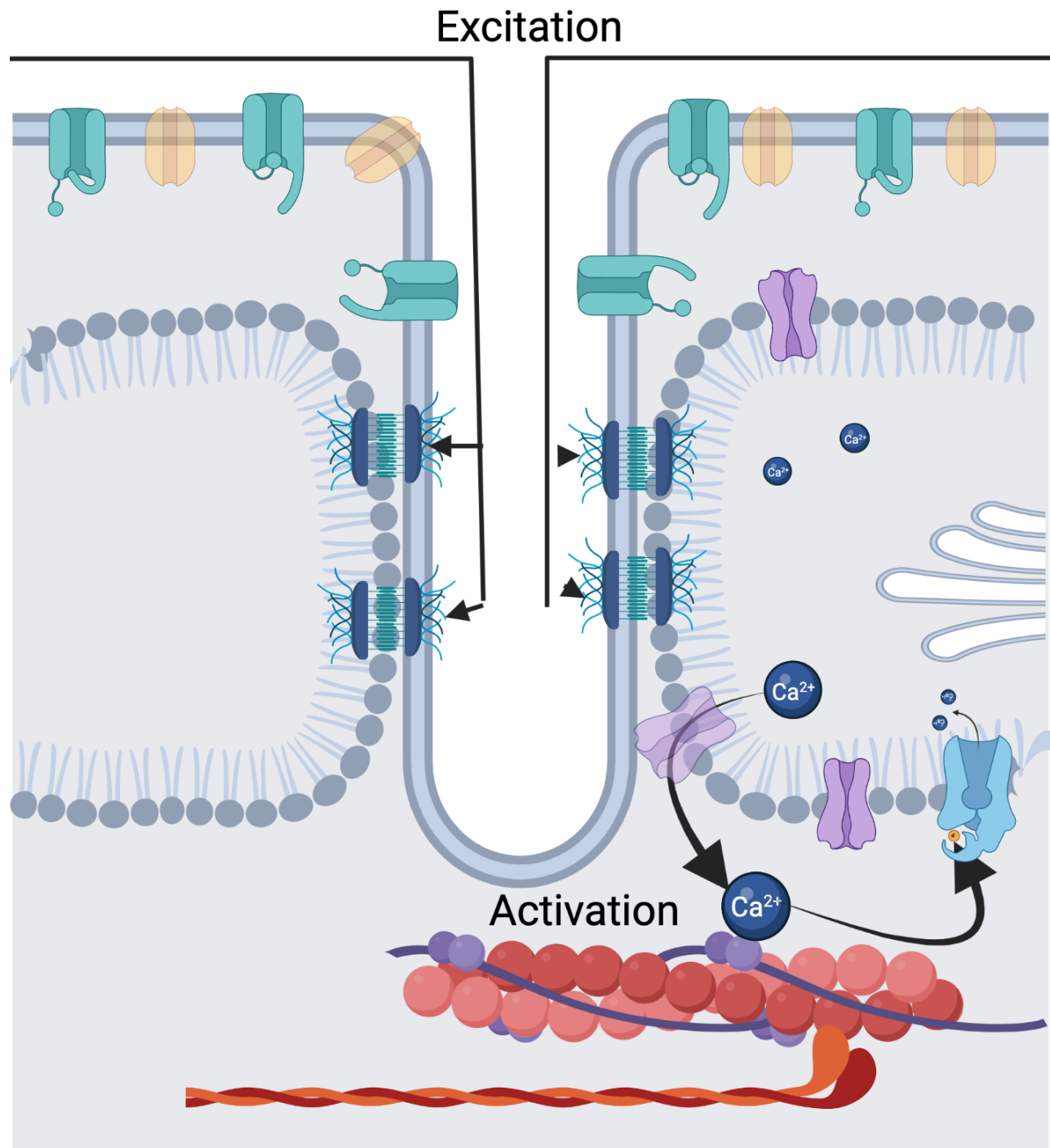


Fig. 4.5: Sarcoplasmic Reticulum as the site of Excitation-Activation Coupling (Created with BioRender.com)

4.1.4.1 Twitch to Tetany

Each excitation activates enough crossbridges to produce a twitch. A twitch is the fundamental unit of active tension. However, a single twitch of active tension is not enough tension when the goal of tension is some sort of voluntary movement or stability. The rise and rate of a twitch depends on the amount and spread of Ca^{2+} released, the available

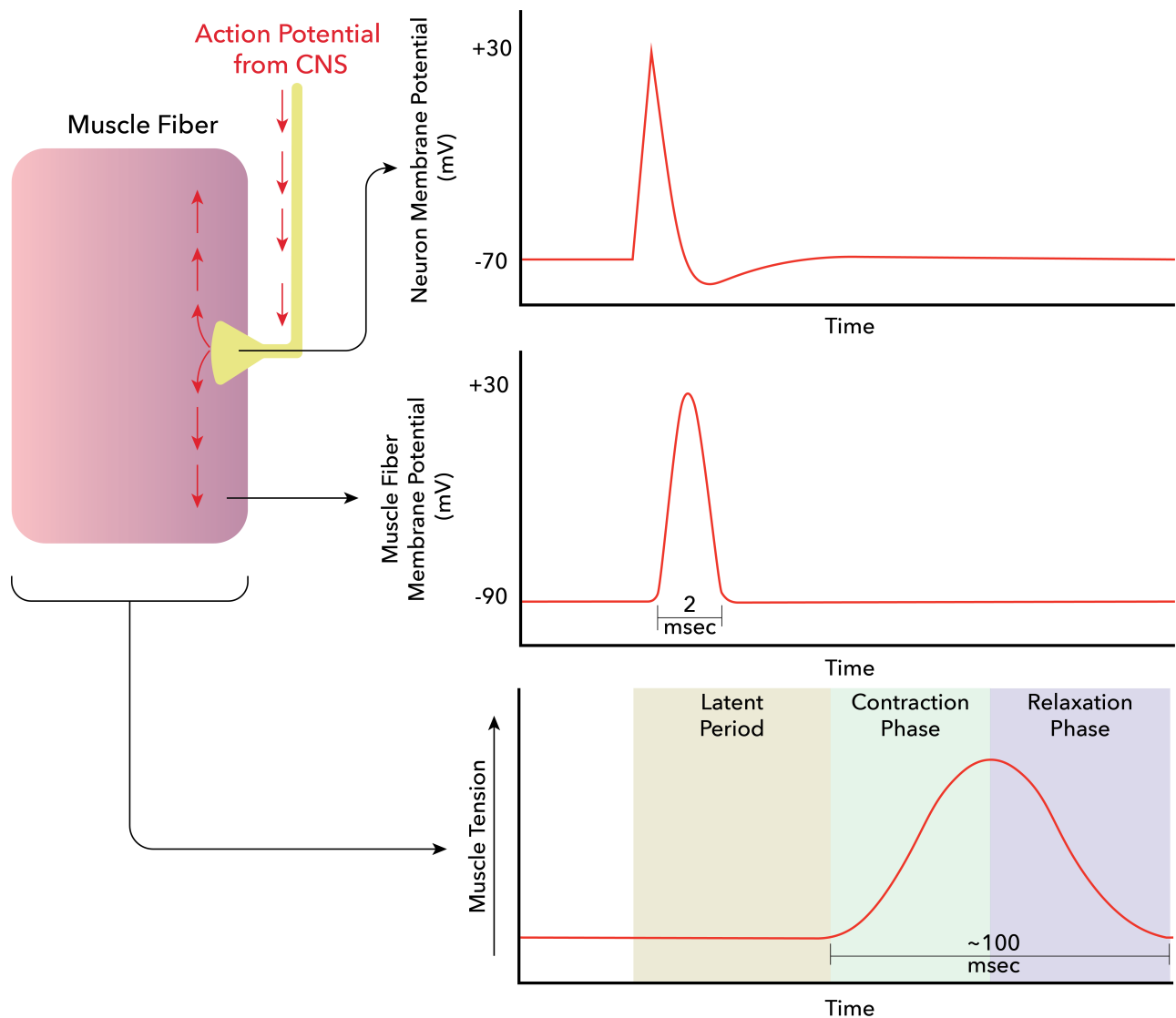


Fig. 4.6: EAC Latency (Wikimedia Commons, CC BY 4.0, [EAC-Latency](#))

troponin sites, the available crossbridges to activate as well as the amount of elasticity in the series. One mechanism of muscle fatigue is a reduction in the release of Ca^{2+} from the SR.

Tetany

Repeated excitation results in repeated activation and therefore continued crossbridge activation and repeated twitches (Figure 4.7). Successive twitches create tension summation and is called tetany. When the force of active tension is measured, such as during a muscle test, it is the result of many muscle fibers being in tetany. Variations of twitch frequency to create tetany is used to regulate active tension (Chapter 5).

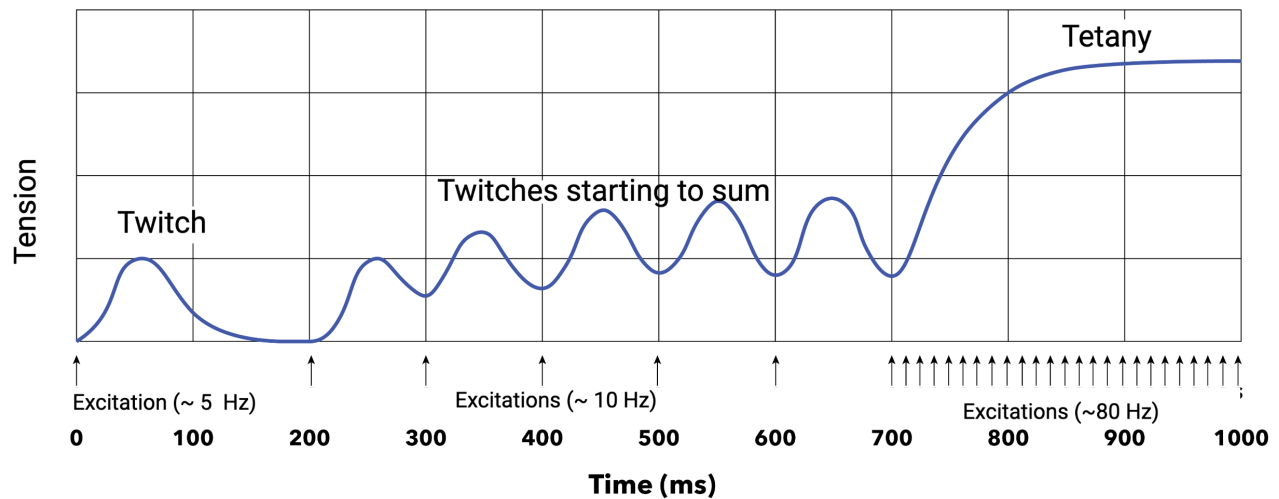


Fig. 4.7: EAC Latency (Modified with BioRender.com from [Wikimedia Commons](#), CC BY 4.0)

4.1.4.2 Removal of Ca^{2+} from Sarcoplasm

Removal of Ca^{2+} from the sarcoplasm requires active transport with Ca^{2+} ATPase. This pump is dependent on Mg^{2+} and moves 2 Ca^{2+} into the SR while moving one H^{+} out at the cost of 1 ATP. Sarcoplasm levels of Ca^{2+} can quickly return to resting levels which is necessary for rapid muscle relaxation and therefore muscle tension regulation and coordination. Situations that result in a delay in the removal of Ca^{2+} from the sarcoplasm can delay muscle relaxation producing extra twitches or spasms (involuntary tetany). Since Ca^{2+} ATPase is dependent on Mg^{2+} , a deficiency in Mg^{2+} can create such a situation.

4.1.5 Return to Resting (Non excitation) State - RMP

The process of returning to the resting state (RMP) is not explicitly discussed in the above sequence of membrane excitation events. It is implied, or taken for granted, once we started discussing the excitation of a twitch to the excitation of tetany (the summation of many twitches). Excitation of tetany requires that the axon, axon terminal, motor end plate, sarcolemma and T-tubule return to a resting (non excited) state. It is important to note that there is no excitation summation, only twitch summation.

Excitation occurs quickly because the resting state includes potential energy for excitation. The RMP is a polarized membrane with large ion concentration gradients (just waiting to run downhill with their gradients). This potential energy comes at a cost of spending energy (ATP) to pump ions against a concentration gradient. The RMP enables rapid excitation because ions are waiting with potential energy due to their concentration gradient to move quickly across the membrane (depolarize).

The RMP also enables rapid return to the resting state with potential energy. The movement of K^{+} and Cl^{-} ions with their concentration gradient can quickly end excitation enabling the membrane to be quickly returned to RMP for another excitation. The long term stability of the RMP relies on ion pumps to sustain concentration gradients for a lifetime. For example, brain and heart death are situations where the cells of the brain and heart no longer return to the resting state, they no longer repolarize to the RMP, and therefore they can no longer undergo excitation.

A deeper understanding of excitation and return to RMP, and clinical connections, requires a molecular and cellular understanding of excitable membranes.

4.2 Excitable Membranes

An excitable membrane is polarized in its resting state, which is the resting membrane potential (RMP). An excitable membrane is excited when the membrane loses its polarity, which is called depolarization. Depolarization is what occurs during an action potential (AP). What makes the membrane able to be depolarized is the fact that it is polarized. What makes the membrane capable of an AP is the fact that it has a RMP. Please excuse the repetition. Since these terms are used rather interchangeably it is important to develop familiarity with the terminology. The use of the various terms in the book reflects various terms being used across fields, papers, and other books.

In summary, a membrane cannot depolarize unless it is polarized. To understand muscle excitation and regulation (since muscles are regulated through variations in excitation) requires understanding how the RMP is established, how the RMP depolarizes with an action potential (AP), and how the AP propagates (spreads as a wave of excitation). Understanding these concepts also opens the door to understanding several additional topics in clinical physiology.

4.2.1 Resting Membrane Potential

The RMP requires several characteristics and actions of the cell membrane, and in the case of a muscle fiber, the sarcolemma.

4.2.1.1 Sarcolemma

The sarcolemma is a semipermeable phospholipid bilayer membrane that forms a boundary for the muscle fiber. Semi-permeability arises for two reasons. First, because the lipid bilayer prohibits the free passage of many molecules (a unique combination of hydrophobic and hydrophilic components). Second, there are transport proteins within the membrane for selective forms of active, facilitated and passive transport.

Cellular Transport Mechanisms

Active transport utilizes protein machines that require ATP and move very specific molecules, typically, against a concentration gradient. **Facilitated transport** utilizes protein machines to move very specific molecules across the membrane without the need for ATP. Facilitated transport requires a concentration gradient for at least one molecule, but can be coupled with another molecule that moves against a concentration gradient. Meaning, facilitated transport can utilize the potential energy of one molecule to provide the energy needed to move another molecule against a gradient. **Passive transport** occurs due to a concentration gradient (such as simple or channel guided diffusion).

The membrane separates the intracellular from the extra cellular space. Through selective transport the membrane establishes the conditions for cellular activities within the cell and thus influences (but does not completely establish) conditions outside the cell. Sarcolemma (all cell membranes) are dynamic structures that adapt in response to the intracellular and extra cellular conditions to meet the needs of the cell.

Components of the sarcolemma that are discussed below play an important role for establishing the resting membrane potential as well as the action potential are depicted in Figure 4.8.

Pumps & Transporters

Na^+/K^+ ATPase pumps are embedded throughout excitable membranes. These protein based molecular machines convert the energy of ATP to the movement of three Na^+ outside the cell and two K^+ inside the cell with each cycle of the pump (each use of ATP). Given the natural tendency of molecules to move based on a concentration gradient (high concentration to low concentration) the resultant higher concentration of Na^+ outside and higher concentration of K^+ inside the cell requires the work of these active transport pumps against the concentration gradient. The $K^+ - Cl^-$

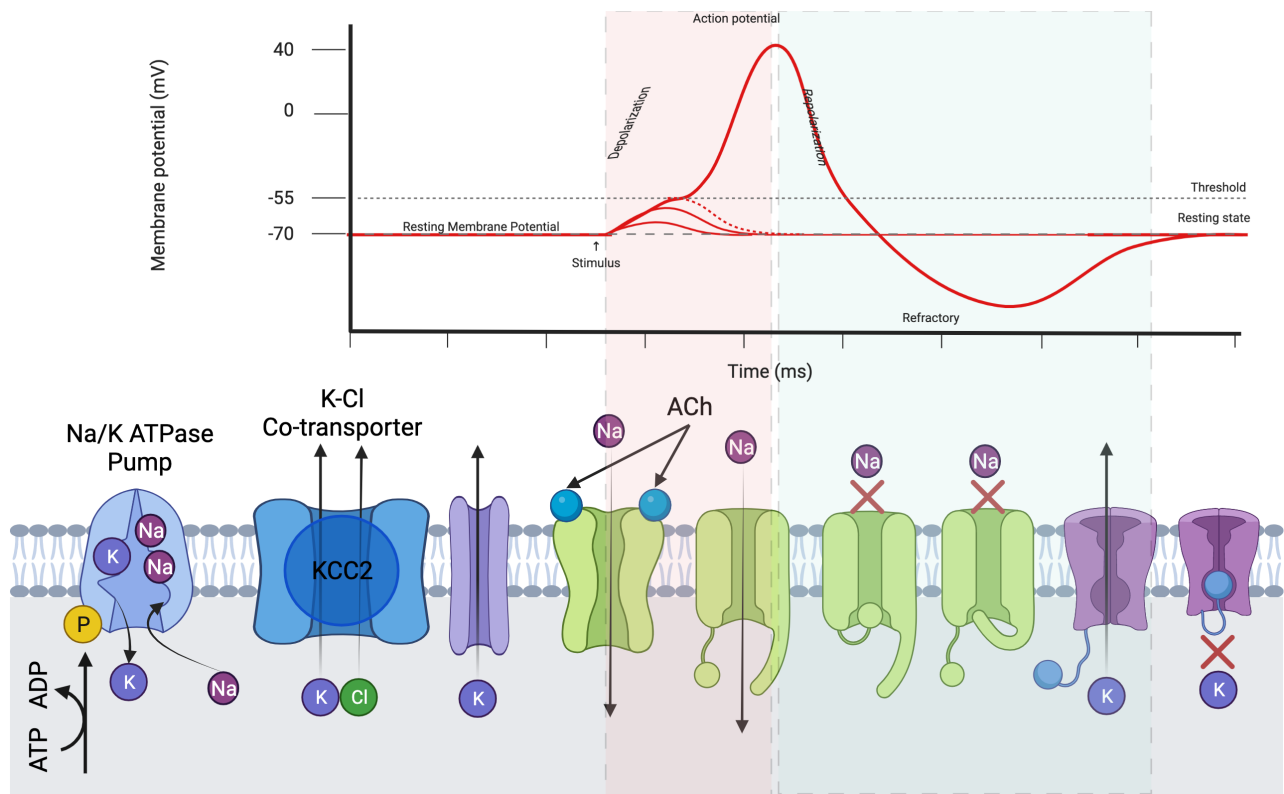


Fig. 4.8: Cell Membrane Components and their general contribution to RMP and an AP Created with BioRender.com

co-transporter (KCC2) utilizes the K^+ concentration gradient created by the Na^+/K^+ ATPase pumps to move Cl^- out of the cell as an exchange with K^+ .

Channels & Receptors

Channels are proteins embedded in the cell membrane that allow passage of molecules between the intracellular and extra cellular space. Channels can be gated and therefore only allow passage under certain circumstances.

Voltage gated channels only allow passage when the membrane potential is at certain values. They may open at a certain value and close a certain period of time later; or they may close at a certain value and open a period of time later. **Ligand gated channels** only allow passage when the channel is activated due to the binding of a ligand (a molecule that can bind). A ligand gated channel is also a receptor, or closely associated with a receptor. Critical channels for membrane excitation include:

- Na^+ channels allow the passage of Na^+ with its concentration gradient and can be either ligand gated or voltage gated.
- K^+ channels allow the passage of K^+ with its concentration gradient and in nerve and muscle excitable membranes are both non gated and voltage gated.
- Cl^- channels allow the passage of Cl^- along its concentration gradient and in the sarcolemma are not gated (Cl^- channels are not depicted in 4.8, but would resemble the K^+ channel to the right of KCC2).

Receptors are proteins that allow ligand binding. Some receptors are ligand gated channels (a receptor where the activity that binding produces is opening a channel gate). Other receptors perform different activities within the cell.

4.2.2 Establishing the Resting Membrane Potential

The process of establishing the RMP in the sarcolemma focuses on roles played by the concentration gradients and electrical potential created by Na^+ , K^+ and Cl^- . Each of these elemental molecules exist in the body as ions (they carry a charge based on whether they have too many electrons (Cl^-) or too few electrons (Na^+ , K^+ , Ca^{2+})).

4.2.2.1 Electrical Potentials & Equilibrium Potential

Movement of molecules across the cell membrane is influenced by the concentration gradient of the ion across the membrane, the permeability of the membrane for that ion and the electrical equilibrium potential across the membrane. When we describe electrical potentials the convention is to speak from inside the cell (intra cellular). A negative potential means at the membrane inside the cell there is a negative charge. A positive potential means that at the membrane inside the cell there is a positive charge. We can, for the sake of understanding, simply invert this and say a negative potential means that at the membrane OUTSIDE the cell has a positive charge; and that a positive potential means that at the membrane OUTSIDE the cell has a negative charge.

When an ion moves across a membrane it carries a charge and creates an electrical potential. An important point is that the electrical potential created at the membrane is not created by the concentration gradient, but from the movement of ions which is based on the concentration gradients. The Na^+/K^+ ATPase pumps work at a rate in response to the concentration gradients (speed up if there is a need to, slow down to resting levels when able). But even for long periods of high frequency excitations there is no risk of significant change to the ionic concentrations outside or inside the cell. Thinking that there is an absolute change in ionic concentration is required for an action potential is a common misconception about action potentials [Sil02].

Membrane Potentials Created by Primary Ions of Interest

- When K^+ moves across the membrane (out of the cell) it creates a negative potential.
- When Na^+ moves across the membrane (into the cell) it creates a positive potential.
- When Cl^- moves across the membrane (into the cell) it creates a negative potential.

To generalize these facts. When a positive ion moves into the cell it creates a positive potential (+in+); when it moves out of the cell it creates a negative potential (+out-). When a negative ion moves into a cell it creates a negative potential (-in-); when it moves out of a cell it creates a positive potential (-out+). Notice that when an ion moves "in" there is a direct relationship between the charge of the ion and its impact on the membrane potential. But when an ion moves out of a cell there is an indirect (or inverse) relationship between the charge of the ion and its impact on the membrane potential. It is always important to remember that the membrane potentials are based on the movement of the ions, not on their concentration.

The electrical potential created at the membrane by the movement of an ion has an influence on the movement of ions. The negative electrical potential created at the membrane by the movement of K^+ out of the cell will eventually resist the continued movement of K^+ out of the cell. How negative that negative potential needs to be to stop the movement of K^+ out of the cell is dependent on the concentration gradient and how permeable the membrane is to K^+ .³ The electrical potential that stops the movement of an ion is called the **equilibrium potential**. Figure 4.9 shows the concentration gradient and resultant equilibrium potential for our primary ions of interest (Na^+ , K^+ , Cl^-). Each equilibrium potential

³ It is also dependent on the temperature (which is kept within tight boundaries in the body), Faraday's constant, and the universal gas constant. From these variables the potential generated by any ion can be determined with the Nernst Equation, which is not covered here at this time.

in 4.9 is based on the assumption that it is the only ion moving across the membrane. In reality they are all moving across the membrane with varying rates based on the permeability of the membrane to each ion at any moment in time.⁴

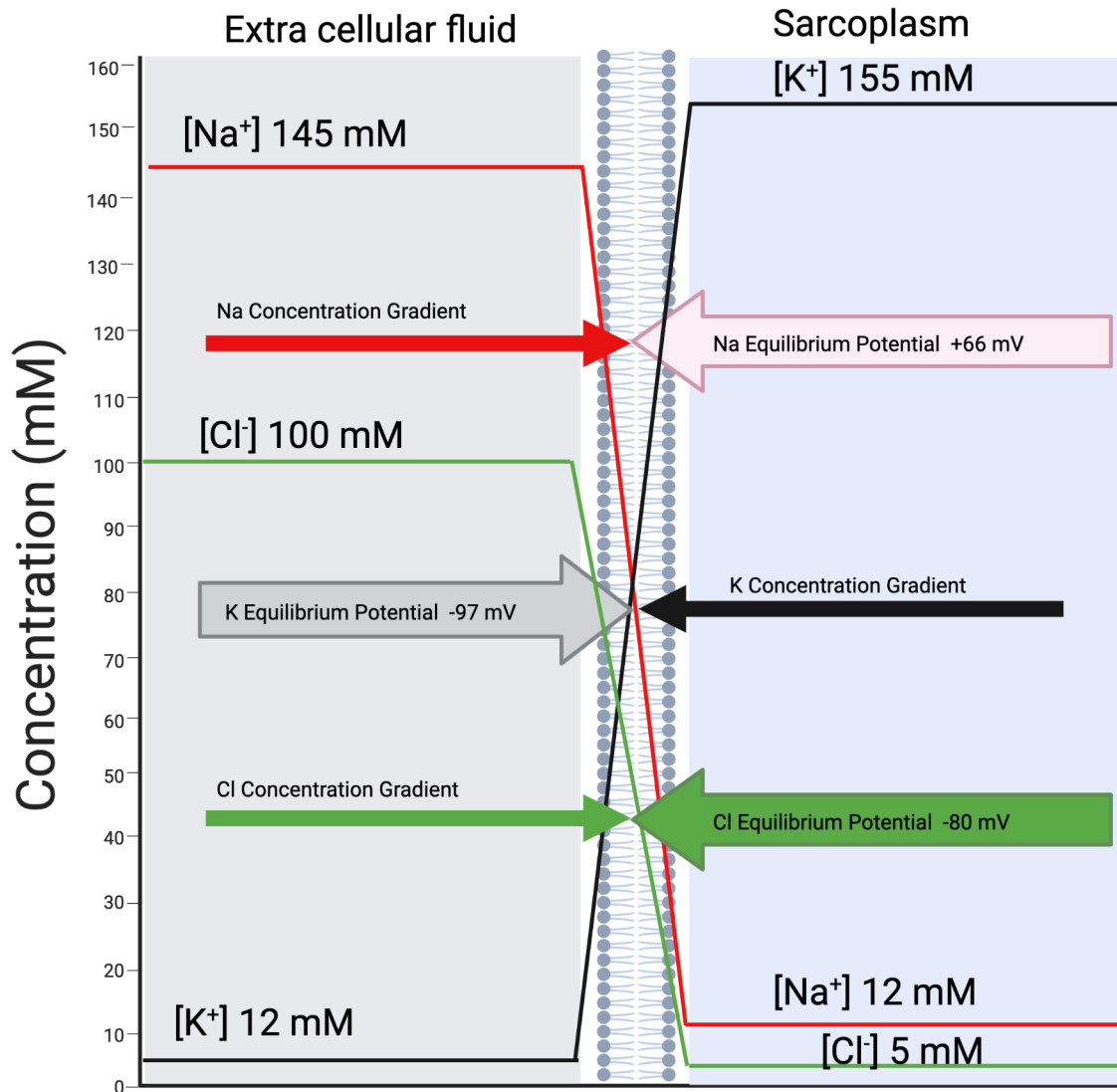


Fig. 4.9: Generation of Equilibrium Potentials (Created with BioRender.com)

The resting membrane potential is an equilibrium potential. It is the consequence of two physiological situations. First, the presence of large concentration gradients for K^+ , Na^+ and Cl^- across the membrane. Second, the relative permeability of the membrane to these ions. For example, based on Figure 4.9, if K^+ was the only ion that the membrane was permeable to at rest the RMP would be equal to the K^+ equilibrium potential (about -97 mV). The concentration gradients for K^+ and Na^+ are established primarily by the Na^+ - K^+ -ATPase pumps. The concentration gradient maintains a large outwardly directed K^+ gradient, and a large inwardly directed Na^+ gradient. The concentration gradient for Cl^- is maintained primarily through the actions of a K^+ - Cl^- co-transporter known as KCC2. Relative to the other ions, there is a moderate inwardly directed Cl^- gradient.

⁴ The Goldman-Hodgkin-Katz (GHK) Equation (model) allows the estimation of the membrane potential based on Na^+ , K^+ , Cl^- concentration gradients and membrane permeability (along with several other relevant variables). A calculator for the GHK Equation can be found at https://www.physiologyweb.com/calculators/ghk_equation_calculator.html.

Since these ions do not pass through the lipid bi-layer of the cell membrane the permeability of the membrane to each ion is tuned by the presence of ion specific channels. Therefore, the relative permeability of the plasma membrane to Na^+ , K^+ and Cl^- is dependent on the number and state (open or closed) of ion-selective membrane channels. Different membranes display different degrees of permeability to different ions, and due to gating mechanisms on many of the ion-selective channels the permeability can be changed. This is an important point considering the entire purpose of an RMP (polarization) is to have the opportunity to generate an AP (depolarization). In resting conditions the membrane of nerve and muscle cells are most permeable to K^+ . In addition to KCC2 allowing movement of K^+ out of the cell, there are also non-gated K^+ ion channels. When combined with the relatively high permeability of K^+ , the concentration gradient results in a K^+ flow out of the cell. As K^+ flows out it creates a negative electrical potential (+out-). However, the negative electrical potential resists further flow of K^+ outside of the cell (and eventually reaches an equilibrium potential). Since K^+ has the highest permeability at rest, the resting membrane potential is primarily influenced by the equilibrium potential of K^+ . It is important to note that while there is no net flow of K^+ at the K^+ equilibrium potential, K^+ is still flowing out of the cell and continues to generate the negative membrane (equilibrium) potential. K^+ also continues to be pumped into the cell by the Na^+/K^+ ATPase pumps. The term "equilibrium" for equilibrium potential means equilibrium in flows (no net flow, no net gain in or outside the cell).

Note from Figure 4.9 that the equilibrium potential of Cl^- is close to K^+ . The implication is that Cl^- , despite its concentration gradient encouraging inward flow, does not at RMP, have net inward flow. Cl^- has a higher permeability than Na^+ at rest, but a much lower permeability than K^+ . An important attribute about Cl^- for the sarcolemma and axon membranes is that the permeability of the membrane does not change much, even during an AP. This contributes to the stability of the RMP and helps prevent recurrent self excitation in these membranes.

The equilibrium potential of Na^+ is positive (opposite of K^+ and Cl^-). At RMP the permeability for Na^+ is very low. Therefore Na^+ has very little impact on the RMP. The importance of Na^+ starts when ligand and then voltage gated Na^+ channels start to open - that's when the action begins.

4.2.3 Trigger & Spread of Sarcolemma Excitation (AP)

In Figure 4.9 the equilibrium potential for Na^+ is positive. When the membrane permeability for Na^+ increases (becomes greater than K^+ and Cl^-) the membrane potential starts to depolarize and even to switch to a positive polarity. This is an action potential. At the molecular level, an action potential is the sudden increase in permeability of the membrane to Na^+ .

4.2.3.1 Triggering the Action Potential (Initial Excitation)

The events leading to an action potential on the sarcolemma are initiated by the opening of ligand gated Na^+ channels by the binding of ACh to these receptors at the motor end plate. The micro potentials created by the opening of the ligand gated Na^+ channels then open voltage gated Na^+ channels due to a change in voltage in the membrane potential (becoming less negative). The number of voltage gated Na^+ channels that open is proportional to the change in the membrane potential. This cycle of opening in response to a voltage change is self perpetuating since once a Na^+ channel opens Na^+ flows into the cell due to the Na^+ concentration gradient. Due to the movement (not a change in the actual gradient) the membrane potential voltage changes (becomes less negative) as the potential moves toward the Na^+ equilibrium potential. The point that approximately half of the voltage gated Na^+ channels open is considered a threshold potential. At the threshold potential the voltage gated Na^+ channels continue to rapidly open until the equilibrium potential approaches that of Na^+ (as opposed to K^+). At this point the membrane is more permeable to Na^+ . The equilibrium potential for Na^+ is approximately 66 mV. Once a membrane reaches the threshold potential an action potential is inevitable. It is at this point that the action potential is said to be "all or none." Prior to the threshold potential membranes can undergo micro potential changes. These micro potentials can be depolarizing (excitatory potentials that move the membrane potential toward the threshold), or hyperpolarizing (inhibitory potentials that move the membrane potential further away from the threshold potential). The use of micro potentials to regulate muscle tension is an important part of Chapter 5 on Regulation since it is a process that occurs at the α -motor neuron dendrites in the spinal cord. It is not a process that occurs at the NMJ.

A peak ahead to an important concept

An α -motor neuron, once excited, will release only one neurotransmitter, ACh at its axon terminal. ACh will only have one effect at the NMJ, creating excitatory micro-potentials on the motor end plate. However, in the Central Nervous System (such as the Spinal Cord) competing neurotransmitters can be released at synapses of the α -motor neuron dendrites. The impact of these neurotransmitters can be excitatory or inhibitory. Since they occur after the synapse they are called excitatory post synaptic potentiations (EPSPs) or inhibitory post synaptic potentiations (IPSPs). Whether, and how frequently, an α -motor neuron sends an excitation to the muscle fibers it innervates is balanced by competition between the EPSPs and the IPSPs.

4.2.3.2 Spread of the Action Potential (Wave of Excitation)

The wave of excitation of the AP across the sarcolemma and T-tubules occurs due to the self-perpetuating opening of voltage gated Na^+ channels. When a region of the membrane depolarizes it opens the nearby voltage gated Na^+ channels which depolarize that region, which then opens the nearby voltage gated Na^+ channels which depolarize that region, "And so it goes."⁵

Of course, there cannot be an AP if there is no RMP. Therefore, the AP must end. The membrane must repolarize to be ready for another depolarization. Considering muscle tension is, partly, regulated by how frequently it experiences excitation (which can reach upwards of 180 Hz, that is 180 excitations per second). For those excitations to occur there need to be an equal number of repolarizations.

The process of repolarization occurs due to two mechanisms. The first mechanism is based on voltage gated Na^+ channels having two gates. One is an activation gate and it is closed under resting conditions and opens during excitation. The second is an inactivation gate and it is open under resting conditions but closes soon after the activation gate opens. The inactivation gate closure is timed, it is not triggered by events outside of the molecule itself. Closing the inactivation gate is the first step to stopping an action potential. The closure of the inactivation gates reduces the membrane permeability for Na^+ . While the activation gate is open and the inactivation gate is closed the membrane is in the absolute refractory period. This means it cannot undergo depolarization regardless of how great an excitation stimulus may be. To be reset and completely ready for another action potential the activation gate must close and the inactivation gate must open. These are also timed events.

The second mechanism to repolarization is opening of voltage gated K^+ channels. These are delayed (take longer to open than the voltage gated Na^+ channels). Once the voltage gated K^+ channels open the membrane permeability to K^+ increases higher than under resting conditions which expedites the repolarization process and results in a hyper polarized state (slightly more negative membrane potential than the normal resting membrane potential). Closing the voltage gated K^+ channels is also a timed event. While they are open the membrane is in a relative refractory period. It is resistant to excitation, however, with an increased stimulus it can be excited.

The overall process of repolarization and return to the resting membrane potential takes approximately 3-4 milliseconds (ms). Assuming 3 ms is the fastest interval for another AP, the upper limit on excitations would be approximately 333 Hz (which exceeds the maximum capacity for human muscle fiber twitches per second).

Spread of excitation occurs in one direction (meaning it does not result in cyclic repeating action potentials of the entire membrane) because of the refractory periods.

Domino Model of the Unidirectional Spread of Excitation

Visualize the membrane as a set of dominoes. The resting potential is when the dominoes are upright. An action potential is when a domino falls. Then the refractory period is the period when the domino is still fallen. Because of the fallen domino on one side of the next falling domino, the dominoes fall in one direction away from the starting stimulus.

A further safeguard against cyclic repeating action potentials in the sarcolemma is the constant permeability for Cl^- in the sarcolemma. Recall that at RMP Cl^- has no net flow because its equilibrium potential is near the K^+ potential. When the membrane potential fluctuates away from the K^+ (and Cl^-) equilibrium potential there is net movement of Cl^- . This

⁵ A quote for other Kurt Vonnegut fans.

Cl^- net movement (into the cell) contributes to repolarization but also to stability of the RMP in the moments following an AP. In situations such as myotonia congenita, altered membrane Cl^- permeability leads to a higher occurrence of myotonia (sustained muscle tetany) due to recurrent muscle excitation from cyclic repeating action potentials [AM76].

4.2.4 Summary of Excitable Membranes & Considerations for Regulation

This section has taken a deeper dive into the molecular events of excitation than the previous section. It should be clear that many events must occur for one excitation, particularly since one excitation creates a twitch, and a twitch is not enough to produce meaningful movement. A more efficient system of excitation-activation coupling exists in the cardiac muscle. However, in the skeletal muscle the efficiency of creating tetany from one excitation is sacrificed for the sake of regulation of muscle tension. Since creating tetany requires many action potentials, then the amount of tetany, and therefore the amount of tension, can be regulated by the number of action potentials. The regulation of muscle fibers is the topic for Chapter 5.

4.3 Clinical Physiology Connections

The concepts of excitable membranes, receptors, ligand channels and signal transduction are pervasive throughout clinical physiology. This final section of the chapter introduces some clinical physiology connections that are now possible given this understanding of excitable membranes. Topics include the neuroendocrine system, sensory receptors and pharmacodynamics. These topics return throughout the book in particular contexts.

4.3.1 Neuroendocrine

The neuroendocrine system regulates several critical physiological systems and coordinates cellular functions throughout the body. Examples include mobilizing the body for fight or flight reactions by getting all cells in the body ready to meet demands that require energy mobilization at the expense of other maintenance activities. For example, release of hormones for fight or flight signals the liver cells to hold off on using energy to maintain the cell membrane and instead use energy to convert glycogen (a stored fuel) into glucose (a usable fuel). The neuroendocrine system can alert cells all over the body to do activities that each of those cells can do for the fight or flight situation, even though those activities are varied between the cells. While the liver cells are releasing glucose, the heart cells are prepared to use more glucose and are working more frequently (higher rate) and harder (more pressure). The differentiation and variety of tasks completed by these cells is due to the specialization of the cells, and due to the specific receptors around the body that are responding to the same ligand. Most cells in the body have a ligand receptor for epinephrine (a major catecholamine of the fight or flight response). Epinephrine is also called adrenaline. Receptors for epinephrine are therefore called adrenergic receptors. The adrenergic receptors on all cells have a similar binding location, but the receptor itself transduces different signals into the cell depending on the receptor and depending on the cell. For example, the adrenergic receptors on smooth muscle in arterioles in the systemic circulation signal contraction which results in vaso-constriction. However, the adrenergic receptors on smooth muscle in bronchioles in the lung signal relaxation, which results in broncho-dilation. The fight or flight (alert) response from the neuroendocrine system is a coordinated response of the component of the central nervous system called the autonomic nervous system, and in particular the sympathetic nervous system (SNS). The SNS is a truly neuroendocrine system because it has both nervous system and endocrine components. Ligands for the SNS, epinephrine and norepinephrine, are both neurotransmitters and hormones.

For homeostasis (balance), if there is a fight or flight (alert) system then there should also be a rest, digest and reproduce (chill) system. The chill system is the parasympathetic system (PNS). The PNS is a nervous (but not neuroendocrine) system. The connections through the body of the PNS are all mediated with nerve function and neurotransmitters (not endocrine, or hormone signals). This is not to say that during a chill situation hormones are not involved. Just that if there are pathways through which the PNS mediates those chill (rest, digest, reproduction) hormones they have not

been discovered yet (at least as far as the author knows as of May 22, 2024). The primary neurotransmitter for the PNs is acetylcholine (ACh). The PNS receptors that bind ACh are called cholinergic receptors.

4.3.2 Sensory Receptors

Sensory receptors are specialized receptors at the terminal end of the axon of a sensory neuron. They differ from ligand and voltage gated receptors in the signals that provokes their excitation. When a sensory receptor is excited it sends an afferent signal (to the central nervous system, as opposed to an efferent signal, away from (or exiting) the central nervous system). The signal that provokes excitation in a sensory neuron is whatever phenomenon or sensation that sensory neuron is associated. If the sensory neuron is excited by changes in temperature then it is a temperature sensory neuron. If by changes in tension, it is a tension sensory receptor. If by vibrations, it is a vibration (as in the ear) sensory receptor. If by photons, it is a light sensory receptor (as in the retina). If by certain chemicals or ions (such as H^+ for pH, or carbon dioxide, or oxygen) they are chemoreceptors. If by changes in pressure they are baroreceptors.

Some sensory receptors terminate in areas of the brain that allow perceptions of the signal in a way that gives them additional meaning. For example, some chemoreceptors proceed to the brain and produce pain signals. Other chemoreceptors that detect an increase in carbon dioxide proceed to areas in the nervous system to produce unconscious reflex responses (changing respiratory rate) and also to the brain (a general state of anxiety and air hunger). Some receptors don't go to the brain at all, but interact at areas in the nervous system where responses are automatic (baroreceptors as part of the autonomic nervous system to regulate blood pressure). Since baroreceptors don't go beyond the brain stem we do not perceive what our blood pressure is directly - making elevated blood pressure a "silent killer." To be a homeostatic regulatory system there needs to be a sensory receptor of some type that contributes to the monitoring of the homeostatic variable. A homeostatic regulatory system includes a homeostatic variable that is sensed by a sensory receptor that sends an afferent signal to be processed (often not perceived), that then sends an efferent signal to exert an action which aims to adjust the homeostatic variable. The neuroendocrine system is commonly involved in such systems, receiving afferent signals, and sending efferent signals. Most physiological homeostatic systems are negative feedback systems, which means sensation of an increase in a variable triggers response to decreases that same variable (and vice versa).

4.3.2.1 Sensory Receptor Dynamics

Sensory receptors are proteins embedded in cell membranes. Therefore, the cell has control over some dynamic and important behavior of these receptors (indeed, all receptors). **Up-regulation** is a term for sensory receptor dynamics that includes an increase in the number of receptors. **Down-regulation** is a term for sensory receptor dynamics that includes a decrease in the number of receptors. With up-regulation or down regulation there are changes to the strength of a excitation and therefore signal sent to the respective neurons. For example, with up-regulation of chemoreceptors for carbon dioxide an individual will have a higher respiratory rate and maintain a lower carbon dioxide level since their response to carbon dioxide is increased. The up-regulation then continues to propagate itself. Since it results in lower carbon dioxide levels, the receptors remain up-regulated in an attempt to be able to respond to the lower carbon dioxide levels. A balance point is eventually achieved. But this new balance point may be enough to result in what is considered altered homeostasis. In individuals taking a beta-blocker (beta-adrenergic blocker medication) the lack of cardiac response to neuroendocrine stimulus results in higher adrenaline levels. The adrenergic receptors of the heart down-regulate in response to higher circulating adrenaline which further blunts cardiac responsiveness to adrenergic stimulation. A balance point is eventually achieved as with the previous example, however, the individual has an alteration to homeostasis.

4.3.2.2 Neuromuscular Sensory Receptors

There are several important neuromuscular sensory receptors that sense signals from the muscle, musculotendonous junction and/or the tendon. Signals for tension (such as golgi tendon organs), signals for changes in length (such as

the muscle spindle), and signals for the chemical constituents of the extra cellular fluid surrounding muscle fibers (chemoreceptors). These receptors and their role in muscle function is covered in upcoming chapters.

4.3.3 Pharmacodynamics

Pharmacology is divided into pharmacotherapeutics and toxicology. Toxicology focuses on the harmful effects that chemicals, such as medications, may have on the body. Pharmacotherapeutics is the development and study of medications. Pharmacotherapeutics is divided into pharmacokinetics and pharmacodynamics. Pharmacokinetics focuses on the absorption, distribution and metabolism of medications. Pharmacodynamics focuses on the intended systemic and cellular effects of medications. Pharmacodynamics is therefore a study of the actions of medications in the body. The cellular and systemic effects of a medication is often related to what receptors it binds, what neurotransmitter or hormone it effects of mimics, or what signal transduction mechanism it influences. Pharmacodynamics often includes changes to the process of excitation discussed in this chapter.

4.3.3.1 Curare

Curare is paralyzing agent. It stops the excitation of the motor end plate. It does so by competitively and reversibly binding to ACh receptors on the motor end plate which blocks those receptors from binding with ACh. When curare binds to an ACh receptor it does not open the ligand gated Na^+ channel and therefore prevents the initiation of depolarization. As a competitive binding, curare blocks the ACh receptor and thus inhibits the ACh binding and response. Another term for competitive binding is that curare is an antagonist (blocks without producing an effect). Some medications are agonists, they bind and create an effect and are thus utilized to facilitate additional cellular activity beyond that of the normal ligand. As a reversible binding medication, curare can be removed from the ACh receptor. Medications can also create irreversible binding. Curare is also selective, it only binds to skeletal muscle motor end plate ACh receptors. Which means it selectively inhibits these receptors and not other ACh receptors throughout the body. This means the systemic response is limited to skeletal muscle paralysis. If used during surgery without other anaesthetics it would produce paralysis but no change in pain sensation, a very unpleasant surgery that would be stopped when the pain and anxiety of the procedure produced a rapid heart rate and rising blood pressure due to the fight or flight response of the neuroendocrine system (which would not be stopped by curare). In addition to surgical or intensive care situations, curare can be used as a poison, causing death by asphyxiation since the muscles of ventilation (breathing) are skeletal muscles.

4.3.3.2 Medication Induced Receptor Dynamics

A consequence of medications is that they can produce changes in receptors. Up-regulation and down-regulation of receptors due to their pharmacodynamics are a common reason for the need to carefully, and only under supervision, withdraw a dose of medications. If the medication led to up or down regulation and depending on whether it is a receptor antagonist or agonist, could have big implications with sudden cessation. For example, if a medication leads to down regulation of receptors, cessation of that medication could result in a large drop in cellular activity (whatever that activity may be). For example, serotonin re-uptake inhibitors are antagonists that block the re-uptake of serotonin in synapses of the central nervous system. This increases the amount of serotonin which can lead to down-regulation of the serotonin receptors. If the medication is suddenly stopped there will be less serotonin and less receptors (due to down-regulation) which magnifies the effect of having less serotonin, potentially worsening of the original reasons for taking the medication.

4.3.3.3 Autonomic Nervous System Medications

There are several medications that influence the autonomic nervous system function for a variety of health conditions. Applying what we know about receptors and pharmacodynamics it should be clear that an adrenergic agonist will increase SNS activity in the cell that the medication binds to, whereas an adrenergic antagonist (blocker) will reduce

the SNS activity in the cell that the medication binds. A cholinergic agonist will increase the PNS activity in the cell, whereas a cholinergic antagonist will decrease the PNS activity in the cell. Very simple. Two types of receptors that can effect the receptors in two ways and therefore can create four different cellular responses.

Thankfully there is enough cholinergic receptor specificity that a cholinergic agonist with the therapeutic goal of changing gut motility is specific to the cholinergic receptors in the gut. If all cholinergic receptors were the same, then such a cholinergic agonist could cause wide spread muscle spasms.

4.4 Summary

Muscle fiber excitation starts with an excited motor axon (α -motor neuron) and ends with the binding of calcium to troponin, which generates a twitch (the fundamental unit of active tension). Binding calcium to troponin allows a crossbridge to go from an inactivated to an activated state. The final step in muscle excitation is the coupling between excitation and activation (Excitation - Activation Coupling). The underlying molecular mechanisms involve the characteristics and actions of excitable membranes. Excitable membranes require ion channels, pumps and receptors. Actions include using transport to establish a resting membrane potential and the ability to have an action potential (excitation). Understanding excitable membranes creates Clinical Physiology Connections opportunities in three areas: 1. system wide regulatory function of the neuroendocrine system; 2. function of sensory receptors; and 3. pharmacodynamics.

A twitch is the fundamental unit of muscle fiber active tension. A twitch is activated by excitation of the muscle fiber by the α -motor neuron. Excitation of the muscle fiber is the fundamental unit of regulating active tension in the muscle fiber. The next chapter considers regulation as the process of varying muscle fiber excitation - both the frequency of excitation as and the number of muscle fibers excited.

References

- [AM76] R H Adrian and M W Marshall. *Action potentials reconstructed in normal and myotonic muscle fibres*. en. The Journal of Physiology 258.1 (June 1976), pp. 125–143. ISSN: 00223751. DOI: [10.1113/jphysiol.1976.sp011410](https://doi.org/10.1113/jphysiol.1976.sp011410). URL: <https://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1976.sp011410> (visited on 06/01/2022).
- [Sil02] Dee U. Silverthorn. *Uncovering misconceptions about the resting membrane potential*. en. Advances in Physiology Education 26.2 (June 2002), pp. 69–71. ISSN: 1043-4046, 1522-1229. DOI: [10.1152/advan.00012.2002](https://doi.org/10.1152/advan.00012.2002). URL: <https://www.physiology.org/doi/10.1152/advan.00012.2002> (visited on 06/01/2022).

Sample Questions

1. Up regulation of ligand gated ion channels on a motor endplate would result in:
2. Rapid upslope of the membrane potential from the resting membrane potential to initiate the action potential is facilitated by which concentration gradient flow.
3. A fascicle is surrounded by?
4. What is the correct order of events for muscle fiber excitation?
5. What occurs during the rapid recovery of the RMP following an action potential?
6. How does elevated extracellular potassium levels with normal intracellular potassium levels and no change in resting membrane potassium permeability change resting membrane potential?
7. What is the effect of an increased resting membrane potential with no change in the threshold potential?

Chapter 5

Regulation

Updated on May 22, 2024

Contents

5.1	Motor Units	62
5.2	Motor Unit Twitch & Tetany	67
5.3	Active Tension Regulation	69
5.4	Regulatory Feedback	72
5.5	Clinical Physiology Connections	73
5.6	Summary	75

This chapter introduces the mechanisms utilized by the central nervous system (CNS) to regulate active tension. To regulate active tension means to vary its force from its lowest value (resting tone) to its maximal value (maximal voluntary contraction (MVC)). Active tension is developed at the level of the sarcomere, excitation at the level of the muscle fiber, regulation occurs at the level of motor units. A motor unit includes an α -motor neuron and all of the muscle fibers it innervates (can excite). The CNS has two strategies to regulate active tension. One strategy is to manipulate the number of twitches per second (frequency summation or rate coding); and the second strategy is to manipulate the number of motor units twitching (motor unit summation or motor unit recruitment). These two approaches allow for nearly continuous variation in active tension possible. Some muscles (fine motor) have a greater capacity for continuous variation than others (gross motor).

Objectives include:

1. Explain motor unit structure–function relationships.
2. Explain muscle fiber differentiation (types) structure-function relationships.
3. Explain motor unit excitation, twitch and tetany.
4. Compare and contrast frequency summation (rate coding) and motor unit summation (recruitment).
5. Explain the order of recruitment.
6. Explain how the muscle spindles and the golgi tendon organs influence motor unit excitation.
7. Explain the physiological basis of the electromyogram (EMG) and limitations of EMG interpretation.
8. Explain the effect of amyotrophic lateral sclerosis and aging on motor units.
9. Explain the effect of peripheral nerve demyelination on muscle regulation.

5.1 Motor Units

A motor unit is an α -motor neuron and all of muscle fibers that its terminal axon branches form a neuromuscular junction at a motor end plate. Each muscle fiber has one motor end plate, and therefore one neuromuscular junction. Each α -motor neuron has many terminal axon branches. The muscle fibers of a single motor unit are randomly dispersed throughout the muscle epimysium. Therefore, the muscle fibers of a motor unit contribute to the tension developed in several fascicles. This arrangement ensures that when only a few motor units are excited there are muscle fibers throughout the entire muscle generating tension at the tendon attachments to avoid asymmetrical pull from one fascicle or isolated location within the epimysium.

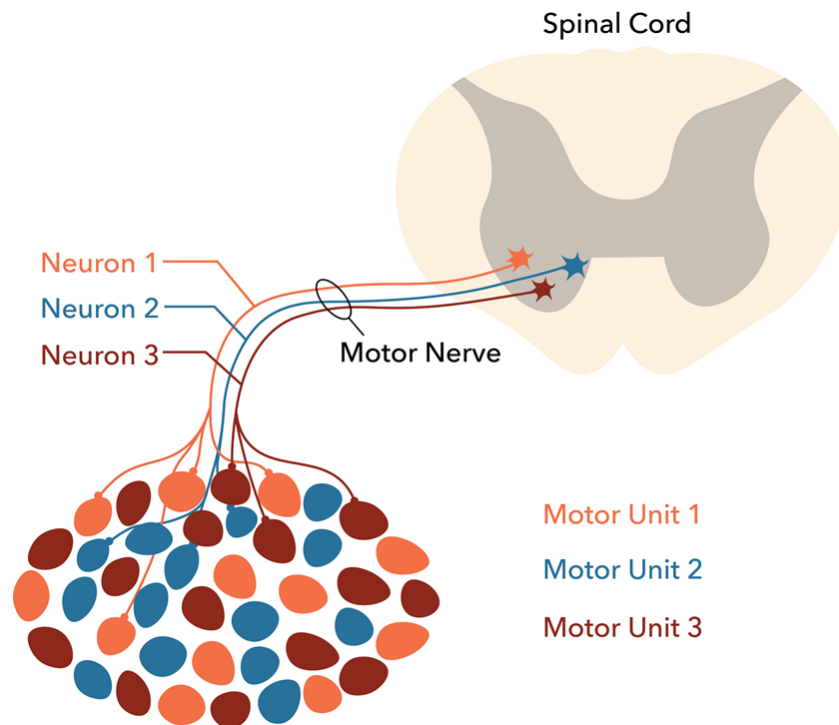


Fig. 5.1: Motor Unit Distribution in a Muscle ([Wikimedia Commons](#), CC BY 4.0)

5.1.1 Innervation Ratio

Motor unit innervation ratio (IR) refers to the number of muscle fibers per motor unit and is usually given as the average number of muscle fibers per motor unit.

$$IR = \frac{N_{mf}}{N_{mu}} \quad (5.1)$$

The IR varies between muscles and differentiates muscles equipped for fine or gross motor control (See Table 5.1). IRs vary from 5 in muscles in the eye (i.e. lateral rectus) to approximately 1700 in muscles in the calf for locomotion (i.e. gastrocnemius).

IRs for whole muscles underlie the concept of fine and gross motor control. Tension is developed by muscle fibers, the number of muscle fibers recruited with each new motor unit recruited influences the size of jumps in active tension. Muscles with lower IRs have greater regulation (control) over the amount of tension they produce. They better accommodate continuous increments in tension with each new motor unit recruited. Of five muscles in Table 5.1, the

Muscle	α -motor neurons	Muscle Fibers ($\times 10^3$)	Innervation Ratio
Lateral Rectus (eye)	30*	.15*	5
Lumbricals (hand)	98	10.3	105
Brachialis	330	130	393
Biceps	774	580	749
Gastrocnemius	580	1000	1724

Table 5.1: Variability of Innervation Ratio in Human Muscles (From [BS80], (* Estimates that need confirmation))

lateral rectus (eye) and lumbricals (muscle in the hand) have the greatest fine motor control. While the gastrocnemius is responsible for the greatest need for high tension (force) output (gross motor control).

5.1.2 Motor Unit & Muscle Fiber Types

While IRs for whole muscles help explain variation between muscles with differences in motor control they ignore the systematic variation in the number of muscle fibers per motor unit within a muscle. There are different types of motor units. The types include differences in the α -motor neurons, the muscle fibers and the IR. The differences are structural and functional. Types exist on a spectrum with transitions in many (but not all) of the characteristics between them. Despite the spectrum, and using different characteristics, experts have identified and agree on classification systems that converge on three types of human motor units and associated muscle fibers [Lie10].

- Fast Fatigable (FF) Motor Units with Fast Glycolytic (FG) (Type 2a) Muscle Fibers
- Fast Resistant (FR) Motor Units with Fast Oxidative (FOG) (Type 2x) Muscle Fibers
- Slow (S) Motor Units with Slow Oxidative (SO) (Type 1) Muscle Fibers

The above classification names are all utilized, however the Type 1, 2x and 2a are probably the most common. The other classifications are informative about the functional characteristics of the motor units and muscle fibers they innervate. Beyond their coherence, the motor unit classifications FF, FR and S share a feature with the muscle fiber classifications FG, FOG and SO. They both use the letters F and S to represent Fast and Slow, and they primarily refer to faster or slower excitation (and in the case of muscle fibers also activation).

The second letter F in FF, and the letters R and S (dual use of the meaning slow for the Slow motor units), and the letters G and O for muscle fibers are related to functional characteristics of fatigue and energetics. For motor unit classifications the second F in FF stands for fatigable, meaning FF fibers are fatigable and cannot sustain tetany for very long. The R in FR motor units stands for resistant, as in fatigue resistant. They resist (for a time) the fatigue of FF motor units. And the S motor units are slow, both in excitation, and slow to fatigue (hence the "dual use" of slow). In Chapter 6 on Energetics, the concept of fatigue denoted in muscle fiber classification letters G and O refer to the biochemical pathways that produce ATP, glycolysis and oxidative (including both the citric acid (Kreb's) cycle and electron transport). The connection between FF and FG, FR and FOG, and S with SO, has to do with the rate and sustainability of ATP production by these pathways. Glycolysis has a relatively high rate and low sustainability, and oxidative pathways have a relatively low rate and high sustainability. These pathways, but not just these pathways, result in the fatigability, fatigue resistant or slow to fatigue characterizations of motor units and their fibers. While fatigability comes up in this chapter, the focus is on the excitation and activation characteristics of the fast and slow motor units and muscle fibers.

5.1.2.1 Fast vs. Slow - An excitable concept

The terms fast and slow refer to relative differences in functional characteristics between α -motor neuron and the muscle fibers. The functional characteristics fast and slow are associated with a set of functionally coherent structural differences. The section below focuses on the most distinguishing structural characteristics that influence the functional characteristics of excitation and activation.

5.1.2.2 Fast & Slow α -Motor Neurons

Fast and slow for the motor unit α -motor neuron refers to how quickly excitations can be sent which influences how frequently they can be excited. The faster an excitation propagates down an axon the sooner it is ready for another excitation. In a sequence of excitations, if a second excitation catches up to the first excitation it will cease propagation once it reaches a section of the axon membrane that is its refractory period from the first excitation. A fast neuron has a higher nerve conduction velocity (NCV) and a higher maximum excitation frequency (Hz) than a slow neuron. NCV is related to the size of the neuron (diameter) because diameter influences resistance to current. Current is the flow of electrical impulses and therefore directly related to velocity. Diameter is inversely proportional to resistance, and resistance is inversely proportional to current. Therefore diameter is directly proportional to current and velocity (See Figure 5.2). The large diameter of the α -motor neurons, combined with the myelin sheaths (The role of myelin was introduced in Chapter 4), contribute to its relatively high NCV.

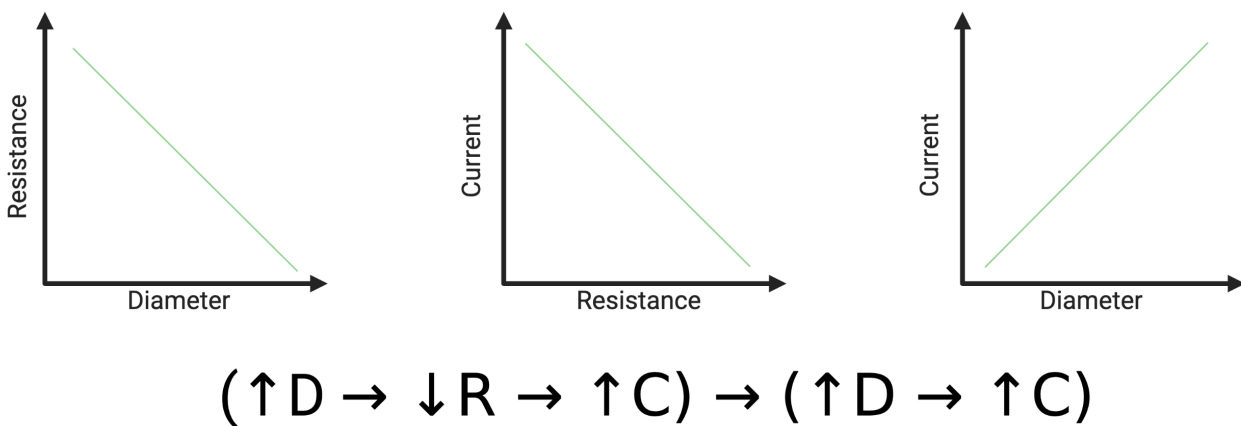


Fig. 5.2: Relationship between Diameter, Resistance and Current (Created in BioRender.com)

Estimated NCVs for three nerve fiber types are presented in Table 5.2. The α -motor neuron has the highest NCV. The variation in diameter in the α -motor neuron is explained by the motor unit types with the fast (FF) motor units innervating the FG fibers having the larger diameter ($\approx 22\mu m$) on the higher end for NCV ($\approx 120 m \cdot s^{-1}$), and the S motor units with the small diameters ($\approx 12\mu m$) and the lower end of NCV ($\approx 70 m \cdot s^{-1}$) (and FR motor units innervating FOG somewhere in the middle). The terms fast and slow for the α -motor neurons of motor units are relative to one another. It is clear in Table 5.1 that both fast and slow α -motor neurons are faster than the other neurons listed. They are faster due to both diameter (as compared to δ -sensory neurons) and additionally myelination (as compared to C sensory neurons).

Nerve Fiber Diameter (μm)		NCV ($m \cdot s^{-1}$)	Function
α -motor	12-22	70-120	motor
δ -sensory	1-5	12-30	sharp pain
C	0.5-1.2	0.2-2	dull pain

Table 5.2: NCV for different nerve fiber types. Note: C fibers are unmyelinated, whereas α -motor and δ -sensory fibers are myelinated. Therefore, C fibers have a larger drop in NCV than would be estimated just from diameter. Data from [Feh17]

A fast α -motor neuron has a larger diameter and myelin (structure) which both contribute to the NCV (function). Since the neurons of slow motor units are also myelinated the primary structural differentiation between fast and slow motor unit neurons are their size (larger diameters in fast motor unit neurons). Structurally, these larger neurons also innervate a higher number of muscle fibers. In mammalian muscles the IR for FF motor units is approximately twice (2x) that of S motor units, but this is an estimate that varies considerably between muscles (range from 1.1x to 3x greater than S motor units) [Bod+87].

A consequence of a larger neuron is that it requires more CNS excitation to become excited (excitation inertia). Excitation inertia has to do with how many ligand voltage gated channels must be opened before the threshold potential is reached. The larger (and thus faster) neuron has more excitation inertia. Fast α -motor neurons are larger. While fast, they take more excitation to be excited. This is an important feature for motor unit recruitment and the size principle of recruitment order covered later in the chapter.

A note on motor unit and muscle fiber type

Fast (larger) and slow (smaller) α -motor neurons play a role for motor unit and muscle fiber type. The narrative of this chapter is that there is coherence between the neuron doing the excitation, and the muscle fibers being excited so that they can be activated. The neuron excitations play a causal role in determining the type of muscle fiber [BS80]. It is not a coincidence that faster and larger motor neurons innervate faster and larger muscle fibers. Studies using electrical stimulation and crossover (changing which neurons innervate which muscles), demonstrate many muscle fiber characteristics change from slow to fast or from fast to slow. These experiments conclude that there is plasticity at the level of muscle fiber type with such manipulations. This plasticity is related to the differences in electrical (not chemical) stimulus. The challenge to anyone wishing to alter their fiber type dominance is finding a way to tap into the plasticity of the neurons. Certain characteristics such as myosin and myosin ATPase isoforms rely on gene expression (one isoform is expressed or the other is expressed) and therefore may provide a limit to how much, or how many, muscle fibers have this plasticity.

Summary of Fast & Slow α -Motor Neurons

The functional characteristics that make the fast α -motor neuron of the fast motor unit fast as compared to the slow classification are a higher NCV and a higher excitation frequency. The predominant structural difference that leads to these functional differences is the size of the neuron (larger, larger diameter). These structural and resultant functional differences fit on a spectrum that, in Table 5.2, result in a wide range of neuron diameters. These neuronal excitation characteristics innervate, and actually contribute to, coherence between the neuron and the muscle fibers of the motor unit.

5.1.2.3 Fast & Slow Muscle Fibers

The functional characteristics of muscle fibers that make them fast vs. slow include both structures related to excitation-activation coupling as well as structures related to activation (crossbridge kinetics).

For muscle fibers in a fast motor unit to respond to the faster and higher excitation frequency of the neurons they require a well developed T-tubule and sarcoplasmic reticulum (SR) system. The SR and T-system of fast fibers can have up to three times the volume of slow fibers. This allows fast motor units to convert the higher frequency of nerve excitation into a higher frequency of muscle fiber activation.

With a higher frequency of muscle fiber activation there must be structural changes that also allow for faster crossbridge kinetics, otherwise the faster activation (faster release and update of Ca^{2+} for example) would not result in faster twitches. The crossbridge kinetic structural changes tend to be discrete rather than on a spectrum. Discrete means that there are no intermediate or transition steps between the structure. This is compared to changes such as the size of a SR or the diameter of a neuron or the number of ACh receptors at a NMJ where there can any number of transitional steps.

Fast fibers express an isoform of the protein myosin (particularly the heavy chain) and the enzyme (also a protein) myosin ATPase (responsible for hydrolyzing ATP on the myosin head). The combination of these two expressions allow for faster crossbridge kinetics (cycling through the sliding filament model) which leads to slightly more tension at the level of the sarcomere and a higher velocity for shortening [LM93; SR96]. Therefore, when isolated or when either of these types of fibers dominate in a particular muscle, the force-velocity curve is shifted upward in fast as compared to slow motor units. Fast (FF) motor units having a higher velocity for any given force, and a higher force for any given velocity (Figure 5.3). These differences result in higher specific tension (T_s) in fast muscle fibers. T_s refers to the tension developed per sarcomere or, more commonly considered, tension developed per unit of cross-sectional area. There are more muscle fibers in fast motor units. Which means there are more sarcomeres and a greater cross sectional area. The T_s makes the

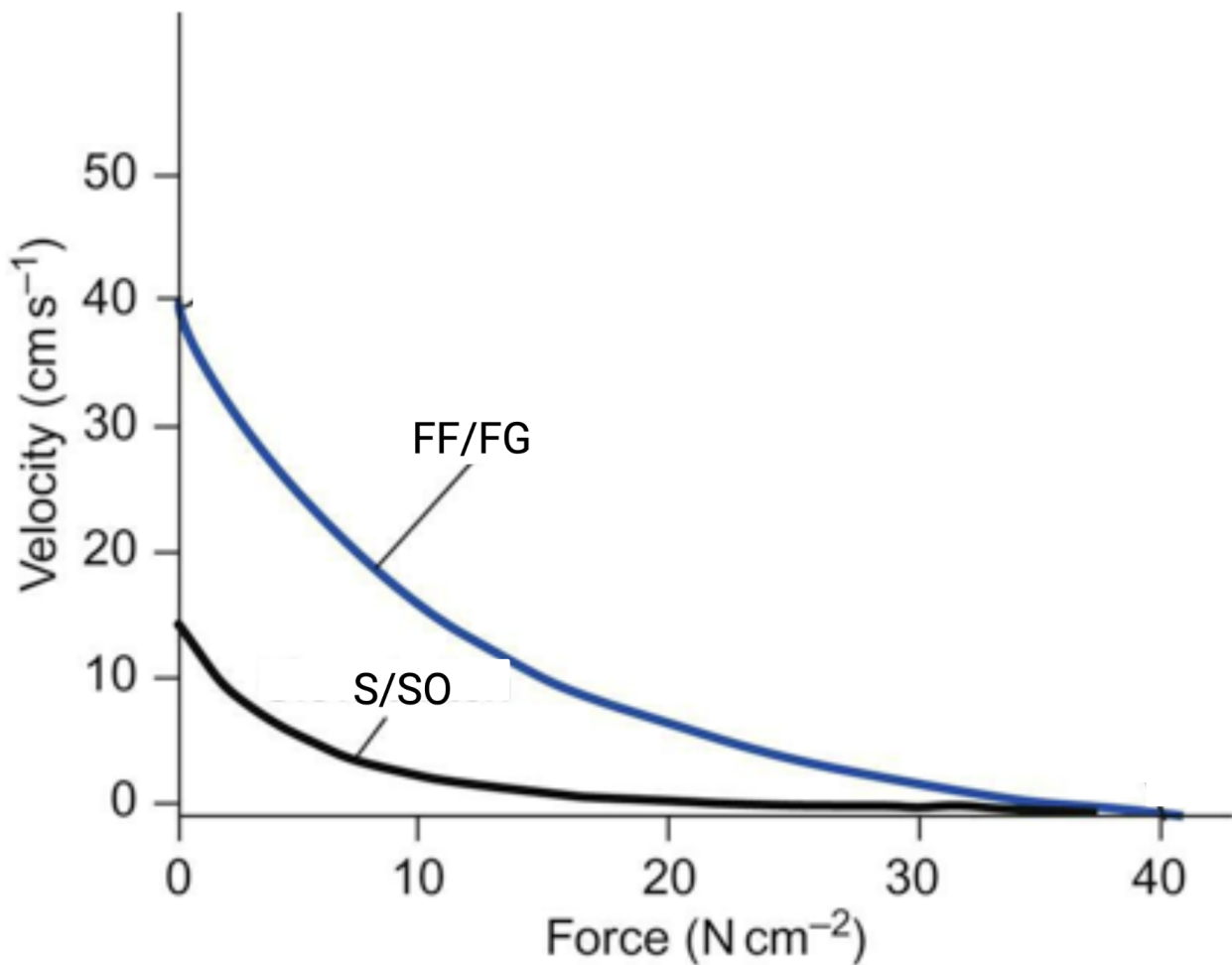


Fig. 5.3: Force and Shortening Velocity of FF/FG and S/SO Motor Units (Created in BioRender.com, Modified from [Feh17])

tension of a motor unit relative to its cross sectional area. When comparing tension of fast vs. slow muscle fibers this allows comparison that avoids the confounding effect of greater cross sectional area.

The molecular (signal transduction) events that lead to the expression of myosin and myosin ATPase isoforms have not yet been completely uncovered. It is assumed that, due to the plasticity of muscle fiber types reported in experimental conditions, at least some muscle fibers retain the ability to express either isoform and that there exists a molecular mechanism that can theoretically be switched [SR96].

Summary of Fast & Slow Muscle Fibers

The functional characteristics of muscle fibers for fast motor units allow for faster excitation-activation coupling and faster crossbridge kinetics. The structural characteristics that enable these include a higher volume of T-tubules and SR (for activation) and an isoform of myosin and myosin ATPase (for crossbridge kinetics). These changes result in an upward shifted force-velocity curve for fast motor units as compared to slow motor units (Figure 5.3).

5.2 Motor Unit Twitch & Tetany

A single excitation of a muscle fiber produces enough activation to create a twitch, making a twitch the fundamental unit of active tension in the muscle fiber. The concept of a muscle fiber twitch is scaled up to a motor unit twitch. A muscle fiber twitch is the fundamental unit of active tension, a motor unit twitch is the fundamental unit of movement.¹ Movements are created by muscle *in situ*, and this means muscle fiber excitations and twitches occur in response to motor unit excitations, and the active tension developed by muscles that results in movement are based on motor unit twitches. When the neuron of a motor unit is excited all of the muscle fibers of the motor unit are excited. Recall from the data in Table 5.1 that for the gastrocnemius this means a motor unit twitch includes approximately 1720 muscle fiber twitches. The fibers that are parallel increase the cross sectional area (A) and the tension.

5.2.1 Motor Unit Type - Twitch & Tetany Characteristics

The three motor unit types display different twitch tension, twitch time and fatigue index (See Table 5.3). These measurable functional characteristics are the result of the structural differences of both the motor unit neuron and the associated muscle fibers discussed in the previous section.

Motor Unit Type	Twitch Tension	Twitch Time	Fatigue Index
Fast-Fatigable (FF)	High	Fast	Low
Fast-Resistant (FR)	Moderate	Fast	Moderate
Slow (S)	Low	Slow	High

Table 5.3: Motor Unit Types

Twitch Time

Twitch time refers to the total time of the twitch which includes its upslope, time at peak or plateau, and time for downslope. A FF motor unit has a fast rise, short lived peak, and fast downslope. An S motor unit has a slower rise, longer plateau, and slower downslope. These characteristic curves (See Figure 5.4) contribute to the excitation frequency required for achieving tetany in the different motor unit types. An S motor unit can achieve tetany with an excitation frequency of 20-30 Hz since the twitches last longer and fuse (summate) more easily. A FF motor unit requires a higher frequency (80-100 Hz) to achieve tetany.

Twitch Tension

Twitch tension refers to the measurable force of the tension developed during a twitch. Since tetany includes a summation of the twitches, a higher twitch tension results in a higher tetanic tension. Twitch tension is highest in the FF motor units for three reasons.

1. These motor units have the highest number of muscle fibers (N, which is directly related to innervation ratio (IR)).
2. These motor units have, overall, a higher cross sectional area. Part of the larger cross area, but not all of it, is due to the higher number of fibers. The IR of a FF motor unit is approximately 2x higher than an S motor unit. However, the cross sectional area (A) is approximately 2.3x higher than an S motor unit [Bod+87; BS80].
3. These motor units have a slightly higher specific tension (T_s). Estimates are that the T_s of FF motor units when tetanized is approximately $22 \text{ N} \cdot \text{cm}^{-2}$ compared to S motor units approximately $17 \text{ N} \cdot \text{cm}^{-2}$. Based on these estimates the T_s of a FF motor unit is approximately 1.3x that of a S fiber.

¹ Just a note that by "unit" we don't mean unit as in a unit of measurement. Other words for how we are using the word unit may be "element" or "atom" or "particle".

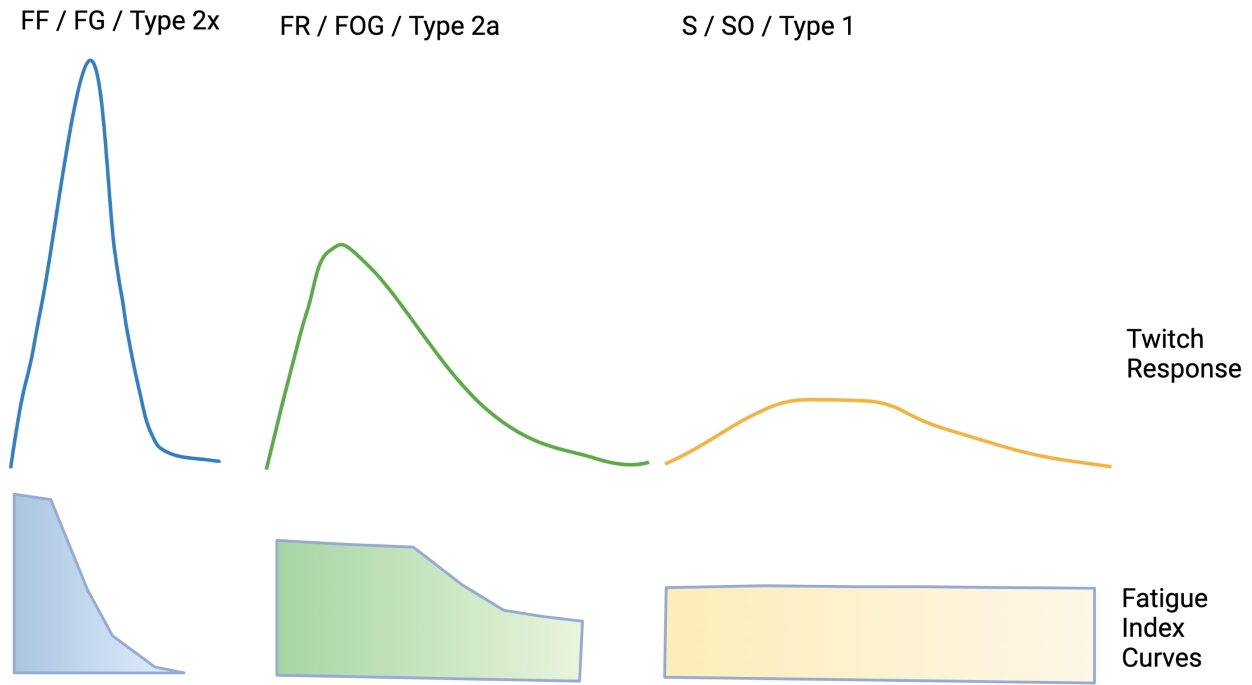


Fig. 5.4: Motor Unit Twitch Characteristics (Created in BioRender.com, Modified from [JRD06])

The tension of a motor unit twitch T_{mu} is the sum of the motor units muscle fiber twitches T_{mf} . The equation:

$$T_{mu} = N \cdot T_{mf} = N \cdot A \cdot T_s \quad (5.2)$$

- N refers to the number of muscle fibers. This is either for one motor unit or can be estimated for an entire muscle by the innervation ratio.²
- A refers to the cross sectional area - an indication of how many sarcomeres are in parallel in the fiber.
- T_s refers to the tension created for each unit of cross sectional area.

Using the above equation and estimates of FF motor units compared to S motor units we can derive an estimate of the FF motor unit T_{mu} relative to the S motor unit.

Assumptions:

- FF motor unit IR, and therefore N is 2x that of S motor units (more fibers)
- FF motor unit A is 2.3x that of S motor units (greater cross sectional area)
- FF motor unit T_s is 1.3x that of S motor units (higher tension per sarcomere)

An equation for the T_{mu} for a FF compared to a S motor unit simply adjusts the variables based on a FF motor unit:

$$T_{FF} = (N \cdot 2) \cdot (A \cdot 2.3) \cdot (T_s \cdot 1.3) \quad (5.3)$$

This equation estimates that a FF motor unit can develop 6x more tension than an S motor unit. If based just on the number of fibers and the cross section area the estimate would be 4.6x. Therefore the number of fibers and the cross sectional area account for 76% of the difference in tension between the FF and the S motor units. And the specific tension accounts for the remaining 24%. Based on these estimates it is not surprising that differences in the specific tension between FF and S motor units is an area of controversy. When considering the likely variability between muscles, between samples, between mammalian species, differences as small as 1.3x are difficult to demonstrate [Lie10]. Some sources equate the T_s between the two fiber types, $T_s \approx 20 N \cdot cm^{-2}$ [Feh17], whereas others report them separately with albeit different estimates, FG $T_s \approx 22 N \cdot cm^{-2}$, and SO $T_s \approx 15 N \cdot cm^{-2}$ [Lie10].

² $IR = \frac{N_{mf}}{N_{mu}}$ (Note, when there is one motor unit, $N = IR = N_{mf}$)

Fatigue Index

Fatigue Index (FI) refers to how long tetany can be sustained for the FF, FR and S motor units. As Table 5.3 indicates, and Figure 5.4 depicts the FF (fatigable) motor units have a low FI, whereas the S motor units have a high FI.

5.2.2 Summary of Motor Unit Twitch & Tetany

The structural and functional differences between motor unit types result in expected differences in motor unit twitch and tetany characteristics. These differences lend themselves to the functional movement benefits of each type of motor units. FF motor units are optimized for fast, high force, high power - ballistic and temporary or intermittent movements. S motor units are optimized for slower, lower force and lower power - sustained, postural, stabilizing postures and movements. Further support for this story comes from how the motor units are recruited, that is, how tension is regulated.

5.3 Active Tension Regulation

Active tension is regulated through motor unit excitation. The two approaches include frequency summation (rate coding) of each motor unit and motor unit summation (recruitment) of more motor units. Both of these approaches start in the spinal cord with motor unit excitation.

5.3.1 Motor Unit Excitation

An α -motor neuron, once excited, releases the neurotransmitter ACh at its axon terminal. ACh has one effect at the NMJ, creating excitatory micro-potentials on the motor end plate. In the CNS, more specifically for motor neurons in the spinal cord, competing neurotransmitters are released at synapses of the α -motor neuron dendrites. The impact of these neurotransmitters can be excitatory or inhibitory. Since they occur after the synapse they are called excitatory post synaptic potentiations (EPSPs) or inhibitory post synaptic potentiations (IPSPs). Whether, and how frequently, an α -motor neuron sends an excitation to the muscle fibers it innervates involves competition between the EPSPs and the IPSPs. For example, the IPSPs of certain protective reflexes associated with receptors in the muscle (see Golgi Tendon Organs below) can reduce the tension developed in a motor unit muscle fibers due to IPSPs lowering both the frequency of stimulation of a motor unit and the number of motor units recruited.

5.3.2 Motor Unit Tetany - Frequency Summation

Motor unit tetany involves the frequency of excitations. For any given motor unit, across all types, there is a range of excitation frequencies (rate coding) that can be used to vary the tension developed. It is estimated that frequency summation accounts for small fine tuning differences in tension. Frequency summation is not capable of resulting in large variations in tension in large part because it involves a set of muscle fibers that have a narrow functional frequency range. S type motor units can tetanize as low as 20 Hz, but don't develop any more force beyond perhaps 60 Hz (with a plateau occurring as low as 40 Hz). Similarly, FF motor units don't tetanize until they achieve upwards of 100 Hz and plateau at approximately 140-160 Hz. These differences, assuming linear increases in tension occur as frequency increases, represent a roughly 50% increase in tension. If someone's grip strength is 100 pounds of force, then getting from 1 to 10 pounds of force requires a 10 times increase, getting from 1 to 100 pounds requires a 100 times increase. These increases in force cannot be accounted for by frequency summation. This has implications in situations discussed later which include a loss of motor units.

Frequency summation fine tunes and adjusts the tension in motor units but it is not capable of large fold changes in tension (or force). But, due to the role that excitation plays in motor unit tetany and the importance of tetany of motor

unit muscle fibers for functional movements, the differences in the frequency of excitation to achieve tetany between the motor unit types is an important concept as we consider motor unit recruitment.

5.3.3 Motor Unit Summation - Recruitment

The primary way that muscle tension is regulated is through motor unit summation (recruitment). The recruitment of more motor units has the capability of greatly increasing the tension. From the example above, to go from 1 pound to a 100 pound hand grip requires the recruitment of few to perhaps all of the motor units involved in grip strength. Figure 5.5 depicts what happens when the three different motor units are recruited, first the SO fibers (S motor units), then the FOG fibers (FR motor units), and finally the FG fibers (FF motor units).³

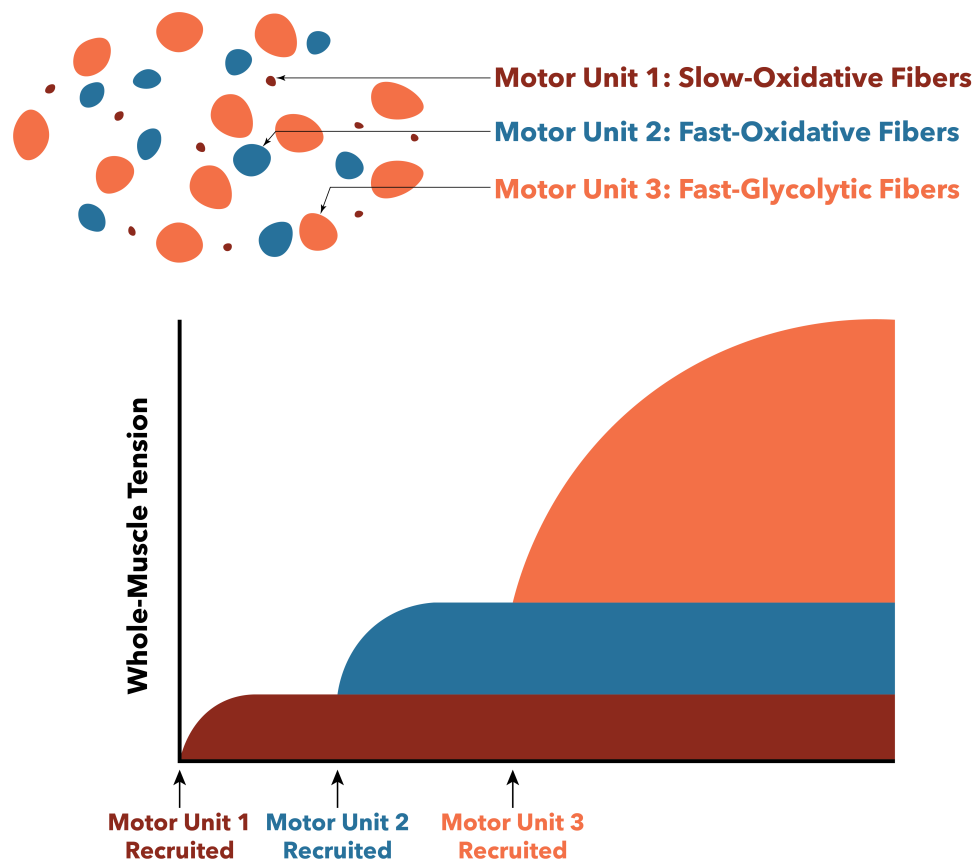


Fig. 5.5: Motor Unit Order of Recruitment (Wikimedia Commons, CC BY 4.0)

Size Principle of Motor Unit Recruitment

The size principle of motor unit recruitment is based on the experimental observation that low tension requirements are met by first recruiting smaller motor units (S) with their associated SO fibers. As tension requirements increase the motor units recruited increase in size [Hen+74]. This is due to two interrelated characteristics of the motor units themselves. First, smaller motor units require less excitation potentials to reach a membrane threshold potential. Lower CNS output into the region of the spinal cord (dorsal horn) with the dendrite pool of α motor neurons for a particular muscle first

³ A future version of this figure will change upslope of these curves to represent the Twitch Time characteristics of the motor unit types).

excites small dendrite membranes (they have less excitation inertia). As the CNS output increases in intensity, signaled as an increase in frequency of excitatory neurotransmitter impulses, there is an increase in excitation of the larger motor neurons. The second characteristic is that the larger motor neurons (FR and FF) require higher frequency excitations in order to achieve tetany.

Tension Developed during Motor Unit Summation - Recruitment

Figure 5.6 is a simulation based on the previously derived estimate that a FF motor unit generates 6x more tension than a S motor unit. This estimate accounts for the higher number of fibers, the greater cross sectional area of the fibers, and the greater specific tension of the fibers in an FF motor unit. Additional assumptions for this simulation include that the muscle has 100 motor units generating up to 100% of its maximal active tension; and that there are 20 S motor units, 20 S to FR transition motor units; 20 FR motor units; 20 FR to FF transition motor units; and 20 FF motor units. The S motor units each contribute 1 arbitrary unit (au) of tension, the FR motor units each contribute 3 au of tension, and the FF motor units each contribute 6 au of tension. The transition motor units are scaled linearly from 1 to 3, and from 3 to 6 in their respective transition zones. What is clear is that based on the order of recruitment (S to FF) there is an upward slope of tension with each new motor unit recruited. Jumps in tension as more motor units are recruited get larger as more motor units are recruited. Frequency summation tuning offers can further adjust these jumps. This tuning is more effective at lower tension with the small motor units since there are smaller jumps to accommodate.

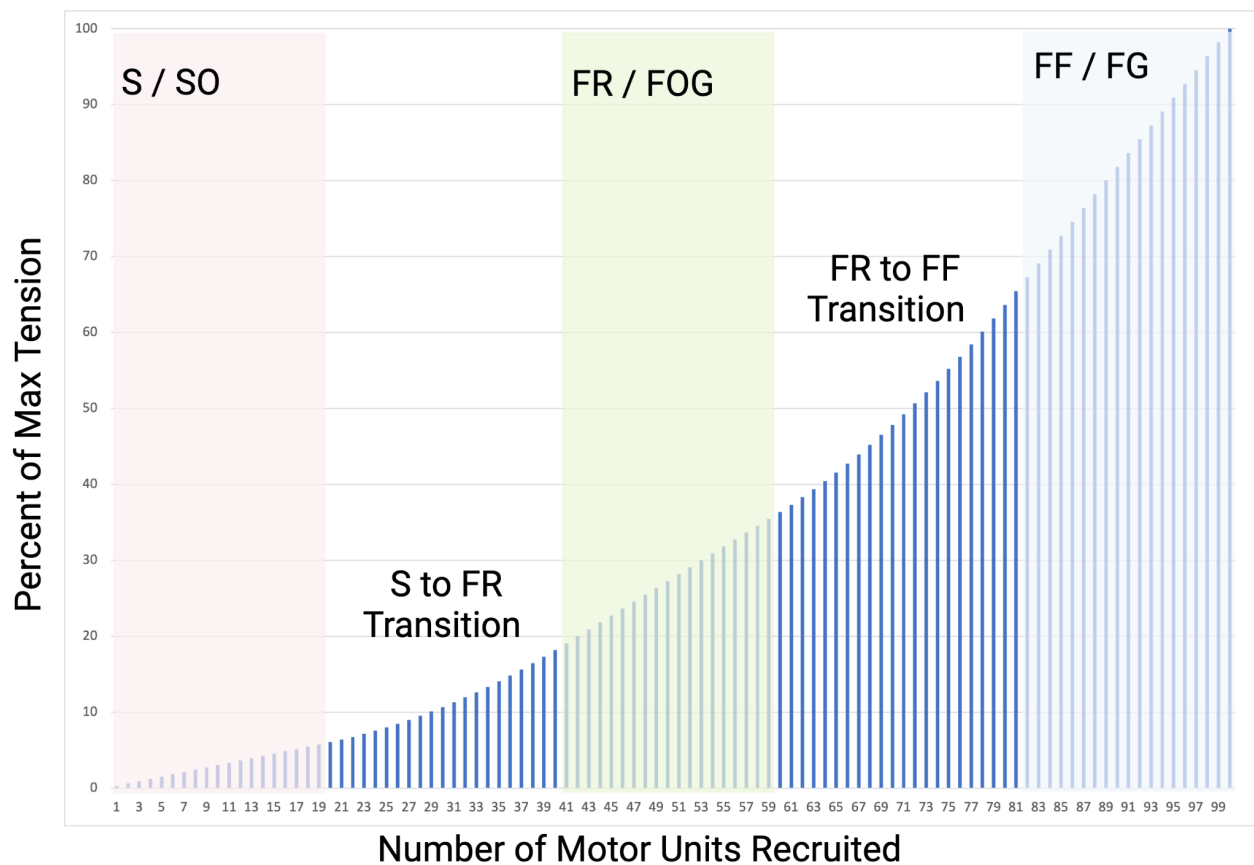


Fig. 5.6: Simulation of Tension Developed during Motor Unit Summation - Recruitment (See text for simulation assumptions) (Created with Excel and BioRender.com)

A commonly experienced consequence of the order of recruitment, that is clear in Figure 5.6, is that regardless of the muscle innervation ratio, there is less precision in tension regulation, and therefore less coordination and fine motor control when a muscle is being near maximal or maximally recruited. Muscles with lower IRs (such as the eye muscle described earlier with an IR of 5), will retain precision to a much higher level of tension, but it will still be reduced at higher tensions. If you have ever tried to fix your apple laptop and remove the small screws that require precision but a lot of force to remove, you understand the challenge of high tension precision.

5.4 Regulatory Feedback

It is estimated that between 40% to 70% of the nerves in a nerve bundle to a skeletal muscle fiber are α -motor neurons (efferent nerves, exiting the spinal cord). The remaining axons are from sensor nerves (afferent nerves, accessing the spinal cord). Regulation of tension is critically dependent on knowing whether the intention of tension is being met (whatever that intention may be) so that tension can be adjusted as necessary. Since neither motor control or neuroscience is a focus of this book (or course) this section is brief and focused on the muscle spindles and golgi tendon organs (GTOs) and how they influence tension regulation in the the spinal cord.

5.4.1 Muscle Spindles

Muscle spindles are receptors that are excited when stretched and provide the central nervous system with information about muscle length and rate of stretching. With this information the muscle spindles are important contributors to proprioception (sensation and perception of position and movement). They are located within fascicles along side (parallel to) muscle fibers. It is conventional, once speaking of the muscle spindle, to refer to muscle fibers as extrafusal fibers (not part of the muscle spindle). The muscle spindles have both afferent and efferent innervation and their own small set of intrafusal contractile fibers. The intrafusal fibers are innervated by γ -motor neurons and allow the muscle spindle to adjust to the length of the muscle to remain sensitive to changes in length (See Figure 5.7).

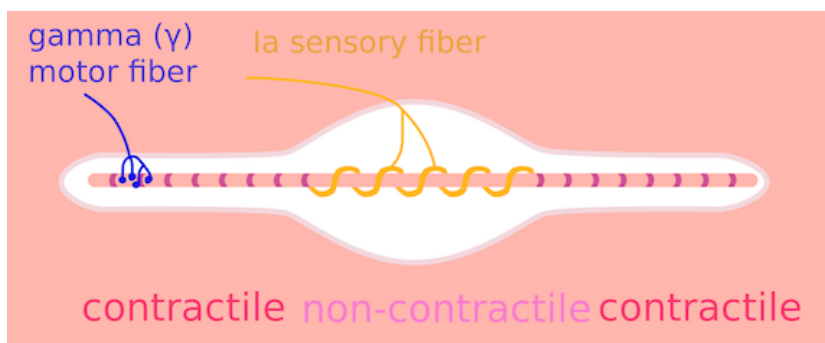


Fig. 5.7: Muscle Spindle [Wikimedia Commons, CC BY 4.0](#))

Afferent signals from the muscle spindles terminate at the spinal cord as well as higher up in the central nervous system for long loop reflexes and coordination (cerebellum) and the cerebral cortex (proprioception). The response of muscle spindles to a rapid increase in length (quick stretch) signals the dendrites of the α -motor neurons of the agonist (same muscle that is being stretched) with excitatory post synaptic potentiations (EPSPs); and the dendrites of the α -motor neurons of the antagonist (opposite muscles of those that are being stretched) with inhibitory post synaptic potentiations (IPSPs). Collectively these actions are referred to as the stretch reflex. If the stretch is quick enough and excites enough muscle spindles the EPSPs are sufficient to excite α -motor neurons which provoke muscle excitation and activation. Testing stretch reflexes is an important component of a physical exam when the integrity of the α -motor neuron, or any part of the sensory - motor reflex loop, is in question.

5.4.2 Golgi Tendon Organs

Golgi tendon organs (GTOs) are on the mostly tendon side of the transition space known as the musculotendinous junction. They are so close to muscle fibers that each GTO is connected in series with 5-20 muscle fibers. In this position the GTOs detect tension in muscle and tendinous fibers [Mac05]. The GTO is excited by tension within the muscle fibers or tendon. Afferent impulses from the GTOs terminate in the spinal cord and cerebellum. GTO afferents in the spinal cord result in IPSPs of the agonist muscle. Inhibiting the agonist limits active tension as a protective mechanism.

5.5 Clinical Physiology Connections

5.5.1 Electromyogram (EMG)

Each time a muscle fiber is activated it is caused by an excitation (action potential) from the nerve to the motor end plate, and along the sarcolemma of all the fibers that are innervated by that α -motor neuron. Therefore, each muscle fiber activation is associated with an electrical signal. An EMG electrode near the fiber can record this electrical activity. The electrode can either be a small needle that acts as an antenna to record local electrical activity or a surface electrode (sEMG). If using a needle, the raw EMG signal can be used to accurately determine when a particular muscle is active (regardless of depth or size of the muscle). sEMG can be used to determine when a particular muscle is active when the muscle of interest is superficial and large enough to avoid cross talk at the sEMG electrode, but cannot be accurately used to determine specific activation of deep or small muscles. For example, sEMG just lateral to the spine cannot be expected to accurately identify activation of deep muscles such as the multifidi.

EMG quantification

It is difficult to quantify the activity of a muscle from a raw EMG signal. The EMG contains electrical signals from every muscle fiber in the vicinity of the EMG electrode which creates a very noisy signal (See Figure 5.8). From the raw signal the Average Rectified Voltage (ARV) EMG can be computed. ARV is a time-windowed mean of the absolute value of the signal (turns all voltages into positive values). ARV is useful for further analysis such as recording the excitation patterns of a muscles that may include changes in excitation or as the first step towards more advanced analyses [MF16]. The relationship between EMG activity and muscle tension is most valuable during isometric activation because the association between activation and tension changes with changes in length (length-tension relationship), and with different velocities (force-velocity relationship).

Biofeedback

sEMG is a valuable form of biofeedback to indicate that a certain muscle has been activated. Commercially available devices provide visual or auditory signals so a patient can learn to activate a muscle. Common applications of the EMG for biofeedback is activation in the vastus medialis after anterior cruciate ligament surgery, or the lower trapezius during shoulder flexion or abduction.

Motor Unit Number and Size Indices (MUNIX & MUSIX)

The motor unit number index (MUNIX) and the motor unit size index (MUSIX) are based on sEMG data interpreted through a mathematical model to estimate an index of the number and size of functional α -motor neurons for a muscle [Nan+04]. MUNIX is not an absolute count of motor units, and MUSIX is not an absolute size of the motor units. But they do provide an estimate that correlates with other measures and they have reasonable reliability. Considering the test only takes 3-5 minutes per muscle, is non-invasive and relatively comfortable (as compared to other approaches), they have gained popularity in clinical neurology and neurophysiology. MUNIX is considered an effective bio-marker

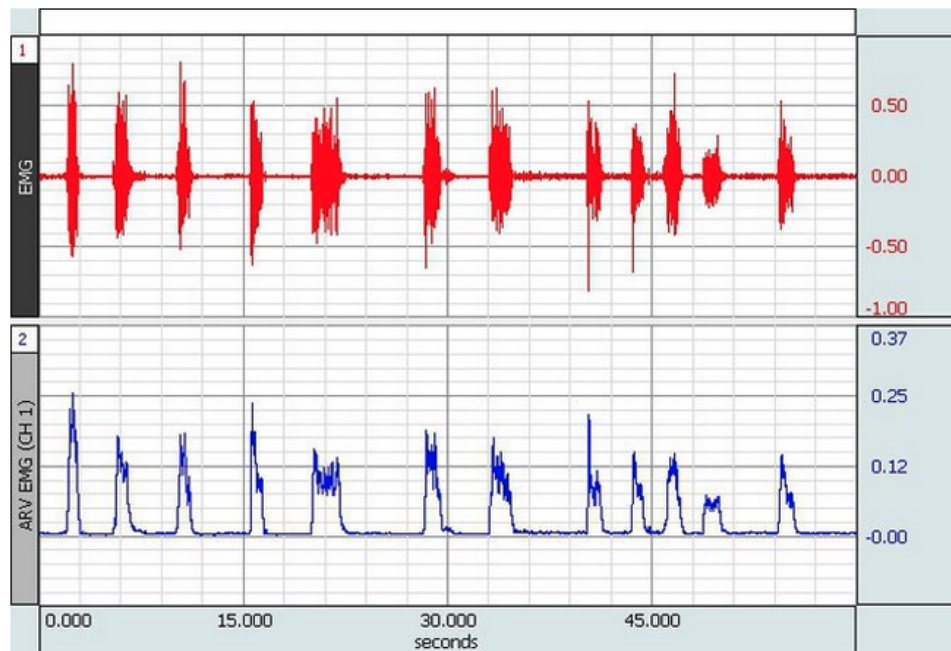


Fig. 5.8: Raw (upper) and ARV (lower) EMG Signals Data collected using Biopaq acquisition hardware and Acknowledge software in the author's lab (SC)

for assessing disease progression and an alternative clinical trial outcome for conditions associated with a decrease in motor units such as amyotrophic lateral sclerosis (ALS). It has also been proposed for use in the evaluation of peripheral neuropathies, particularly in demyelinating neuropathies with conduction block. While there have been less studies, MUNIX is starting to be evaluated in a variety of central nervous system disorders (such as spinal cord injury, stroke, and cerebral palsy [Fat+18].

A study of healthy individuals between 20 and 80 years of age demonstrated a reduction in MUNIX and MUSIX values with aging, particularly over the age of 60, indicating that they may be valuable biomarkers for research on the prevention and treatment of sarcopenia [Cao+20]. Sarcopenia is an age-related progressive loss of muscle mass and strength [Dao+20]. It is a type of muscle atrophy primarily caused by the aging process and certainly confounded by physical inactivity and malnutrition. Sarcopenia tends to result in a larger loss of FF motor units and FG (Type 2a) muscle fibers. If a loss of motor units is inevitable with aging, then healthy aging would strive for lifestyles that promote increasing MUSIX despite an inevitable decline in MUNIX. A consequence of this shift (reduced MUNIX and increased MUSIX) would be a reduction in muscle tension regulation. Therefore, the choice of lifestyle would need to consider the motor control (balance, coordination) capabilities in such individuals.

5.5.2 Nerve Conduction Velocity (NCV)

Nerve conduction velocity (NCV) refers to the time it takes for an excitation to travel a certain distance. It can be measured clinically with electrodiagnostic testing that includes EMG. A stimulus is presented proximally at a peripheral nerve and the distal muscle excitation (via EMG) or axonal excitation is measured.

The tight regulation of active tension for coordinated movement requires fidelity in the signals coming from the central nervous system.⁴ If the myelin sheath of an α -motor neuron is damaged the NCV is reduced. If the NCV is reduced then the ability to activate muscles with high fidelity signals is impaired resulting in delays in the generation of active tension, limiting the ability to increase active tension when needed. These changes occur for two reasons, first, limitations in frequency summation, and second, limitations in motor unit summation. This is precisely what happens in peripheral

⁴ Fidelity in this context means that intended signals are accurately sent where and when they are requested.

demyelinating diseases such as Guillain-Barre syndrome, Chronic Inflammatory Demyelinating Polyradiculoneuropathy, and Anti-Myelin Associated Glycoprotein Neuropathy. The demyelinating disease known as Multiple Sclerosis (MS) also impairs signal fidelity, however, it is a central nervous system condition. Therefore, in MS the loss of signal fidelity does not directly alter the peripheral NCV, it impairs central NCV which interrupts signals that excite the motor unit dendrites in the spinal cord which contributes to a more complicated, nuanced and variable clinical presentation in terms of the impact it can have on motor function such as coordination.

5.6 Summary

Motor units connect the central nervous system to muscle fibers. Varying the frequency of excitation and the number of motor units excited allows the CNS to vary the active tension provided by a muscle. Tight regulation of this active tension is required for motor control. There are three accepted classifications of motor units and they correspond to three accepted accepted classifications of muscle fibers. The structural characteristics of these motor units and muscle fibers are related to distinct functional characteristics in twitch and tetany. Recruitment of motor units is the primary mechanism for varying tension with variations in the frequency of excitation providing a small range of tension variation that allows for fine tuning the nearly continuous range of possible tensions. Muscle spindles and golgi tendon organs provide the CNS with feedback about muscle tension, stretch and length. EMG can capture valuable information about motor unit and muscle excitation as well as the determination of NCV. The next chapter expands on the metabolic characteristics of muscle fibers and the impact muscle energetics has on function as well as the demands that muscle energetics places on other body systems.

References

- [Bod+87] S. C. Bodine et al. *Maximal force as a function of anatomical features of motor units in the cat tibialis anterior*. en. Journal of Neurophysiology 57.6 (June 1987), pp. 1730–1745. issn: 0022-3077, 1522-1598. doi: [10.1152/jn.1987.57.6.1730](https://doi.org/10.1152/jn.1987.57.6.1730).
- [BS80] F Buchthal and H Schmalbruch. *Motor unit of mammalian muscle*. en. Physiological Reviews 60.1 (Jan. 1980), pp. 90–142. issn: 0031-9333, 1522-1210. doi: [10.1152/physrev.1980.60.1.90](https://doi.org/10.1152/physrev.1980.60.1.90).
- [Cao+20] Bei Cao et al. *Reference values for the motor unit number index and the motor unit size index in five muscles*. en. Muscle & Nerve 61.5 (May 2020), pp. 657–661. issn: 0148-639X, 1097-4598. doi: [10.1002/mus.26837](https://doi.org/10.1002/mus.26837).
- [Dao+20] Tam Dao et al. *Sarcopenia and Muscle Aging: A Brief Overview*. en. Endocrinology and Metabolism 35.4 (Dec. 2020), pp. 716–732. issn: 2093-596X, 2093-5978. doi: [10.3803/EnM.2020.405](https://doi.org/10.3803/EnM.2020.405).
- [Fat+18] Farzad Fatehi et al. *The utility of motor unit number index: A systematic review*. en. Neurophysiologie Clinique 48.5 (Oct. 2018), pp. 251–259. issn: 09877053. doi: [10.1016/j.neucli.2018.09.001](https://doi.org/10.1016/j.neucli.2018.09.001).
- [Feh17] Joseph J Feher. *Quantitative human physiology: an introduction*. Academic press, 2017.
- [Hen+74] E Henneman et al. *Rank order of motoneurons within a pool: law of combination*. en. Journal of Neurophysiology 37.6 (Nov. 1974), pp. 1338–1349. issn: 0022-3077, 1522-1598. doi: [10.1152/jn.1974.37.6.1338](https://doi.org/10.1152/jn.1974.37.6.1338).
- [JRD06] David Jones, Joan Round, and Arnold De Haan. *Skeletal muscle: from molecules to movement*. Vol. 40. 2006.
- [Lie10] Richard L Lieber. *Skeletal muscle structure, function, and plasticity, 3rd edition*. Lippincott Williams & Wilkins, 2010.
- [LM93] L Larsson and R L Moss. *Maximum velocity of shortening in relation to myosin isoform composition in single fibres from human skeletal muscles*. en. The Journal of Physiology 472.1 (Dec. 1993), pp. 595–614. issn: 00223751. doi: [10.1113/jphysiol.1993.sp019964](https://doi.org/10.1113/jphysiol.1993.sp019964).
- [Mac05] Vaughan G Macefield. *Physiological characteristics of low-threshold mechanoreceptors in joints, muscle and skin in human subjects*. Clinical and Experimental Pharmacology and Physiology 32.1-2 (2005), pp. 135–144. doi: <https://doi.org/10.1111/j.1440-1681.2005.04143.x>.
- [MF16] Roberto Merletti and Dario Farina. *Surface electromyography: physiology, engineering, and applications*. John Wiley & Sons, 2016.
- [Nan+04] Sanjeev D Nandedkar et al. *Motor unit number index (MUNIX)*. IEEE transactions on biomedical engineering 51.12 (2004). Publisher: IEEE, pp. 2209–2211.

- [SR96] S. Schiaffino and C. Reggiani. *Molecular diversity of myofibrillar proteins: gene regulation and functional significance*. en. *Physiological Reviews* 76.2 (Apr. 1996), pp. 371–423. issn: 0031-9333, 1522-1210. doi: [10.1152/physrev.1996.76.2.371](https://doi.org/10.1152/physrev.1996.76.2.371).

Sample Questions

1. Based on axon diameter, order alpha motor neurons based on NCV.
2. Based on the innervation ratios (IRs) identify which muscle is best equipped for fine motor control.
3. What differences in FG muscle fibers allow faster crossbridge kinetics and higher specific tension (Ts)?
4. When isolated or when Fast fibers dominate in a muscle the force-velocity curve is shifted in which direction and what does that mean for the velocity at a given force, or force at a given velocity?
5. What does activation of Ia sensor fibers by a quick stretch result in?

Chapter 6

Energetics

Updated on May 22, 2024

Contents

6.1	Energetic Transformation Overview	78
6.2	ATP Regeneration Pathways	78
6.3	Motor Unit & Muscle Fiber Types	85
6.4	Clinical Physiology Connections	86
6.5	Summary	91

Muscles transform chemical energy into mechanical energy. The chemical energy utilized in this transformation is ATP. However ATP is not ingested and then circulated through the body to the cells. It is present in the cells in small quantities and regenerated in the cells at a rate that approximates its rate of utilization. The primary utilization of ATP in the muscle cell for active tension occurs when it is hydrolyzed by myosin ATPase during crossbridge cycling. The change in shape of the myosin head is potential mechanical energy that becomes kinetic mechanical energy during crossbridge activation. Muscle excitation and activation requires ATP in two other key steps. The Na^+/K^+ -ATPase pumps to establish the resting membrane potential, and the Ca^{2+} -ATPase pumps to return Ca^{2+} into the sarcoplasmic reticulum. In addition to these clear ATP needs for excitation, activation and tension the muscle fibers require ATP for for membrane transport and synthesis of chemical compounds. These maintenance energy needs are critical for cellular integrity. The chemical energy to regenerate ATP comes from food we ingest, digest, absorb and metabolize into substrates. The absorption of glucose can be directly utilized as a source of energy. All other substrates require metabolism, fatty acids, proteins, glycogen (stored form of glucose). This chapter considers the rate and capacity of pathways to regenerate ATP in muscle cells. These pathways vary across a spectrum with rate of regeneration inversely related to capacity. They are all always being utilized to some extent during muscle cell energetic transformations. Structural variations that exist on a spectrum between muscle fibers result in functional variation in their ability to carry out these five pathways with a high degree of adaptability. Clinical physiology connections emerge based on an understanding of energetics, this chapter focuses on the implications of sarcopenia, the concept of fatigue, and the pathophysiological concepts of hypoxia, hypoxemia and ischemia.

Objectives include:

1. Explain the factors underlying the storage, use and regeneration of ATP.
2. Compare and contrast the energetic system pathways that transform macro nutrient substrates to ATP.
3. Compare and contrast muscle fiber differentiation (types) based on energetic structures and resultant functional capacities.
4. Provide examples of possible differences in muscle tension energetic characteristics and performance based on hypothetical changes to muscle fiber types or aspects of the energetic pathways.
5. Apply the concepts of mass balance, flow gradients, and energy to the analysis of patient/client problems related to the muscular system.

6. Evaluate the reasons different energetic pathways are primary sources of ATP in different activities, and analyze the response to different activities based on the primary energetic pathway being used.
7. Explain the implications of sarcopenia on muscle strength and endurance.
8. Explain the concept of, and the energetic basis of, fatigue.
9. Explain the energetic basis of hypoxia, hypoxemia and ischemia.

6.1 Energetic Transformation Overview

Energetic Transformation is a slightly more explanatory way of describing what is commonly referred to as metabolism. Of course, there are many different types of energetic transformations occurring when considering all metabolism throughout the body. And not all metabolic reactions are specifically about the regeneration of ATP. Our focus is on energetic processes and pathways that transform substrate energy to regenerate ATP in muscle. Most of these processes and pathways occur in cells throughout the body.

Muscle fibers continuously utilize ATP. Resting utilization of ATP results in a steady demand for ATP which is met by an equal rate of ATP regeneration. The balance of demand and regeneration (supply) is regulated by the cell. The muscle fiber stores ATP at relatively low concentrations ($5 \text{ mmol} \cdot \text{L}^{-1}$) [Feh17; JRD06]. At rest there are low energetic needs (low ATP demand) so the tendency to store a small amount of ATP is not a problem. But muscle fibers can increase their energetic needs (rate of ATP utilization, demand) 60-100 fold during transitions from rest to high levels of tension. When ATP demand increases, supply (rate of ATP regeneration) must increase in response. Resting muscle has approximately $1 \text{ mmol} \cdot \text{L}^{-1}$ of ADP and $10 \text{ mmol} \cdot \text{L}^{-1}$ of P_i . Regulating the rate of ATP regeneration is based on maintaining this 5:1 ratio of ATP:ADP. When the muscle fiber starts to use more ATP the ratio drops and the rate of regeneration increases throughout all pathways until the ratio is reestablished. If it cannot be reestablished by increasing the rate of regeneration and the amount of ATP falls below the ATP demand, then whatever functions need ATP start to fatigue (fail to meet their objectives).

There are two proposed reasons that ATP is not stored in large quantities. First, it is a heavy molecule. It is estimated that an individual regenerates their body weight worth of ATP [TN08]. ATP is 507.18 g/mole. Whereas glucose is 180 g/mole and even fatty acids with up to 29 carbon chains and double bonds (high energy) have molar masses lower than ATP (less than 500 g / mole). Since mass (and weight) influence the energetic costs of movement then storing ATP is a less efficient approach to energy storage than storing glucose (as glycogen) and fatty acids (as adipose tissue). Second, ATP is a highly reactive molecule that triggers several cellular events by binding and hydrolyzing. Too much ATP being stored could cause "chaos" in cellular functions [JRD06]. The implications of not storing ATP includes the continuous need to regenerate ATP at the rate it is being utilized while maintaining a reserve (surge) capacity.

6.2 ATP Regeneration Pathways

There are six ATP regeneration pathways in the muscle fiber. The myokinase pathway is briefly described as part of the immediate ATP/Phosphocreatine (PCr) pathway. The pathways are summarized in Figure 6.1. The rates and capacities of the pathways are later directly compared in Table 6.1.

Some of the pathways occur in the sarcoplasm (outside of the mitochondria, including PCr, Myokinase, Muscle Glycogen $\rightarrow$ Lactate Pathway). These pathways are not rate limited by the structural requirements of mitochondria. Pathways occurring outside of the mitochondria do not involve cellular respiration and are referred to as anaerobic (not requiring O_2). The anaerobic pathways follow standard chemical substrate phosphorylation. The chemical reactions of these pathways are catalyzed by enzymes and regenerate ATP at high rates but low overall yield. The low yield is due to the fact that they tend to be self limiting. They either recycle existing energy and quickly exhaust their capacity (myokinase and PCr) or they create byproducts that inherently limit their process (the glycolytic process within the Muscle Glycogen $\rightarrow$ Lactate pathway). With either self limitation they can temporarily regenerate ATP at a higher rate than their support systems allow. They meet a particular need, provide a high rate of ATP regeneration, but fatigue quickly and must recover. Such work:recovery cycles are a recurring theme for the concept of fatigue.

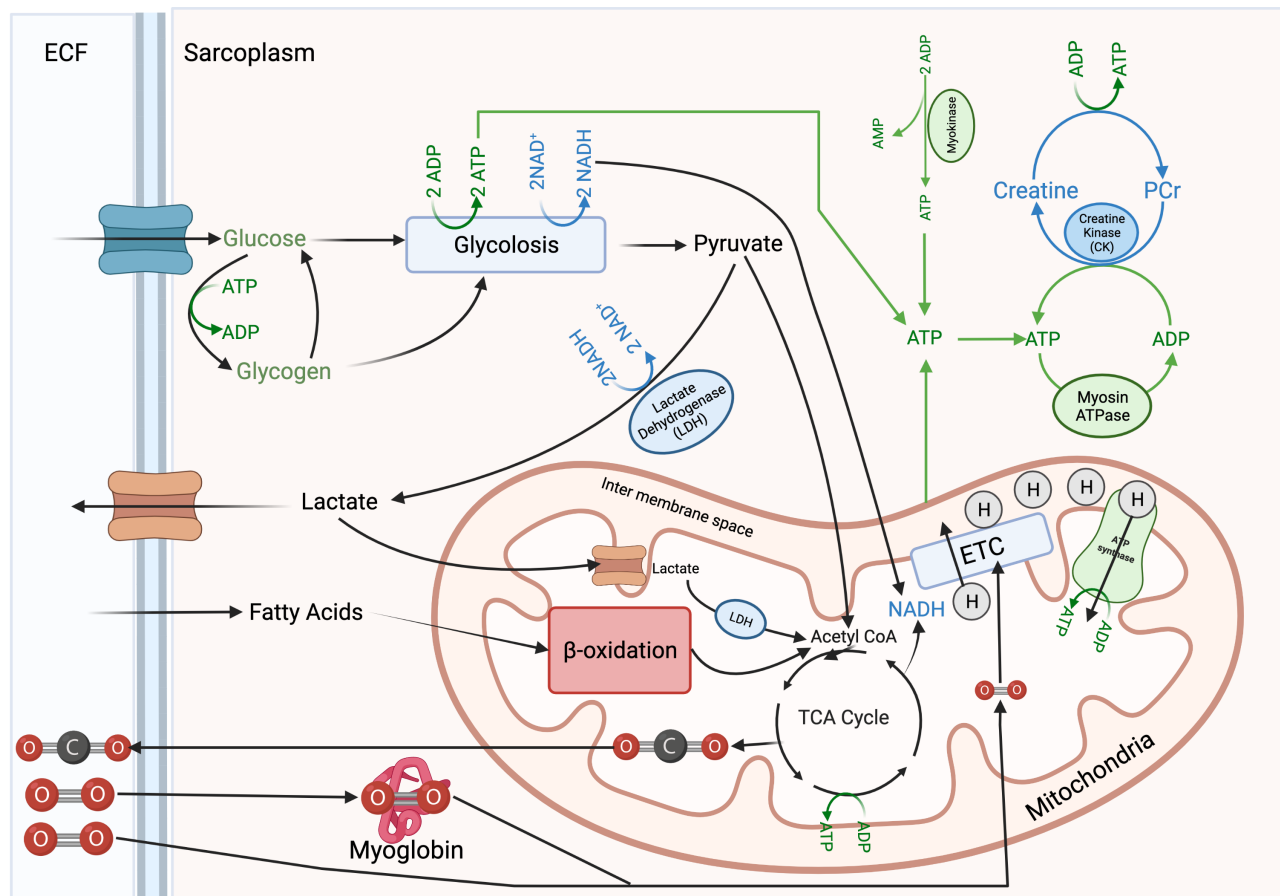


Fig. 6.1: Overview of Energetic Transformation Pathways in a Muscle Fiber. An important pathway not explicitly included in this graphic is the pathway that starts with the breakdown of glycogen to release glucose into the blood which ultimately enters cells through the glut4 channel. This and other important metabolic processes and pathways occur in the liver and are covered in the chapter on Digestion-Absorption-Metabolism. (Created with BioRender.com)

Pathways occurring in the mitochondria involve cellular respiration and are referred to as aerobic (Muscle Glycogen → CO_2 ; Liver Glycogen → CO_2 ; Fatty Acid → CO_2 pathways). All pathways that produces Acetyl Co-A enter into the citric acid cycle (TCA, Kreb's Cycle). TCA contributes some ATP, but primarily contributes more $NADH$ and creates some CO_2 in the process. $NADH$ is utilized to generate large quantities of ATP by sending it through the Electron Transport Chain (ETC). The ETC utilizes $NADH$ to build a concentration and electrochemical gradient of H^+ in the mitochondrial inter membranous space. ETC rate dependent on the availability of O_2 to accept the transported electrons in the final step. The energy of this gradient is pushed through ATP-synthase like water through a turbine at a dam. The reason there are fluctuating estimates on the number of ATP produced through aerobic pathways (oxidative phosphorylation) is that the ATP is produced by the function of a molecular machine, not with standard chemical substrate phosphorylation.

There is no switch!

There is no switch that turns on anaerobic energy transformation and turns off aerobic; or that turns on aerobic or off anaerobic. There are no exercises that are exclusively anaerobic or exclusively aerobic. A slow walk still utilizes anaerobic pathways. A heavy lifting session regenerates ATP and PCr utilizing the underlying continuous activity aerobic pathways. Dispel any notions that these socially constructed concepts entail an exclusive monopoly of energetic pathways to regenerate ATP. All energetics are on a spectrum. The primary demand for ATP during the slow walk are met by aerobic pathways, and the primary demands of the heavy lifting session are met by anaerobic pathways. But all ATP is ultimately regenerated by aerobic metabolism. It is this concept of meeting primary demands that has been utilized

for the social construction of aerobic and anaerobic exercise and the misinterpretation that these are mutually exclusive pathways.

There is no ATP Debt

There is no such thing as ATP debt. ATP utilization during crossbridge activation by myosin-ATPase is absolute. If there is not enough ATP there will be less crossbridges involved in generating active tension. The active tension cannot be generated with the energy for the tension paid back later. Either the pathways provide a way to regenerate the ATP at a rate needed for the demand, or the active tension is not generated as intended. The primary determinant - in a muscle fiber - of the rate of ATP utilization, and therefore need for regeneration, is the frequency of excitation - activation. The primary determinant - in an entire muscle - of the rate of ATP utilization, and therefore the need for regeneration, is the frequency excitation across all the motor units recruited. If the ATP is not regenerated at the rate it is being utilized, cellular functions do not occur.

Energetic Pathways - Directed Graph

Figure 6.2 is a directed graph (DG) of the energetic pathways. Details are excluded. All models are wrong, some models are useful. This directed graph is useful for tracking the different energetic paths. This DG does not have clear boundaries for intracellular and extra cellular, or sarcoplasm and mitochondria spaces. A general pattern of the DG is that things entering the fiber come from the top or left of the graph, and things exit the fiber to the right. It is recommended that the reader compare and contrast Figure 6.1 and 6.2.

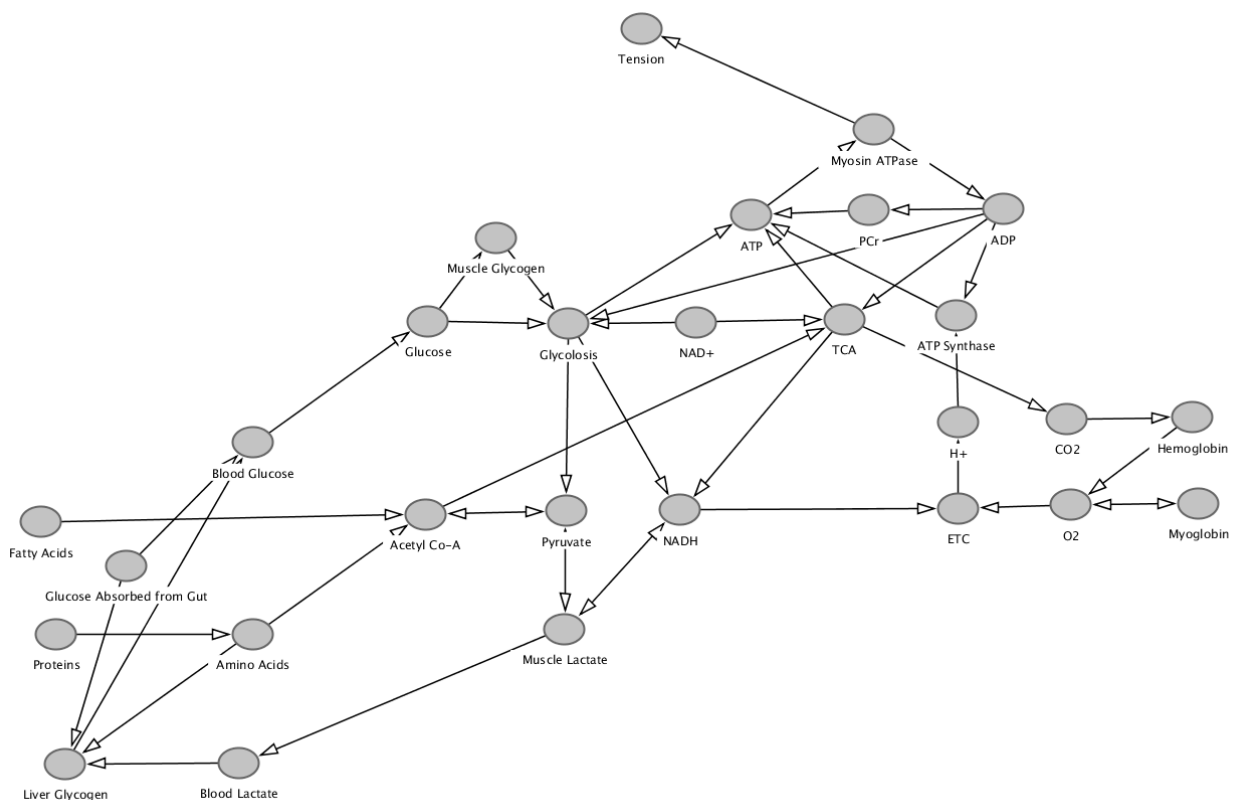


Fig. 6.2: A directed graph (DG) of the energetic pathways (Created with DAGitty.com and available for modifying at http://dagitty.net/mZ_hp70)

6.2.1 ATP / Phosphocreatine (PCr) (immediate) Pathway

Creatine is synthesized in the liver, or ingested, digested and then absorbed from meat or ingested and absorbed from supplementation. Creatine is a molecule that can be absorbed directly through the gut into the blood. A high energy bond can form between creatine and phosphorous (P_i), forming PCr.

The reaction: $PCr \rightleftharpoons ATP$ is a reversible high rate reaction catalyzed by the enzyme creatine kinase (CK). The rightward reaction (PCr to ATP) regenerates ATP for cellular functions. At rest about 80% of creatine is in the energized (PCr) state and there is approximately five times as much PCr than ATP (PCr:ATP ratio of 5:1) in a muscle fiber. Based on the weight and reactive volatility of ATP, PCr provides a very important reserve of high energy phosphate bonds for ensuring the muscles ability to meet high energy demands, and even to smooth out transitions between lower and higher ATP utilization rates.

Primary determinant muscular ATP utilization

In a muscle fiber the primary determinant of ATP utilization is the frequency of activation which is directly proportional to frequency summation. It is a stable and relatively light form of energy to store in the muscles.

PCr is considered an immediate source of ATP and the flux rate of PCr to ATP is determined by the ATP:ADP ratio. As ADP levels increase the rate of PCr to ATP reactions increase and will continue until PCr is expended or ADP drops to a normal level.

The maximum rate of $PCr \rightleftharpoons ATP$ is much higher than other pathways, upwards of 4.4 moles/minute of ATP. However, it is only capable of regenerating approximately 0.7 moles of ATP overall. Therefore, when working at maximum capability (rate of 4.4 moles/minute) it is sustained for 9.5 seconds (See the Summary of Max Rates of ATP Regeneration by Pathway in Table 6.1). This time estimate is dependent on the assumption of the max rate (4.4 moles/minute). Once the capacity of PCr to regenerate ATP is exhausted (fatigued) the muscle cannot continue to generate a tension that requires that rate of ATP utilization.

If a lower rate of ATP was being utilized, say 0.5 moles/minute, then PCr would be able to sustain ATP regeneration much longer. And since that rate of ATP utilization is lower than other pathways then PCr would continue to be regenerated by ATP coming out of the mitochondria. In this situation the need for ATP from PCr would not approach maximum because the aerobic pathways are contributing to the ATP regeneration.

Not just for high muscle tension activities

Even though PCr is most often thought of as the source of ATP for high muscle tension activities, it also is the initial source of ATP for any changes in the rate of utilization of ATP because other pathways have a delay (when compared to PCr) in how quickly they can start regenerating more ATP. Therefore, PCr is the source of ATP for a sudden burst of tension such as that needed to increase pace during a run, or increase velocity of a movement. PCr ensures that ATP is regenerated at the rate needed during any transitions in the rate of ATP utilization. In this sense, it buffers differences between ATP demand and supply and thus allows smooth active tension transitions.

ATP-PCr Sarcoplasm Shuttle

The PCr pathway is also utilized as a ATP transport mechanism throughout the sarcoplasm (ATP-PCr shuttle). There are substantial diffusion barriers in a muscle fiber due to the number of highly structured protein complexes. The mitochondria are near the sarcolemma (along with the nuclei largely because the rest of the fiber is packed tightly with myofilaments). The diffusion of ATP from the mitochondria to myosin-ATPase is challenging. PCr is diffused through the entire sarcoplasm and due to its rapid regeneration rate can easily shuttle (or pass) the energy of ATP through the sarcoplasm from just outside a mitochondria to near the myosin ATPase. During muscle activation when the rate of ATP use by the myosin ATPase increases the nearby PCr regenerate ATP and then that regeneration can propagate from all around the sarcoplasm, and from near the mitochondria, to shuttle ATP toward the myosin ATPase (See Figure 6.3). This movement - either the diffusion of ATP directly or the use of an ATP-PCr shuttle - is based on an ATP flow gradient from

the mitochondria (source) toward any steps utilizing ATP (sinks). The more ATP being utilized by the myosin ATPase, the Na^+/K^+ -ATPase or the Ca^{2+} ATPase the lower the concentration of ATP near those proteins (sinks) and the greater the diffusion, and shuttling, of ATP towards those locations.

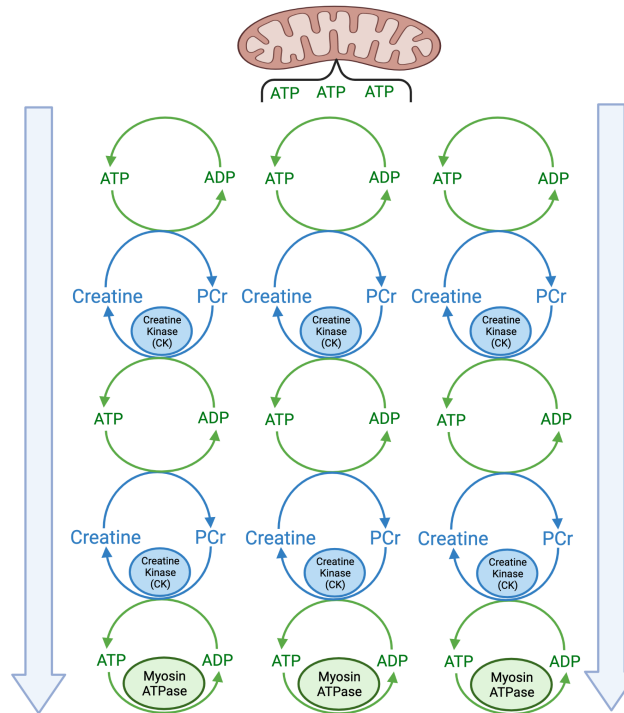


Fig. 6.3: Movement of ATP through the sarcoplasm using an ATP-PCr shuttle

Myokinase (Adenylate Cyclase)

Another immediate high rate reaction to regenerate ATP is the myokinase enzyme (also referred to as adenylate cyclase) that catalyzes the reaction: $2ADP \rightleftharpoons AMP + ATP$. During this reaction a P_i is hydrolyzed from ADP and the energy regenerates ATP. This reaction can happen at a high rate, but it is limited. It is most likely only utilized in extreme ATP demand situations because in most situations ADP is being utilized by other pathways and not available for the myokinase reaction. But if there is a high enough accumulation of ADP myokinase allows continued cell functions for a short period of time. Estimates are that under conditions of maximum ATP utilization (upwards of 4.4 moles/minute), this system could be sustained for milliseconds (less than a second). It is self limiting because it is recycling a dwindling source of energy. For our purposes it is therefore added to the ATP/PCr row of Table 6.1 given its relatively insignificant contribution to capacity. Though it may play a role in smoothing (buffering) responses to variable ATP demands.

6.2.2 Muscle Glycogen → Lactate Pathway

The muscle glycogen (or glucose) → lactate pathway provides ATP regeneration at a moderate rate. It is only capable of maximum rates for a relatively small amount of ATP and therefore for a short period of time. The key biochemical process for this pathway is glycolysis. Glycolysis transforms glucose or glycogen into several important end products including ADP to ATP, NAD^+ to $NADH$ and pyruvate. When regenerating ATP at high rates glycolysis ends up creating more $NADH$ and there is not enough NAD^+ to continue with glycolysis, and the accumulation of $NADH$ results in a

drop in intracellular pH. As a remedy to this situation $NADH$ is converted back into NAD^+ which includes the conversion of pyruvate into lactate.

The glycogen (or glucose) $\rightarrow$ lactate pathway, while always involved a bit, is primarily involved when it is functioning at a rate of ATP regeneration that exceeds the ability of the muscle glycogen $\rightarrow CO_2$ pathway. A simplifying approach, based on the estimates in Table 6.1, is that if the glycogen (or glucose) $\rightarrow$ lactate pathway is generating over 1 mole / minute of ATP it is exceeding the rate that the muscle glycogen $\rightarrow CO_2$ pathway. In this situation it is beneficial to the muscle fiber for lactate to form and be shuttled out of the cell to stabilize the pH and to make NAD^+ available to continue glycolysis.

If the glycogen (or glucose) $\rightarrow$ lactate pathway towards lactate continues it eventually cannot shuttle lactate out of the cell fast enough and cellular pH drops (acidity), and there is not enough NAD^+ , so the rate of glycolysis slows down (forcing a recovery period). A consequence for the muscle is that it cannot continue to generate tension at a rate that requires such a high rate of ATP regeneration.

Anaerobic Threshold - Lactate Threshold - Onset of Blood Lactate Accumulation

Anaerobic threshold (AnT), lactate threshold (LT) and onset of blood lactate accumulation (OBLA) are three related concepts. AnT is primarily theoretical and can be estimated by LT or OBLA which are both based on measurements of blood lactate. Ventilatory threshold (VT) is another approach to measuring AnT discussed in Chapter ???. The concept of AnT is directly related to the previously discussed idea that when the glycogen (or glucose) $\rightarrow$ lactate is regenerating ATP at a rate higher than capable by the glycogen (or glucose) $\rightarrow CO_2$ pathway, the pH of the cell drops and an increased amount of lactate is shuttled out of the cell into the blood. Active tension requiring ATP at a rate of regeneration that needs the glycogen (or glucose) $\rightarrow$ lactate pathway is not sustainable. The time that the active tension can be sustained is based on the different rates of the glycogen $\rightarrow$ lactate pathway and the glycogen $\rightarrow CO_2$ pathway.

6.2.3 Muscle Glycogen $\rightarrow CO_2$ Pathway

The muscle glycogen $\rightarrow CO_2$ pathway provides ATP regeneration at less than half the rate as the muscle glycogen to lactate pathway. But it provides a sustainable source of ATP regeneration. It is not self limiting. There are factors and situations that emerge such as substrate (glycogen) availability that lead to fatigue and a need for recovery of this system. But as long as there is glycogen this pathway will continue to provide ATP regeneration to its max rate. This pathway includes several biochemical processes. It begins with glycolysis which provides ATP, $NADH$ and pyruvate. The ATP is available for cellular energy. $NADH$ and pyruvate enter the mitochondria. Pyruvate is transformed into Acetyl-CoA which can enter the TCA cycle. The TCA cycle generates CO_2 for removal, ATP and more $NADH$. $NADH$ from glycolysis and TCA enter the ETC which transports electrons along the inner membrane of the mitochondria until they reach the final electron acceptor O_2 . The transport of electrons pumps protons (H^+) between the mitochondrial membranes and develops a concentration and electrical gradient that pushes protons through ATP synthase (like water rushing through a dam). The ATP synthase rotates like a turbine and energizes the bond between ADP and P_i to regenerate ATP.

CO_2 is created as part of the citric acid cycle (TCA), O_2 is consumed as part of the electron transport chain (ETC). Note that CO_2 is produced and O_2 is consumed at different steps in the pathways that produce CO_2 as a byproduct and consume O_2 . Together TCA and ETC are referred to as oxidative or aerobic metabolism. ETC is also oxidative phosphorylation. The availability of O_2 is critically important to the rate of ATP generated in all pathways using ETC. The use of O_2 in this pathway is the only use of O_2 in the human body. All O_2 consumed through respiration and measured through ventilatory tests such as indirect calorimetry are based on the use of O_2 in the mitochondria for the ETC. All estimates of metabolic rate that are based on O_2 consumption, including those using heart rate and steps taken, are founded in the fact that the ultimate source of most ATP (energy) passes through ETC and consumes O_2 . The muscle glycogen $\rightarrow CO_2$ pathway produces the same volume of CO_2 as it consumes O_2 . The ratio of $\frac{CO_2}{O_2}$ is referred to as the respiratory exchange ratio (RER). The RER for the muscle glycogen $\rightarrow CO_2$ pathway is 1.0. However, it is not the only pathway that results in an RER = 1.0.

6.2.4 Liver Glycogen $\rightarrow$ CO_2 Pathway

The liver glycogen $\rightarrow$ CO_2 pathway is similar to the muscle glycogen $\rightarrow$ CO_2 pathway. One difference is that the rate of ATP production is lower due to the need to breakdown glycogen in the liver and circulate it through the blood to the extracellular fluid of the muscle for transport into the muscle. There is also a lower overall capacity because more glycogen is stored in muscles (in total) than in the liver. Breakdown of liver glycogen into blood glucose can be used by muscles to either immediately enter glycolysis. If not needed for glycolysis this glucose can be stored as glycogen. Once glucose enters glycolysis it basically follows the same path as the muscle glycogen $\rightarrow$ CO_2 pathway. Whether starting with liver glycogen or with muscle glycogen these two pathways produce the same volume of CO_2 as they consume O_2 , at they both of an RER = 1.0. These are the only two pathways with an RER = 1.0. Therefore, at rest an RER of 1.0 would indicate that a glucose or glycogen pathway to CO_2 was the dominant energetic pathway. Variations to this during activity are discussed in Chapter ?? on the Fick Equation.

6.2.5 Fatty Acids $\rightarrow$ CO_2 Pathway

The fatty acid $\rightarrow$ CO_2 pathway has the lowest overall rate of ATP regeneration. However, this finding is based on years of experiments in individuals consuming typical carbohydrate dominant diets. Research in subjects that are fat adapted challenge this paradigm, suggesting that it is possible, with fat (keto)adaptation that the fatty acid $\rightarrow$ CO_2 pathway may not have as significant a drop in the rate of ATP utilization as previously believed [McS+18]. The substrate for this pathway are fatty acids carried in the blood as ketones. The process of generating ketones is covered in Chapter ??. Despite the lower rate there is a huge capacity for continued ATP regeneration with fatty acids due to the generally larger storage of fat in adipose tissue in even lean individuals, and the higher energy density of fat (approximately twice the energy yield per gram than glucose or glycogen). Fatty acids enter the muscle fiber and the mitochondria and are converted to several acetyl-CoA molecules (exact number depends on the number of carbons in the fatty acid). Acetyl-CoA is the first step to TCA which produces CO_2 for removal and $NADH$ for ETC. Once a fatty acid is converted to acetyl-CoA the pathway is identical to the pathways starting with liver or muscle glycogen. The higher energy (ATP) yield when using fatty acids is related to the number of acetyl-CoA molecules formed when using the same number of grams of fat as compared to glucose or glycogen.

The RER when fatty acids are utilized tend to be less than 0.75 (the exact RER varies depending on the fatty acid utilized). A mixed diet at rest typically results in an RER $\approx$ 0.85. A predominantly glucose (carbohydrate) diet will increase the RER toward 1.0. A fat based (keto) diet lowers the RER [Ale+15].

6.2.5.1 Pathways for Protein Energetics

There are a variety of pathways for protein energetics, but they are not an optimal source of energy and not the best use of protein. The pathways are reserved for extreme situations. These pathways involve protein breakdown into amino acids and the use of amino acids at various steps of other pathways (requires removal of nitrogen). The two primary paths for protein involve the use of amino acids to produce acetyl-CoA, and in the liver for gluconeogenesis. Gluconeogenesis is the generation of glucose from non carbohydrate carbon sources including amino acids (as well as other molecules such as lactate, pyruvate, acetyl-CoA and fatty acids).

6.2.6 Rates and Capacities

Rates and capacities of ATP regeneration for each of the pathways are included in Table 6.1.

When considering the time estimates based on the quantity of ATP can be regenerated by various energetic pathways it is important to consider that two factors influence the sustainability of the pathway. First whether the pathway, at or near its max rate, is rate limiting. Second, the availability of the substrate (macro-nutrients) in storage. The time estimates in Table 6.1 are calculated based on the assumption of the max rate. For example, the 10 second estimate for ATP/PCr is based on the max rate of ATP regeneration from ATP/PCr. If the demand for ATP by the muscle is 2.2 mol/min then

Source	Max Rate of ATP (mol/min)	Amount of ATP (mol)	Time at Max (s or min)
ATP/PCr/Myokinase	4.4	0.7	10 s
Muscle Glycogen $\rightarrow$ Lactate	2.4	1.6	40 s
Muscle Glycogen $\rightarrow CO_2$	1.0	84	84 min
Liver Glycogen $\rightarrow CO_2$	0.5	30	60 min
Fatty Acids $\rightarrow CO_2$	0.4	4000	10,000 min

Table 6.1: Summary of Max Rates of ATP Regeneration by Pathway (Data from [Feh17])

the time that ATP/PCr could contribute would be approximately 19 seconds. If the demand for ATP by the muscle is 1.1 mol/min then the time that ATP/PCr could contribute would be approximately 38 seconds. By this time the muscle glycogen $\rightarrow$ lactate pathway could be contributing and meeting most of the ATP demand of the muscle.

Table 6.2 provides estimated rates of ATP consumption and the amount of ATP utilized for a variety of timed running events. Comparing this data to Table 6.1 allows the reader to consider which energetic pathways are capable of regenerating ATP at rates high enough for the various events.

Events	Rate of ATP Consumption (mol/min)	Amount of ATP (mol)	Estimated Time (s or min)
Rest	0.12	173	One Day
100 Meter Sprint	2.8	0.5	10 seconds
800 Meter Run	2.0	3.4	192 seconds
1500 Meter Run	1.7	6	210 seconds
42,000 Meter Run	1.0	150	150 minutes

Table 6.2: Estimated Rate of ATP Consumption during Various Timed Running Events (Data from [Feh17]). Resting data is estimated based on basal energy expenditure of 1200 calories, and a conversion rate of about 6.93 calories per mole of ATP.)

A 100 meter sprint requires a mix of ATP/PCr and muscle glycogen $\rightarrow$ lactate pathway. The 800 meter and 1500 meter runs require a mix of ATP/PCr, muscle glycogen $\rightarrow$ lactate and muscle glycogen $\rightarrow CO_2$ pathways. The 42,000 meter run will rely on the slower rate but higher capacity pathways of muscle glycogen $\rightarrow CO_2$, liver glycogen $\rightarrow CO_2$ and may require some fatty acid to CO_2 depending on the exact quantity of ATP capable of being generated by muscle and liver glycogen. There is debate about the ability of fat adaptation to improve the performance of endurance athletes. No one questions whether there is a higher overall capacity (storage) when using fat as the substrate of choice, such as with fat adaptation and a low carbohydrate diet. The question is whether the fatty acid $\rightarrow CO_2$ pathway is capable of providing the same, or at least close enough, rate of ATP regeneration as the muscle and liver glycogen pathways [McS+18].

6.3 Motor Unit & Muscle Fiber Types

Chapter 5 discussed the fast and slow twitch characteristics of motor units and muscle fibers based on the concepts of excitation and activation. There are also energetic differences between motor unit muscle fiber types. These characteristics can be inferred by the labels "O" for oxidative and "G" for glycolytic. The energetic differences between the three classifications of muscle fibers are provided in Table 6.3. Four energetic characteristics are compared based on high (+++), medium (++) or low (+) relative capacities of the fiber types.

Muscle Fiber Type	Mitochondria	Myoglobin	Glycolytic Enzymes	PCr
Fast Glycolytic (FG)	+	+	+++	+++
Fast Oxidative Glycolytic (FOG)	++	++	++	++
Slow Oxidative (SO)	+++	+++	+	+

Table 6.3: Energetic Characteristics of Muscle Fiber Types

Slow Oxidative (S/SO/Type 1)

SO fibers have a greater number of mitochondria and more myoglobin than FG and FOG fibers. These fibers are better equipped to sustain ATP production. It would be expected that muscles with more SO fibers would have higher rates of ATP production utilizing the energetic pathways that rely on TCA and ETC. This supports the ability to provide tension for a longer period of time.

Fast Glycolytic (FF/FG/Type 2a)

FG fibers have a greater number of glycolytic enzymes and more PCr available which maximizes the anaerobic (high rate of ATP production outside of the mitochondria) pathways. They are higher rates of ATP production for the PCr and muscle glycogen to lactate pathways, which result in higher rates of ATP production overall. However, these high rates of ATP regeneration are not sustainable. This supports the higher tension capability of these fibers. Based on the ATP-PCr shuttle (Figure 6.3) the elevated PCr throughout FG muscle fibers can quickly propagate ATP regeneration through the muscle fiber to the areas most needed, and the regeneration of PCr from ATP released from the mitochondria during recovery.

Fast Oxidative Glycolytic (FR/FOG/Type 2x)

FOG fibers have a moderate number of all energetic characteristics. They can attain more tension than the SO fibers but cannot sustain it as long as SO fibers. They can sustain tension longer than FG fibers, but cannot attain as high a tension as FG fibers.

6.4 Clinical Physiology Connections**6.4.1 Sarcopenia**

Sarcopenia is an age related loss of muscle fibers with a bias toward FG fibers (FF motor units). A consequence of sarcopenia is a loss of the ability to attain the upper ranges of muscle tension. Expected consequences are a reduction in peak tension, high forces and high velocities of movement. An unexpected, paradoxical, consequence is a reduction in the ability to sustain lower ranges of muscle tension such as those required during activities of living such as walking. This paradoxical consequence is based on the need to recruit the remaining motor units with increased frequency to meet the tension requirements of movement. Recruiting a slow motor unit and its SO fibers more frequently (higher frequency summation) requires a higher rate of ATP than can be sustained. There is no question that S motor units and SO muscle fibers can achieve a higher rate of ATP regeneration than FF motor units and FG muscle fibers. But at a high level of tension (relative to the motor units capacity) the higher rate of crossbridge activation requires a higher rate of ATP regeneration. If that higher rate of ATP regeneration is greater than the rate capacity of the glycogen $\rightarrow$ CO_2 pathway, then the fiber will need to utilize the glycogen $\rightarrow$ lactate pathway, which is self limiting (cannot be sustained). Let's walk through this with a simple model example. Assume walking requires 50% output (tension) of 1/4 of the motor units. At this output the fibers require 0.7 moles of ATP / minute. In this scenario the Glycolysis $\rightarrow$ CO_2 pathway is sufficient and the activity is sustainable. Now assume there has been a loss of motor units. Now there are 1/2 of the prior motor units, and they are predominantly lower tension generating (S/SO) because the loss (due to sarcopenia) results in a loss of FG/FOG muscle fibers. Now walking (at the same pace as before) requires 75% output from 3/4 of the fibers (since a twitch generates less tension there needs to be more frequency and motor unit summation). At this output the fibers require 1.3 moles of ATP / minute. In this scenario the Glycolysis $\rightarrow$ CO_2 pathway is not sufficient and the Glycolysis $\rightarrow$ Lactate pathway must be utilized to supplement ATP regeneration, which is not sustainable. Continuing to walk at this pace will result in fatigue. That is the paradoxical consequence of the loss of FF motor units in sarcopenia. Everyone expects a loss in the ability to attain tension as measured by peak force. But there is also a loss in the ability to sustain tension as measured by the ability to continue developing tension for an activity.

6.4.2 Fatigue

Biological systems include cycles. Ca^{+} is released and taken in, muscle fibers develop active tension and then they don't, movements happen then they don't, we are awake and then asleep. Across all scales in biological systems there are cycles. These cycles include using resources, capabilities, and reserves and then building them back up. The concept of fatigue is fundamentally about these cycles. At the most general level fatigue occurs when resources, capabilities and reserves have been utilized and it is time for the cycle to be completed and go to its recovery phase. If the cycle that leads to fatigue is not completed so that recovery can begin, then fatigue occurs.

Fatigue is a widely utilized concept because it can be applied to any aspect of a system with cycles occurring at any scale, which is any system achieving objectives. Reading about fatigue is sure to result in several different interpretations. Muscle fatigue is the inability continue to sustain a tension with repeated excitation, and with fatigue the inability to attain a previously attainable tension. Running fatigue is the inability to sustain a pace with continued running, and the inability to attain that pace until recovery has occurred. Mental fatigue influences the ability to concentrate. Cognitive fatigue influences the ability to do complex tasks. Physical fatigue can reduce the maximal heart rate that can be achieved with exertion, or a higher heart rate at a given sub maximal workload. Chronic fatigue syndrome is characterized by a severe lack of energy following exertions. The descriptions of fatigue can be like the description of the elephant by blind monks in Figure 6.4. One monk describes the elephant as like a rope (while feeling the tail), another like a snake (while feeling the trunk), another like a wall (while feeling the body), another like a tree trunk (while feeling a leg), another like a spear (while feeling the tusk), and another like a large sheet (while feeling the ear).



Fig. 6.4: Five Blind Monks Describing an Elephant (Public Domain from [From Martha Adelaide Holton & Charles Madison Curry, Holton-Curry readers, Rand McNally & Co. \(Chicago\), p. 108](#))

A general definition of fatigue is an inability of a system (In our case a biological system) to achieve objectives. That this meaning is objective. The fact that the system cannot achieve its objective is measurable in some way. This is an objective definition of fatigue that can be applied in many contexts.

There is also a subjective perspective of fatigue. The subjective perspective, at least in terms of human fatigue, is the feeling (perception) of fatigue. Sometimes people confuse the subjective and the objective. The objective perspective of fatigue includes the inability of the system to achieve some measurable objective. If the system is meeting its objectives then it is not fatigued. However, there may be a perception of fatigue, and the person can report that they are fatigued.

Both need to be considered and both pieces of information are important.

For example, if a person walking 4 mph on a treadmill starts to feel fatigued but they can continue walking at 4 mph then they are not objectively fatigued (they are meeting the objective), but they are subjectively feeling fatigued (or perhaps tired; or out of breath which can be confused symptomatically with fatigue). The symptom of (perception of) fatigue in this individual despite being able to meet the objective means that the person is not fatigued for walking at the point they are doing the walking. The perception is probably arising because there are signals about the use of a reserve capacity to alert the person that its nearly time to enter a recovery phase. It could be some supporting system for walking is fatigued and perhaps in an objective measurable (but not currently measured) component such as muscle tension. It does not take much muscle tension to walk relative to maximal muscle tension. However, with fatigue of the muscles the maximal muscle tension decreases. To continue walking there is a greater percentage of the muscle tension being utilized (since the maximal value has decreased). At some point a support system is fatigued walking cannot continue at 4 mph, so the person slows down the treadmill to 3 mph and they are meeting a new objective. At this new walking speed, the energetic requirements may be low enough that all of the support systems recover to full capability and the person can go back to 4 mph for a time being. The change in the ability - previously being able to do 4 mph without fatigue to now not being able to do 4 mph without fatigue - can be due to a large number of factors. Relevant to this chapter it may be due to the amount of substrate available, or state of the muscle fibers (perhaps there has been a loss of mitochondria due to reduced use of the muscles), or perhaps deficiency in creatine that is slowing down transport of ATP through the sarcoplasm so that the muscle glycogen $\rightarrow$ CO_2 pathway ATP needs to be supplemented with the muscle glycogen $\rightarrow$ lactate pathway. Based on the above example it is hopefully clear that the applications of the concept of fatigue vary based on time and space scales. These scales are superimposed. As we consider longer time scales all of the objectives that can occur within that time scale are important. As we consider higher spatial scales, all of the objectives at spatial scales below it are important. If walking for 4 hours at a 4 mph pace then all of the objectives within 4 hours, at all of the scales must be met to the extent that they support walking at 4 mph for 4 hours. Walking requires tension in a particular set of muscles, each muscle must generate the right tension, each motor unit contributes the tension it must contribute, each muscle fiber contributes its share of the tension which requires the fiber to be utilizing and regenerating ATP, etc. Walking today for 4 hours at 4 mph may be achieved (no fatigue for walking at 4 mph for 4 hours) with the perception of fatigue (subjective, so most likely some objectives in supportive roles were fatigued). The next day walking for 4 hours at 4 mph may not be achieved - there may be objective fatigue the next day. An activity today with no fatigue, but performed everyday, may result in fatigue in upcoming days. The point is that system must have time to recover from fatigue.

Table 6.4 provides experimentally derived estimates of fatigue recovery times for a variety of physiological components important for considering low back injury. These times were used to build a model of low back fatigue and recovery as a dynamic predictor of low back injury at work based on the work recovery exposure cycles. The basic idea is that it is not just the work demands, but the time to recovery between demands, that influences injury risk.

The perspective of objective and subjective fatigue does not intend to diminish the importance of subjective fatigue experiences. Just to keep both perspectives in mind. These perspectives give rise to the perspectives of absolute and relative workload. Walking at 4 mph is an absolute workload. If that workload can be achieved without objective fatigue at the level of the required objective for 4 hours then that absolute workload is 4 mph for 4 hours. The subjective fatigue is the relative workload, how hard did the person need to work (or push, or strain) to complete that absolute workload? Relative workload can be measured with a rating of perceived exertion (technically different but relatable to subjective fatigue), or by measures of other systems that support the absolute workload of walking such as heart rate, blood flow, ventilation rate, blood lactate, oxygen consumption, carbon dioxide production.

6.4.2.1 Energetic Fatigue

Table 6.1 provides estimates of ATP regeneration rates based on the energetic pathways. Each energetic pathway leads to muscle fatigue when its capacity (Amount of ATP) has been utilized. When its capacity has been utilized it must, as

Fatigued System	Recovery Time
Myoglobin stores	10-15 seconds to 1 minute
Conduction Velocity	5 minutes
PCr	10 minutes
ATP stores	Minutes (2-5)
Lumbar flexion creep	30+ minutes
Lactate removal	25 minute half life
Intramuscular pH	20-30 minutes
EMG Amplitudes	30 minutes
Body Height after lumbar disc compression	4 hours
Ligament laxity	7 hours
Muscle glycogen	5-24 hours

Table 6.4: Experimentally Derived Estimates of Recovery from Fatigue in Various Systems. Readers should not attempt to memorize the actual times. This data is presented to emphasize the point about fatigue from different systems across scales and times for recovery. Data is from [KW08].

a pathway, recover. The time it takes for a pathway to utilize its capacity is dependent on the rate of ATP regeneration required by the muscle fibers need for ATP. When you are preserving battery life on your phone you shut down apps in the background to minimize the rate of energy depletion and extend batter life. Subjective fatigue (the perception of fatigue) is often felt prior to complete depletion. Most likely as a warning sign to consider "shutting down some apps" prior to the point when the objectives can no longer be met due to objective fatigue.

The recovery of each system is specific to the system. For the PCr pathway, ATP is needed to regenerate PCr and can take several minutes (Table 6.4). For the muscle glycogen $\rightarrow$ lactate pathway recovery would include a fresh supply of NAD^+ from ETC and the reduction in lactate. For muscle and liver glycogen $\rightarrow CO_2$ pathways recovery would be new stores of glycogen (hours to days). Fatty acid $\rightarrow CO_2$ pathway is one of those systems that tends to function with little need for recovery for a really long time. Efforts to preserve fatty acids (a form of recovery through slower utilization) could include breaking down protein and generating glucose with gluconeogenesis in the liver. But even during sustain activity other systems, or other aspects of muscle function, are most likely going to fatigue before fatty acid energetics fatigue (such as the brain requiring sleep; or muscle tissue breakdown, or connective tissue breakdown both due to prolonged overuse, for example). Since walking is a pace that people can sustain with the fatty acid $\rightarrow CO_2$ pathway it is unlikely that the fatigue during a 20 mile walk is due to energetics. However, that does not mean that a person untrained for such a walk would not have blisters, dehydration, chaffing and several other problems due to various systems being fatigued. How an activity influences muscle fatigue is based on the overload. The overload training concept is the stimulus for training related muscle adaptations. Training adaptations often occur in response to training fatigue in an attempt to build system capacity to avoid future fatigue.

6.4.3 Hypoxia, Hypoxemia & Ischemia

The oxidative energetic pathways provide most of the ATP for cellular functions and are critically involved in the restoration of the energetic resting state following short term rate increases in ATP production with PCr, myokinase or glycolysis. The oxidative (mitochondrial) pathways require oxygen and macro-nutrient substrate (seconds-minutes). Over longer time periods (hours-days) they require cellular synthesis of enzymes which require ATP and amino-acids (to build proteins). And over longer time periods (days-weeks) they require maintenance of the health of mitochondria and replacing damaged or dead mitochondria. Interruptions in the availability of O_2 interrupts the process of ATP production with these pathways and can lead to the accumulation of metabolic waste products that alter the pH of the cell and the extra-cellular fluid. These interruptions form the basis of a large number of relatively common and life-threatening chronic medical conditions such as heart disease, stroke, peripheral vascular disease, COVID-19, pulmonary disease, and hematologic (blood) conditions; as well as sudden onset (acute) conditions such as a heart attack and acute altitude sickness. Some of the conditions have a rapid onset and immediately threaten life; others exert their effect gradually. The variation is related to rate of onset of O_2 deprivation, the magnitude of O_2 deprivation (how much deficit, how many cells), and whether the impact is just in the availability of O_2 or whether there is also an impairment in waste product

removal. But they can all be analyzed based on an understanding of cellular energetics and the role that ATP plays in cellular fidelity, efficacy and integrity.

Hypoxia

Hypoxia refers to the situation in which there is not enough O_2 getting to cells for them to sustain the oxidative production of ATP. It is a form of fatigue since the resources of ETC are not available it cannot achieve objectives (ETC cannot generate enough ATP because there isn't enough O_2). Hypoxia can be caused by a variety of situations. It is a local condition because it depends on local O_2 levels and local O_2 needs. Local O_2 needs are dependent on the local need for ATP (metabolism). Cells of the body that have relatively high and constant O_2 needs, and are therefore more susceptible to hypoxia, are the heart and brain. Metabolism is always related to cellular activity and the cells of these two organs generally have a higher resting metabolism than other cells. Muscle metabolism can far exceed that of both heart and brain, that is only during periods of high muscle activity which require high levels of ATP production. Common causes of hypoxia are hypoxemia and ischemia.

Hypoxemia

Hypoxemia is a specific situation in which the blood isn't carrying adequate oxygen to body tissues. Common causes of hypoxemia include pulmonary and blood conditions (for example, obstructive pulmonary disease and anemia) or environmental conditions (altitude, carbon monoxide).

Hypoxemia can cause hypoxia. The severity of hypoxia in a cell caused by hypoxemia is dependent on the severity of hypoxemia as well as the energetic activity of the cell (ATP regeneration rate which determines O_2 demand), which fluctuates based on several factors. In someone with mild hypoxemia there may be no hypoxia in cardiac or skeletal muscles. However, with exertion that increases cardiac and skeletal muscle activity and thus need for O_2 (O_2 demand) there may be hypoxia. In these situations cardiac and skeletal muscle function will be impaired by the lack of O_2 . The cellular adjustment will be to provide ATP using a higher rate of anaerobic pathways which is not sustainable. The by-products of these activities decrease the cellular and extra-cellular pH which further impair ATP regeneration. The acidosis and reduced ATP relative to need can impact the ability of the cells to repolarize, reduce the frequency of excitations, reduce the pumping of Ca^{+} back into the sarcoplasmic reticulum, reduce the rate of myosin head release. The consequences of these changes include lower tension production, spasm and potentially damage to the cell membrane. But typically, if the problem originates with hypoxemia that is adequate for resting levels of O_2 demand then simply ceasing the activity will restore balance and not result in damage to the cell membrane.

Ischemia

Ischemia is a specific situation in which blood supply to cells is reduced. The extent of cellular involvement depends on the extent of the reduction. For example, if an entire artery is impacted than an entire limb, or muscle can be involved. If the reduction occurs in capillaries then the reduction in blood flow is to far fewer cells. Common causes of ischemia are atherosclerosis, arteriosclerosis,¹ blood clots (arterial thrombosis or embolus, or in the case of pulmonary blood flow venous thrombosis or embolus), blood vessel spasm, and micro-circulatory inflammation (a clinical manifestation of hypoxia itself and seen in COVID-19).

Ischemia can cause hypoxia. The severity of hypoxia in a cell caused by ischemia is dependent on the severity of ischemia as well as the metabolic activity of the cell (O_2 demand). A sudden and complete blockage of blood flow to cells, with no alternative pathways to provide blood flow to the cell, is a serious situation that results in cellular death due to the inability to produce ATP for cell membrane functions (sudden complete hypoxia) and to remove waste products. The combination of these two situations results first in reversible damage to the cell membrane and then to irreversible damage to the cell membrane. Without the cell membrane the cell has lost its integrity. Less extreme reductions in blood flow can create a wide variety of hypoxic and waste removal situations that allow sustained but reduced function for a

¹ Arteriosclerosis refers to thick and stiff arteries that can restrict blood flow. Atherosclerosis is a type of arteriosclerosis that includes buildup of fats, cholesterol and other substances in and on artery walls (plaque). The subsequent reduction in vessel diameter limits blood flow, and if a plaque becomes an embolus it can lodge and completely block blood flow (blood clot).

cell (resulting in long term problems in cell maintenance), and reduced function of the cells. For example, a limit on how much activity the cardiac or skeletal muscle can perform prior to having hypoxia. It is common to simply refer to such situations as ischemia (which means local blood flow is reduced and O_2 demands are high enough to cause hypoxia).

Given the variable nature of blood flow supply and cellular O_2 demand there are two situations that can arise for cardiac muscle cells in particular. Stable ischemia refers to the situation that blood flow is sufficient for resting O_2 demand, but not sufficient during elevated O_2 demand (increased cardiac muscle activity). The situation is considered stable because simply reducing cardiac activity will reduce cardiac O_2 demand and restore balance to allow recovery. Unstable ischemia refers to the situation that blood flow is not sufficient for resting O_2 demand. Unstable ischemia is unstable because balance cannot be restored by reducing the cardiac muscle activity back to rest since it is the resting demand that cannot be met. In such situations blood flow must be restored, or cardiac muscle O_2 demand must be reduced below resting levels. Restoring blood flow is highly situational and can involve breaking up a blood clot (thrombus) with medications (thrombolytic therapy); or restoring the diameter of blood vessels with an angioplasty or by re routing the blood (by-pass graft). Reducing cardiac muscle O_2 demand can be accomplished by lowering blood pressure (blood pressure is the resistance that the cardiac muscle must work). Nitroglycerine is a very powerful and fast dilator of blood vessels that quickly lowers blood pressure and allows the cardiac O_2 demand to be lowered below resting values. The hope is this restores balance between O_2 supply and O_2 demand and the cells to recover before damage.

A consequence of ischemia that is not present in hypoxemia is the removal of waste products. Since blood flow to a region of the body does two things for the local extra cellular fluid - deliver O_2 and nutrients and remove waste - ischemia results in a loss of both. Therefore, the reduction in pH associated with prolonged anaerobic energetic pathways without sufficient aerobic pathway contribution is worse with ischemia than with hypoxemia. With hypoxemia the lactate in the surrounding extra-cellular fluid is removed. But with ischemia removal is impaired.

Hypoxia Summary

Hypoxia is the basis of the homeostatic imbalances caused by many conditions and diseases. It is based on the cellular requirements for ATP, which are based on the mitochondrial requirements for O_2 . Cells can produce ATP without O_2 , however they cannot sustain the production of ATP without O_2 . Understanding these mechanisms and that of hypoxia offers a wellspring of conceptual insights for many diseases and pathophysiological conditions that go well beyond the above discussion. Hypoxia caused by hypoxemia or ischemia continues to be topic for several of the upcoming chapters on Muscle Support.

6.5 Summary

ATP is critical to muscle fibers and all cells. Muscle fibers have limited ATP reserves and can experience a 60 to 100 fold increase in the utilization and therefore need to regenerate ATP. There are several energetic pathways for the regeneration of ATP that are all always functioning. Each pathway contributes different max rates of ATP production and different capacities. The max rate of ATP regeneration and the overall capacity of ATP regeneration of the energetic systems are inversely proportional. The pathways can be generally classified as anaerobic (occurring outside the mitochondria and not using O_2 or producing CO_2) and aerobic (occurring inside the mitochondria and using O_2 and producing CO_2). The anaerobic pathways have higher rates of ATP regeneration and lower capacities, whereas the aerobic pathways have lower rates of ATP regeneration and higher capacities. The three muscle fiber types are each uniquely adapted to three different combinations of energetic characteristics. These adaptations are well matched to the tension generating capabilities of these fibers as discussed in Chapter 5.

Sarcopenia results in the loss of predominantly FG muscle fibers. Combining the recruitment required to meet tension needs it becomes possible to explain the paradoxical reduction in endurance in people with sarcopenia.

Fatigue is objectively the inability of a system to accomplish its objective(s). Fatigue occurs when a system resources, capabilities or reserves are used up and must be recovered. This pattern of using up and then restoring resources, capabilities or reserves is cyclic. The cycles of fatigue and recovery occur at all levels of biological systems, and across several time scales. Subjective fatigue refers to the perception of fatigue (feeling fatigue in some way). Subject fatigue is based on the sensations that arise when system resources, capabilities or reserves are being depleted whether or not

there is objective fatigue. Energetic fatigue of muscles is based on the ability of the energetic pathways to continue to regenerate ATP.

Hypoxia occurs when there is not enough O_2 for ETC to function normally and therefore provide enough ATP. It is based on O_2 availability and how much ATP is needed by the cell. It is a form of fatigue since the resources of ETC are not available it cannot achieve its objectives (ETC cannot generate enough ATP because there isn't enough O_2). Hypoxemia is a reduction in O_2 in the blood and therefore can cause hypoxia. Ischemia is a reduction of blood flow to cells and therefore can cause hypoxia. Ischemia additionally disrupts the removal of cellular waste products and therefore the maintenance of the local extra-cellular fluid surrounding the ischemic cell.

Next Steps

To this point the highlight has been the muscle of the muscle centered approach to clinical physiology. Muscle tension produces the forces necessary for movement. Muscle tension is produced with excitation and activation. Excitation, and therefore activation and tension, are regulated by motor units. Muscle tension, activation and excitation all require ATP and therefore the continuous regeneration of ATP. These processes require a particular intra and extra cellular environment for the muscle fiber. The extracellular electrolytes and availability of substrates such as glucose and fatty acids, the availability of O_2 and removal of CO_2 all require extra cellular maintenance. The concept of fatigue and the situations causing hypoxia often include what is happening in the muscle fiber, and also what is happening just outside of the muscle fiber.

We now proceed to the physiological processes and systems that support the muscle as a tension generating system. These systems support the creation of muscle tension by supporting the extra cellular fluid. The first step is Chapter 7 on Micro-circulation.

References

- [Ale+15] Rubini Alessandro et al. *Effects of Twenty Days of the Ketogenic Diet on Metabolic and Respiratory Parameters in Healthy Subjects*. en. Lung 193.6 (Dec. 2015), pp. 939–945. ISSN: 0341-2040, 1432-1750. DOI: [10.1007/s00408-015-9806-7](https://doi.org/10.1007/s00408-015-9806-7). URL: <http://link.springer.com/10.1007/s00408-015-9806-7> (visited on 06/18/2022).
- [Feh17] Joseph J Feher. *Quantitative human physiology: an introduction*. Academic press, 2017.
- [JRD06] David Jones, Joan Round, and Arnold De Haan. *Skeletal muscle: from molecules to movement*. Vol. 40. 2006.
- [KW08] S. Krajcarski and R. Wells. *The time variation pattern of mechanical exposure and the reporting of low back pain*. en. Theoretical Issues in Ergonomics Science 9.1 (Jan. 2008), pp. 45–71. ISSN: 1463-922X, 1464-536X. DOI: [10.1080/14639220600758720](https://doi.org/10.1080/14639220600758720). URL: <http://www.tandfonline.com/doi/abs/10.1080/14639220600758720> (visited on 06/16/2022).
- [McS+18] Fionn T. McSwiney et al. *Keto-adaptation enhances exercise performance and body composition responses to training in endurance athletes*. en. Metabolism 81 (Apr. 2018), pp. 25–34. ISSN: 00260495. DOI: [10.1016/j.metabol.2017.10.010](https://doi.org/10.1016/j.metabol.2017.10.010). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0026049517302986> (visited on 06/18/2022).
- [TN08] Susanna Törnroth-Horsefield and Richard Neutze. *Opening and closing the metabolite gate*. en. Proceedings of the National Academy of Sciences 105.50 (Dec. 2008), pp. 19565–19566. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.0810654106](https://doi.org/10.1073/pnas.0810654106). URL: <https://pnas.org/doi/full/10.1073/pnas.0810654106> (visited on 06/18/2022).

Sample Questions

1. Pyruvate is the entry point of which energy substrate into which energetic pathway?
2. Why is glycolysis is self-limiting?

3. Objective fatigue during an activity that relies on the maximal rate of ATP generated by the Muscle ATP/PCr/Myokinase pathway is most likely to occur in how long?
4. Following a CVA (stroke, cerebrovascular accident) there is a loss of motor units to affected muscles. In this situation, what type of summation would be relied upon more for tension regulation, and how would the demand for ATP in a muscle fiber be impacted?
5. A person is doing their usual treadmill workout after a week of being sedentary with the flu. They usually walk at 4 mph at a 4% grade for 30 minutes with an average HR of 135 bpm and an RPE of 3/10. If walking at that intensity for that duration is the objective, what situations would indicate subjective fatigue? What situations would indicate objective fatigue?

Part III

Muscle Support

Muscle Support Overview

The chapters of Part III are about the supportive environment of the extracellular fluid (ECF). Every chapter describes a set of functions that serve the purpose of delivering, removing or regulating the contents of the ECF.

Figure 6.5 is a graphic overview of the ECF supporting functions. Filtration relies on micro-circulation which relies on circulation. Circulation is the flow of blood through the entire circulatory system. Circulation of blood through the entire body ensures regular mixing of the ECF contents throughout the body. This mixing ensures consistency of ECF contents and a stable environment for all the cells of the body. For example, if blood Na^+ concentration is tested, it is properly assumed that concentration is the concentration of Na^+ throughout the entire ECF. Circulation also allows each supporting system contribute - either to adding water, heat, ions, molecules or nutrients to the ECF, or removing by-products from the ECF. Supporting systems are also in communication with one another based on the circulating substances and messages (hormones). For example, if a muscle working at high tension is shuttling lactate into the ECF, it is picked up by the circulation and can then enter other cells (heart, liver, non working muscles) where lactate levels are low. In these cells the lactate can be used to produce an $NADH$ from NAD^+ , and a pyruvate to proceed into the mitochondria.

The contents of blood, and therefore ECF, is regulated and supported by renal filtration (kidneys) and the endocrine system. The regular circulation of blood through the kidneys involves several mechanisms that selectively filter the blood and in doing so regulate the ECF, and therefore the ICF volume. The regulation of fluid volume includes the regulation of several important ions (electrolytes) and the removal of waste (i.e. urea). Working with the endocrine system the kidneys provide long term regulation of blood pressure, which is important for circulation; as well as red blood cells, which is important for oxygen carrying capacity.

Cellular respiration in the mitochondria is constantly using O_2 and producing CO_2 . These molecules are called gases (blood, alveolar) because their natural state is in the gaseous form. The micro circulation ensures the regular delivery (O_2) and removal (CO_2) of these gases from the interstitial fluid. The circulation to the lungs ensures the regular exchanges of these gases with the environment. Blood gases rely on blood flow through the lungs for respiration (gas exchange) between the blood and alveoli. Respiration depends on ventilation, the regular movement of these gases into and out of the lungs with the bulk transport of air.

Maintaining blood nutrients relies on digestion, absorption, metabolism and elimination by the gastrointestinal organs and the liver. Based on the basic physiological concept of mass balance the regular elimination of water requires the regular intake of water. The regular transformation of the carbon bond energy from carbohydrates and fats requires the regular intake of carbohydrates and fat. The regular use, breakdown and elimination of proteins for cellular functions requires the regular intake of protein.

Together these systems support the ECF and therefore muscle fibers and muscle function. They maintain homeostasis of several important electrolytes (ions, including H^+ as measured by pH), nutrients, molecules, water and heat (temperature). The systems are integrated with several interactions and redundancies to provide support across a wide range of muscle demands in a wide range of circumstances, both good and bad.

One difference between Part II and Part III of the book is that many topics throughout the chapters in Part III are inherently *Clinical Physiology Connections*. Since the concepts described in Part III are traditional concepts in a clinical physiology course, in many ways these chapters are, in their entirety, clinical physiology connections. As the occasion arises, in pace of *Clinical Physiology Connections* there will be *Muscle Connections*.

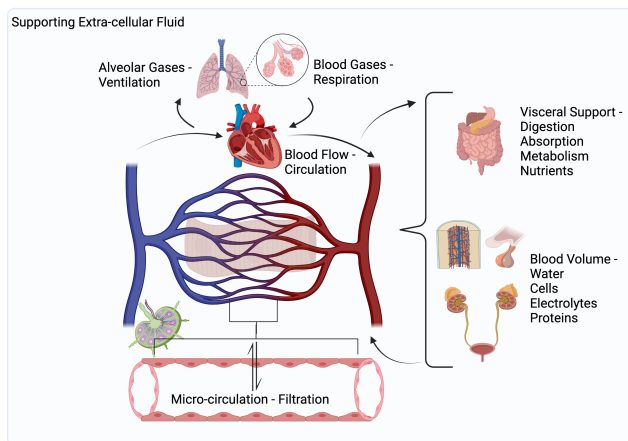


Fig. 6.5: Part III: Muscle Support Overview (Created with BioRender.com)

Chapter 7

Micro-Circulation

Updated on May 22, 2024

Contents

7.1	Micro Circulation	98
7.2	Filtration	104
7.3	Muscle Connections: Smooth Muscle	110
7.4	Clinical Connections: Compression	112
7.5	Summary	112

This chapter focuses on the process of micro circulation of water and filtration of ions, molecules, nutrients and heat between the vascular compartment of the extracellular fluid and the interstitial compartment (the non vascular compartment of the ECF). Micro-circulation supports the ECF through a continuous circulation between its two compartments, vascular and interstitial. Water, ions, molecules and nutrients are constantly circulating between the vascular (capillaries and lymphatics) and interstitial compartments of the ECF.

The extracellular fluid (ECF) supports muscle fibers by having the correct resources necessary for muscle function such as the correct concentration of ions, the availability of nutrients and molecules such as oxygen. The ECF is also where things moved from the muscles end up until they can be metabolized or removed from the body. Exchange between the ECF and intra-cellular fluid (ICF) is selective for ions, molecules and nutrients but comparatively free for water. The function of muscle fibers depends on the homeostasis of ECF oxygen, carbon dioxide, sodium, potassium, calcium, water, pH, and temperature.

Objectives include:

1. Explain the basic dependency of muscle fibers on the extra-cellular fluid.
2. Explain the basic dependency of extra-cellular fluid on the micro-circulation.
3. Explain the difference between vascular, interstitial and intra cellular fluid.
4. Explain the role of osmosis, osmolarity and osmotic pressure in micro circulation between ECF-ICF and various situations leading to intra-cellular edema.
5. Explain the components, roles and implications of the parameters of the filtration equation to explain filtration.
6. Use the filtration equation to explain various causes of extra-cellular edema.
7. Compare and contrast smooth muscle and skeletal muscle.
8. Explain the underlying mechanism for the effectiveness of compression in edema management.

7.1 Micro Circulation

Micro-circulation is the continuous movement of ions, nutrients and molecules between the blood plasma (vascular part of ECF) and interstitial fluid. Micro-circulation occurs through the semi permeable capillary membranes in a process called filtration. The capillary membranes allow relatively free passage of water, ions, molecules and nutrients (other than proteins). The contents of plasma and interstitial fluid is similar other than cell and protein content (particularly red blood cells and albumin), due to an inability of these larger items to pass through the capillary membrane. There are also differences in O_2 and CO_2 due to the constant exchange with the cell.

Normal values of several important ions, molecules and nutrients are depicted in Figure 7.1. Vascular and interstitial values are typically equal unless otherwise noted. O_2 and CO_2 pass easily through both the semipermeable capillary membrane and the semipermeable sarcolemma. There is a relatively large gradient for these molecules that varies due to the muscle fiber energetics. Normal function of circulation, respiration and ventilation at sea level are capable of equilibrating the vascular O_2 and CO_2 content even extreme high energetic situations.

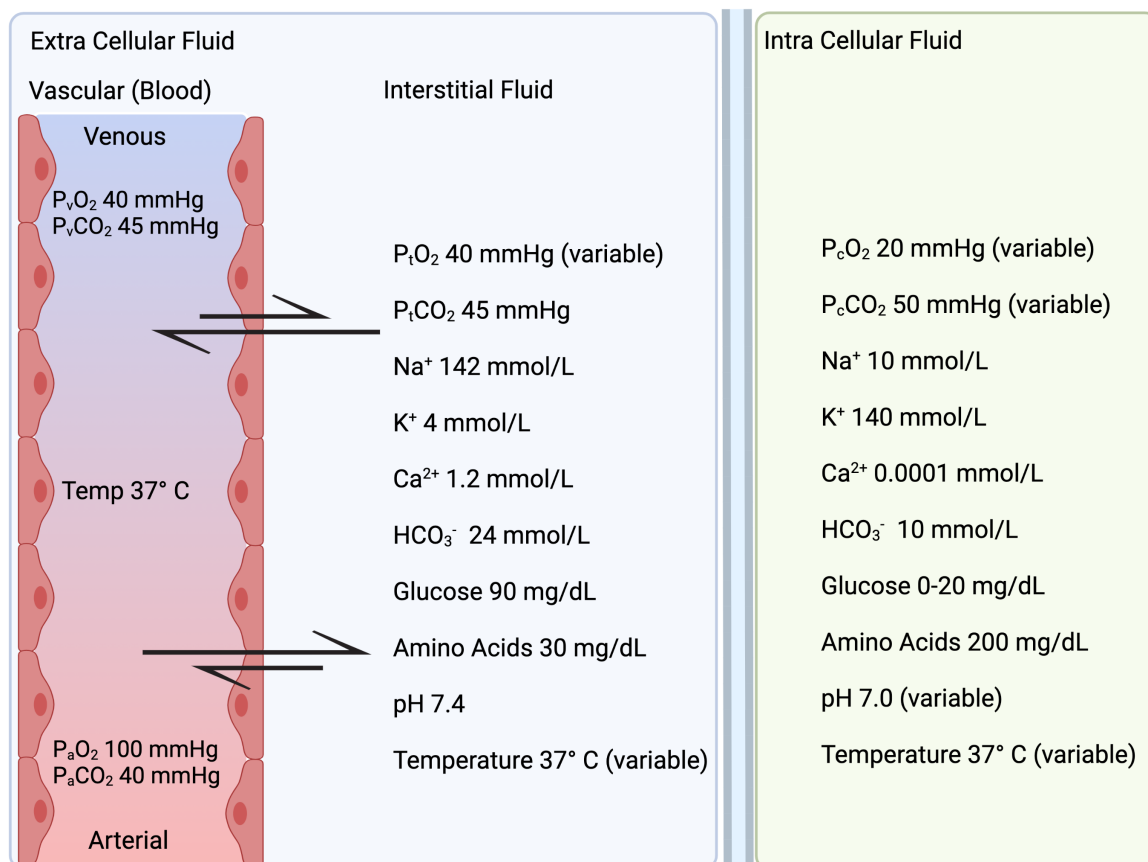


Fig. 7.1: Extra-cellular (Vascular & Interstitial) and Intra-cellular Fluid Created with BioRender.com

7.1.1 Extra Cellular - Intra Cellular Fluid Movement

Exchange between the interstitial fluid and intracellular fluid occurs through the semi permeable cell membrane (sarcolemma). The sarcolemma restricts movements of ions and even with changes in the permeability for Na^+ and K^+ during excitation does not change the concentration difference between the interstitial and ICF fluid for these ions (at

least under normal circumstances). The sarcolemma regulates the movement of nutrients. Glucose uptake by a cell is influenced by the activation of the GLUT4 receptor channel for glucose by insulin.

Between 50-70% of an adult body is water. The variation in the percent of the adult body that is water is based on variations in lean (mostly muscle) tissue. In individuals with more muscle (compared to adipose tissue) there is a higher percent of water. Of the total body water approximately 2/3 is intracellular (ICF) and 1/3 is extra cellular (ECF). Since water moves freely between the interstitial and ICF due to a large quantity of aquaporins (water channels) in the cell membrane the balance of ICF and ECF water volume is based largely on solute concentrations and resultant balance of osmotic pressure.

7.1.1.1 ECF - ICF Water Volume Balance

The sarcolemma is freely permeable to water but not to the ions. Ions require passage through protein channels. The osmolarity between the ECF and ICF determines whether there is net water movement by osmosis and the water (fluid) volume shifts between ECF and ICF. Under steady state conditions the osmolarity between the ECF and ICF is equal. Changes in solute concentration (ions, molecules, nutrients) in the ECF and ICF changes the osmolarity and results in movement of water until osmolarity balance is achieved.

7.1.2 Osmosis, Osmolarity & Osmotic Pressure

At this point most people benefit from a review of these terms and the associated process of osmosis and osmotic pressure. Osmosis is the net movement of water across a selectively permeable membrane caused by a concentration difference across the membrane. Osmotic pressure is the pressure required to prevent osmosis of water through a membrane that is permeable to water but not to the solute (similar concept to an equilibrium potential). The osmotic pressure is an indication of how quickly osmosis occurs (driving force); and osmosis does not occur if there is no osmotic pressure (or osmolarity) gradient. Osmotic pressure is exerted by particles and is determined by the number of particles per unit volume of fluid. Water moves from low concentration (low osmolarity) to a high concentration (high osmolarity).

Osmotic pressure (π) is directly proportional osmolarity difference (concentration in Osm in a volume (L) of solvent) (C); the ideal gas constant (R); and absolute temperature ($K = 310^\circ$ at body temperature) in the equation: $\pi = C \cdot R \cdot T$. Since R is a constant, and K in the body varies within a relatively narrow boundary, the osmotic pressure is determined by the osmolarity (Osm/L) difference on each side of a membrane permeable to water, but not permeable to all solute in a solution. At body temperature each 1 mOsm/L (milli osmole per liter) difference results in approximately 19.3 mmHg of osmotic pressure. In Figure 7.2 the solution on the left has 1 mOsm/L more osmolarity. At body temperature this means 19.3 mmHg of pressure is required to stop the movement of water. Differences in osmolarity due to differences in solute concentration across the sarcolemma results in osmosis of water until there are no differences in osmolarity.

Since water moves freely across the sarcolemma osmosis occurs and balances the osmolarity of the ICF with that of the ECF. When the osmolarity of ECF and ICF are equal there is no osmosis or net movement of water across the membrane since the osmotic pressures cancel each other out.

The total intake of water and ions (electrolytes) are carefully matched by equal outputs from the body to prevent fluid volumes and ion concentrations from fluctuating beyond acceptable ranges (mass balance, and homeostasis). There are times when there are differences between intake and output of water and ions that eventually must be remedied. Small shifts in water between ECF and ICF ensure that both compartments have the water required and maintain appropriate ion concentrations.

These shifts typically start with changes to the ECF. It is usually the ECF that we are adding, or removing, water and ions to, or from, (ingestion, absorption, renal filtration). Water and ions are also removed from ECF in the kidneys, the colon (small amount under normal circumstances) and water is removed from ventilation. These changes in water volume change the concentration and therefore osmolarity of ECF.

The response to changes in osmolarity between the ECF and ICF is movement of water by osmosis until there is once again equalize the osmolarity. Large water intakes to the ECF such as large water intake or IV infusions; or decreases (dehydration) associated with sweating, GI fluid loss, or excessive urine formation by the kidneys (i.e. in diabetes mellitus) change the ECF water volume. If these changes alter the osmolarity of the ECF then water will pass between ECF and ICF until osmolarity is equalized.

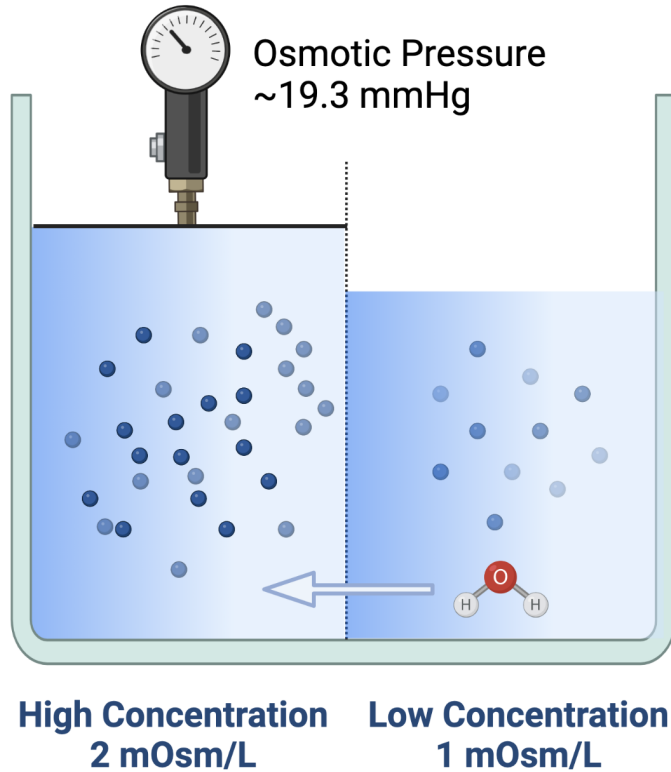


Fig. 7.2: Osmotic Pressure. It is important to realize that osmotic pressure is the consequence of water movement from a low to a high concentration gradient through a semi-permeable membrane. The membrane is permeable to water and perhaps some solutes, but not to all solutes. It is not the movement of particles from low concentration to high concentration when a membrane is permeable to those particles. Nor is it movement from low pressure to high pressure. (Created with BioRender.com)

Figure 7.3 depicts the changes that would occur to the osmolarity with the addition of a volume of isotonic (same osmolarity), hypotonic (lower osmolarity), and hypertonic (higher osmolarity) solution to the ECF.

- **Top:** Adding an *isotonic solution* to ECF does not change osmolarity of the ECF so there is no osmosis. The end result is an increase in ECF volume but no change to ICF volume.
- **Middle:** Adding a *hypertonic solution* to ECF increases the osmolarity of ECF and results in osmosis of water out of the cells (ICF) and decreases ICF volume. This continues until the osmolarity of both compartments are equal. The overall effect is decreased ICF volume and increased ECF volume; and an increase of osmolarity of both ICF and ECF.
- **Bottom:** Adding a *hypotonic solution* to ECF decreases the osmolarity of ECF and results in osmosis of water into the cells (ICF) and increases ICF volume. This continues until the osmolarity of both compartments are equal. The overall effect is increased ICF and ECF volume, and decreased ICF and ECF osmolarity.

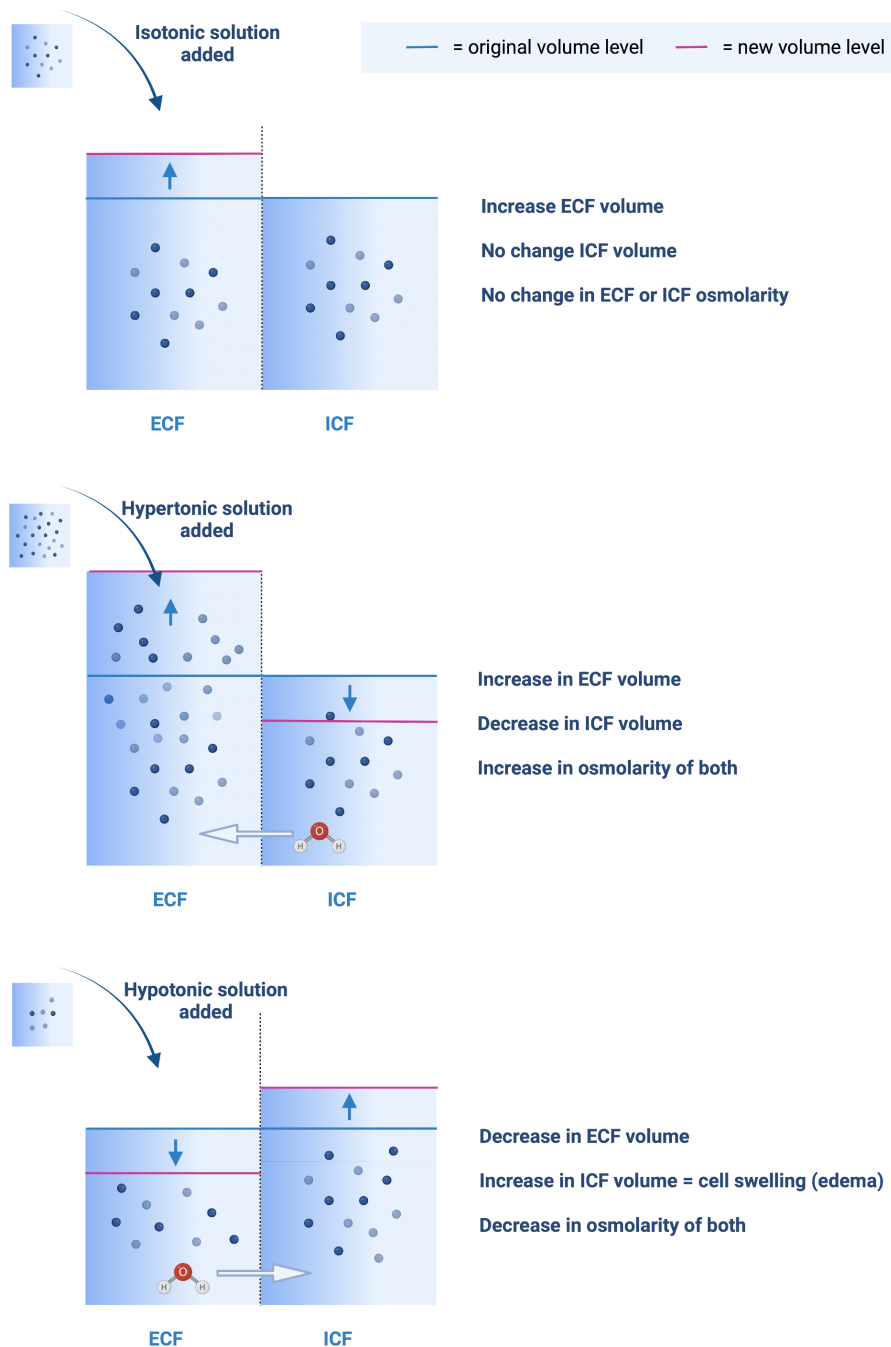


Fig. 7.3: Impact of ECF Infusions of Iso, Hyper and Hypotonic Solutions. (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

7.1.2.1 Importance of Osmosis to Muscle Excitation

Changes in the osmolarity of the ECF and ICF are based on the concentration of these solutions. The concentration changes with changes to the water volume or changes to the amount of solute. A hypertonic solution increases the concentration of Na^+ in both the ECF and ICF, but retains the concentration gradient. A hypotonic solution decreases the concentration of Na^+ in the ECF and ICF, but does not change the concentration gradient. Since the concentration gradient of Na^+ (and all ions) across the sarcolemma is critical to excitation, it should be clear that these osmotic

pressure directed fluctuations to water volume serve an important buffer capacity to maintain these ion concentration gradients. Dehydration, a reduction in water volume, ultimately results in a slightly higher concentration of ions, but this higher concentration is balanced between the ECF and the ICF, so gradients are kept constant. Ion abnormalities exist (electrolyte imbalances), but they do not have as profound an impact as they would if water did not move freely between ECF and ICF to normalize concentration gradients.

7.1.3 Intra Cellular Edema

Intra cellular edema is when there is increased fluid in the intra cellular space. Edema of the ICF can occur locally or systemically. As described above, ICF edema can occur with the addition of a hypotonic solution to the ECF (See Figure 7.3). In general, any situation that reduces the osmolarity of the ECF can result in ICF edema. Hyponatremia (low ECF Na^+) reduces the ECF osmolarity, which results in a fluid shift from ECF into the ICF until the osmolarity between these compartment equalizes. This is an example of systemic ICF edema (effects all cells). Since this fluid shift from ECF to ICF is systemic it tends to be self limiting - the osmolarity between the ECF and ICF will eventually reach a balance.

Local depression of energetic systems which can occur due to a lack of adequate nutrition or O_2 (i.e. hypoxia) can increase the osmolarity of the cells and pull more water (locally) in from the ECF. In this situation there are two contributing problems. First, is a local depression of the Na^+/K^+ -ATPase pumps (for lack of ATP) which leads to an accumulation of Na^+ inside the cell. Second, the accumulation of lactate and H^+ in the cell cannot be balanced by the removal. Keep in mind, this situation entails depression of energetic systems due to a lack of oxygen. If that lack of oxygen cannot be remedied with a reduction in ATP demand then the imbalance will persist until the cell dies. Cell death occurs when the cell membrane is no longer intact. The above scenario includes an acid pH (damaging to the cell membrane) and cellular edema (damaging to the cell membrane). The increase in cellular osmolarity (since it is a local occurrence) does not (or cannot) lead to a balance of osmolarity between the ECF and ICF because it is not systemic. The local cells cannot pull in enough fluid to alter the osmolarity of the entire ECF.

Extracellular edema can also cause intracellular edema if the extra cellular edema lowers the ECF osmolarity. Similar to adding a hypotonic solution to the ECF or in the case above of hyponatremia, with reduced ECF osmolarity the ICF osmolarity would pull water into cells until the osmolarity is equalized. A hypotonic solution is water (isotonic is water with 0.9% NaCl or 5% glucose). Adding too much water to the ECF (drinking too much water) before adequate renal filtration and secretion can occur can lead to extra cellular edema and reduced ECF osmolarity. Note, problems with renal filtration and secretion (i.e. kidney failure) does not create this problem because there is also a build up of urea and other solutes in the ECF. Kidney failure does not cause a reduction in osmolarity of the ECF, but might cause an increase in the osmolarity of the ECF.

7.1.4 Extra cellular Values & Ranges

Table 7.1 provides normal values and ranges and non lethal limits for several important components of the ECF. The wide normal ranges and non lethal limit ranges is a good example of the degree of overall robustness. Moving outside of a range without an easy temporary explanation (such as low P_vO_2 in response to exercise) is an indication that something is wrong. But even when something is wrong, the body adjusts and even large deviations in normal values are tolerated before the changes become lethal. Several of these components of the ECF have larger non lethal range in one direction. For example, P_vO_2 can increase tremendously (if given a large amount of oxygen in hyperbaric (high pressure) systems) before it becomes lethal. The criteria for lethal for these values is death. Some of the secondary effects may lead to premature death.

For example, while a Na^+ value of 165 mmol/L is within the non lethal range. This concentration of Na^+ is buffered by the movement of water that would occur due to changes in the osmolarity of the ECF so its impact on the concentration is not as great as it may seem. But with such a high concentration it would be associated with, amongst other signs and symptoms of hypernatremia, a high blood volume and therefore high blood pressure despite attempts to maintain normal blood pressure. The elevated blood pressure can then lead to premature death. Tolerance for temperature is greater with cold than with heat. With heat proteins denature and metabolic functions cannot be catalyzed. With decreased temperature reactions slow down from reduced kinetic energy, but at least the enzymes (proteins) are functioning.

Value	Normal Value	Range	Non-Lethal Range
P_vO_2 (mmHg)	40	25-40	10-1000
P_vCO_2 (mmHg)	45	41-51	5 - 80
Na^+ (mmol/L)	142	135 - 145	115 - 175
K^+ (mmol/L)	4.2	3.5 - 5.3	1.5 - 9.0
Ca^{2+} (mmol/L)	1.2	1.0 - 1.4	0.5 - 2.0
HCO_3^- (mmol/L)	24	22 - 29	8 - 45
Glucose (mg/dL)	90	70 - 115	20 - 1500
Acid-Base (pH)	7.4	7.3 - 7.5	6.9 - 8.0
Temperature F (C)	98.6 (37)	98-98.8 (37)	65-110 (18.3-43.3)

Table 7.1: Normal Range and Non Lethal Range of Values in ECF (Data from [Feh17])

7.1.5 Oxygen & Carbon Dioxide

The gradient for O_2 from the vascular fluid into the cell exists because the cell is constantly utilizing O_2 . The gradient for CO_2 from the cell to the vascular fluid exists since the cell is constantly producing CO_2 . The other gradient for O_2 and CO_2 depicted in Figure 7.1 is from the arterial end of the capillary to the venous end of the capillary. As shown, with healthy capillaries, normal capillary blood flow and normal blood the vascular O_2 and CO_2 partial pressure equilibrate by the time blood reaches the venous end of the capillary. The cellular and interstitial partial pressure for oxygen (P_cO_2 & P_tO_2) are variable based on how much O_2 is being consumed in the muscle fiber, which is dependent on the rate of ATP regeneration occurring in electron transport (ETC). With a higher than resting ATP regeneration in ETC the P_cO_2 will drop, which drops the P_tO_2 below 40 mmHg, and then the partial pressure of oxygen at the venous end P_vO_2 of the capillary will also drop. P_vO_2 will equilibrate to P_tO_2 . However, the P_aO_2 will remain at 100 mmHg as long as the circulation and pulmonary systems are providing the additional support that is required.

Similarly, for ETC to be regenerating ATP at a higher rate, the citric acid cycle (TCA) has to be functioning at a higher rate and thus producing more CO_2 . This results in a higher P_cCO_2 . However, because P_tCO_2 and P_vCO_2 equilibrate, as long as the circulation and pulmonary systems are providing the additional support that is required there will not be an increase in either P_tCO_2 or P_vCO_2 . The difference between arterial O_2 and venous O_2 is an important indicator of how much O_2 is being consumed. How much O_2 is being consumed is an indicator of metabolism and energetic demands from rest to peak exercise.

7.1.6 Temperature

Temperature is heat unless there is a temperature of 0 degrees Kelvin. Heat is a byproduct of all energetic transformations, including regenerating ATP and hydrolyzing ATP. At rest, temperature in muscle fiber tends to equilibrate (or be slightly lower) with that of the body because the flow of heat from the fibers into the ECF is moved to the blood and out of the region. Some of the heat also dissipates through the ECF and tissues across the temperature gradient directly out of the body. When core body temperature falls, a reflex action is to shiver, which utilizes muscle activity to generate more heat for sharing with the rest of the body through the transport of ECF. When muscle temperature rises during activity, the circulation of blood carries much of that heat away from the muscle for dissipation throughout the body. As heat rises, more of the circulation is sent to the skin to facilitate this process.

Temperature of muscles varies more than core temperature. While core temperature is highly regulated at 37 degrees Celsius muscles in the extremities, and more so the distal extremities, can vary from as high as 40 degrees in a hot environment to as low as 20 degrees in the cold. Muscle function when muscle temperature is less than 37 degrees impacts the rate of tension development and the rate of tension recovery (relaxation). At the extreme of 22 degrees a the rate of tension development and recovery is approximately 25% of that at 37 degrees [JRD06]. This is thought to be due primarily to the dependency on these rates on kinetic energy of moving molecules throughout the chain of events from excitation to activation. The overall impact is that a reduction in muscle temperature has an impact on high power activities which are dependent on the rate of tension development and recovery.

7.1.7 Acid Base (pH)

Intra cellular fluid has several buffering systems for regulating pH through a wide range of energetic rates and thus do not, in resting or even low to moderate energetic situations, send H^+ ions out of the sarcoplasm for acid base balance. In situations that force the utilization of the Glycogen $\rightarrow$ Lactate pathway such as a high ATP demand or low ETC capacity, lactate (lactic acid) removes excessive H^+ ions from the sarcoplasm to the ECF which can then be circulated to other cells for recycling. But this lactate shuttling is not perfect and intra cellular pH can drop if those situations are not balanced. There are several processes involved with the regulation of ECF acid base balance (pH). Each is presented in the upcoming chapters such as buffers in the blood, filtration in the kidneys, and respiration in the lungs.

7.2 Filtration

Micro-circulation relies on filtration between the ECF vascular and interstitial compartments. Filtration occurs through capillary membranes. It includes the flow and exchange of water, ions, molecules and some nutrients (glucose, fatty acids, amino acids). Larger items such as proteins (albumin) and blood cells do not normally filtrate through the capillary membranes. However, in certain circumstances white blood cells, as part of an immune response, will leave the vascular compartment and enter the interstitial fluid.

There are two pressure gradients that drive filtration: osmotic and hydrostatic pressure. Osmolarity creates osmotic pressure that has the net effect movement of water and solutes between the vascular and interstitial compartments. The osmolarity between the vascular fluid and interstitial fluid is one driving force for filtration. The second pressure gradient is hydrostatic pressure (fluid pressure).

7.2.1 Hydrostatic pressure

Hydrostatic pressure is the force that fluids exert on the walls of the space they are in, and is related to the size of that space (volume) in and the compliance of the walls of that space. Pressure is also related to temperature. However, pressure varies much more with volume and compliance in the human body due to the relative stability of body temperature. Pressure, volume and compliance are related by the equations (C = compliance, V = volume, P = pressure, Δ = delta meaning "change in"):

$$C = \frac{\Delta V}{\Delta P} \quad (7.1)$$

$$\frac{1}{\Delta P} = \frac{C}{\Delta V} \quad (7.2)$$

$$\Delta P = \frac{\Delta V}{C} \quad (7.3)$$

Compliance is determined by the change in pressure to change in volume (Equation 7.1). It is useful to consider how changes in characteristics of a structure change its compliance and then influence Δ pressure starting with the assumption of a stable volume. Or the influence on volume with the assumption of a stable pressure. In equations 7.2 and 7.3, compliance and Δ pressure are inversely proportional. Compliance of a tissue (or compartment which in the body is created by tissue) is inversely proportional to stiffness. Fibrosis, which occurs when scar tissue replaces normal tissue in response to injury or inflammation, makes the tissue more stiff and less compliant.

The volume of all capillaries is related to the sum of their cross sectional area. The volume of one capillary is related to its cross sectional area. The cross sectional area is determined by the radius. The radius also influences resistance to flow. And flow influences the amount of fluid entering and exiting. These relationships are complicated, in part, because the radius of the capillaries influences pressure both changing cross sectional area and volume, and by changing resistance to flow. The effect of these relationships is that changing a capillary radius, or the total capillary cross sectional area, has a powerful, and complicated, effect on hydrostatic pressure in the capillaries.

7.2.2 Capillary Blood Flow

The capillaries have a thin endothelial membrane that does not include smooth muscle. Without smooth muscle the radius of a capillary is based on its structure which influences its compliance and the amount of blood flowing into and out of the capillary (volume). Blood flow into the capillaries is regulated by local and central (neuroendocrine) factors. Changing the radius of the arterioles (vessels that flow into capillaries) and pre-capillary sphincters changes the volume of blood entering the capillaries (Figure 7.4). Blood flow out of the capillaries is also regulated by local and neuroendocrine factors. The venules (vessels capillary blood flows into) have a less developed smooth muscle layer than arterioles and therefore influence less effect on the radius of the venules. Venules, and the entire venous system, have higher compliance and are described as a low pressure system. This helps to ensure that flow proceeds through the capillaries from the lower compliance (higher pressure) arterial circulation to the higher compliance (lower pressure) venous circulation. Neuroendocrine control of the capillary blood flow is from the sympathetic branch of the autonomic system. Neural and endocrine influences are exerted through the catecholamines (epinephrine and norepinephrine). The effect of catecholamines is vasoconstriction. But this effect is overridden locally as needed.

The volume of blood in the capillaries changes by changing the flow of blood into and out of the capillaries. A strong local determinant of blood flow, that overrides sympathetic vasoconstriction, are energetic byproducts that indicate the need for O_2 such as interstitial O_2 being low, CO_2 being high, a drop in pH and a rise in temperature. All of these changes, locally, promote dilation of arterioles, pre-capillary sphincters and venules. In this situation the increased blood flow generally does not increase the capillary hydrostatic pressure because flow into and out of the capillary is balanced and associated with vasodilation. If blood flow to capillaries increases the capillary blood volume beyond its normal range of compliance then hydrostatic pressure can increase which may lead to local edema.

Vasoconstriction of the vessels leading into and out of the capillaries reduces blood flow. But since it is a change in flow in and out of the capillaries the hydrostatic pressure also remains relatively stable. The overall effect of this neuroendocrine vasoconstriction and local vasodilation is delivery of blood flow to the capillaries in the body with the greatest need during fight or flight situations that lead to increased vasoconstriction overall. Though there is always a baseline amount of sympathetic vascular tone that helps maintain blood pressure. These mechanisms have an impact on micro-circulation, however they are primarily used to regulate overall circulation to ensure appropriate distribution of cardiac output to the regions of the body with the greatest energetic (and therefore O_2) demands.

Blood flow to the capillaries is well balanced (in and out) and under a wide range of conditions does not change the hydrostatic pressure within the capillaries. Situations that alter the hydrostatic pressure of capillaries that lead to edema are those that influence venous return. Even though the veins have a high compliance and accommodate a high volume with a small change in pressure, there are situations where the volume that accumulates does increase venule and then capillary hydrostatic pressure. If blood flow is slow through the venous system for any reason the elevated venous pressure is transmitted to the capillaries. This can happen in a minor way with every day situations such as standing. Standing increases the venous pressure in the lower extremities due to gravity making it more difficult for venous blood to return to the heart. This results in increased capillary hydrostatic pressure. More severe situations, such as heart failure, result in increased venous pressure which increases capillary hydrostatic pressure and is discussed in the section on extra cellular edema.

7.2.2.1 Capillary Dilation

The capillary membrane can undergo changes to the size of its openings which can allow cells that could not previously leave the vascular compartment to enter the interstitial fluid. When this occurs, such as with capillary dilation in response to cellular inflammatory mediators or other threats to the well being of cells, the white blood cells enter the interstitial space and change the osmolarity and osmotic pressure gradients. It is also possible that a high volume of blood in the capillaries, due to high pressure in the venules, will dilate capillaries. In this situation the high hydrostatic pressure tends to be more complicated by a low osmolarity as larger blood proteins of cells are able to leave the vascular compartment.

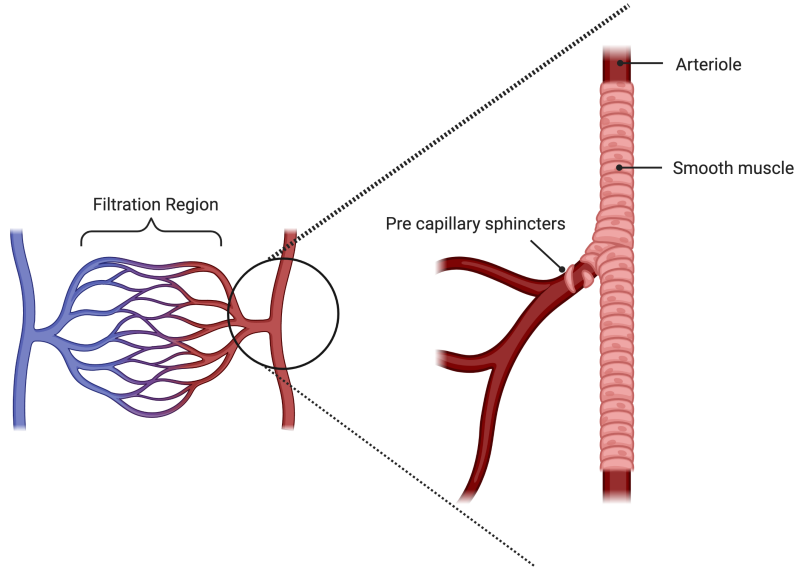


Fig. 7.4: Capillary Anatomy (Created with BioRender.com)

7.2.3 Filtration Pressures

Filtration is the normal process of exchanging vascular and interstitial fluid (and solute) for micro-circulation. When filtration does not work well there is either hypoxia due to ischemia (not enough O_2 being delivered due to no or limited capillary blood flow), extra cellular edema (excessive fluid in the interstitial space), or (rarely) increased blood volume from excessive fluid moving into the vascular space).

To promote normal filtration the overall filtration pressures (hydrostatic (P) and osmotic (π)) are slightly unbalanced. The unbalance pushes fluid and solute out of the vascular compartment and into the interstitial compartment at the arterial end of the capillary. It then tends to pull fluid and solute into the vascular compartment at the venous end of the capillary. However, the return to the vascular compartment is incomplete leaving some of the fluid and solute in the interstitial compartment. The volume of fluid and solute left in the interstitial compartment then enters the lymphatic system and reenters the vascular system after circulation and filtration through the lymph vessels and nodes (See Figure 7.5).

An equation that shows the relationship between the determinants for filtration is:

$$\text{Filtration} = K_f \times (P_c - P_{if} - \pi_c + \pi_{if}) \quad (7.4)$$

There are no constants in this equation, each parameter is a variable. The parameter K_f is the filtration coefficient.

- K_f refers to the permeability of the capillary membrane. While not constant, it is relatively stable unless there are inflammatory mediators caused by injury or infection or very high capillary volumes that cause dilation. Dilation increases the K_f by opening channels between endothelial cells of the capillary membrane. An increase in K_f does not push or pull fluid into a particular compartment. It simply allows more fluid to be moved with any particular balance of filtration pressures.
- P_c refers to the hydrostatic pressure inside the capillary, and P_{if} refers to the hydrostatic pressure inside the interstitial fluid (space). The hydrostatic pressures encourage flow from the compartment of higher pressure to lower pressure. P_c pushes solute out of the capillary and into interstitial fluid when it is higher than P_{if} . P_{if} pushes solute out of the interstitial fluid and into the capillary if it is higher than P_c .
- π_c is the osmolarity of the blood in the capillary. It exerts a pressure that pulls water into the capillary (from the interstitial fluid).
- π_{if} is the osmolarity of the interstitial fluid. It pulls water into the interstitial fluid (from the capillary).

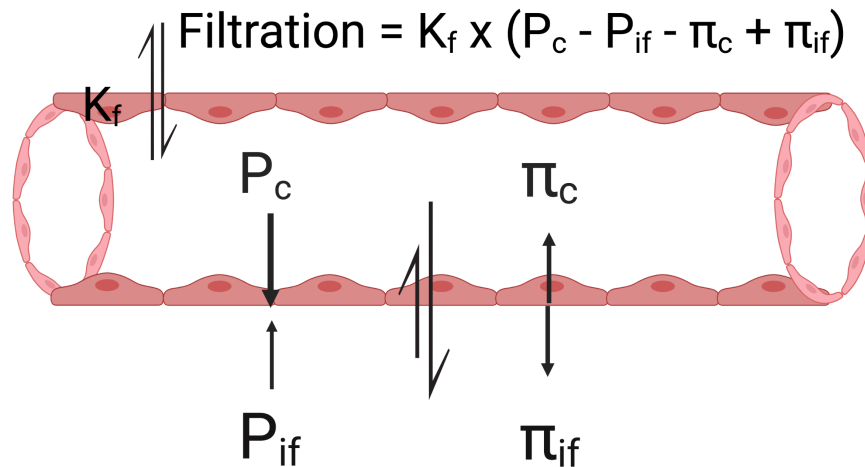


Fig. 7.5: Filtration Equation (Created with BioRender.com)

Since all four of these pressures work with or against one another the overall balance of these four pressures determines filtration.

If $(P_c - P_{if} - \pi_c + \pi_{if}) = 0$ there is no net change in vascular or interstitial volume and there would be no filtration.

If $(P_c - P_{if} - \pi_c + \pi_{if}) > 0$ there is filtration with a net gain in the interstitial space. If $(P_c - P_{if} - \pi_c + \pi_{if}) < 0$ there is filtration and a net gain in the vascular space. Under normal circumstances, across the entire capillary, $(P_c - P_{if} - \pi_c + \pi_{if})$ is just slightly greater than 0 so that there is filtration and a net gain in the interstitial space that is picked up by the lymphatic system.

7.2.4 Micro-circulation Dynamics Through the Capillary

The overall process of filtration is a bit more dynamic than depicted in Figure 7.5. The hydrostatic and osmotic pressures change as blood flows through the capillary, as depicted in Figure 7.6. At the arterial end of the capillary (left side of the Figure) $P_c - \pi_c - P_{if} + \pi_{if} > 0$ which drives solute out of the capillary into the interstitial space. Since the larger items in the blood such as cells and proteins do not move into the interstitial space the π_c of the capillary blood increases as the fluid that leaves has a lower osmolarity. Also, as the capillary blood loses volume the P_c drops. The movement of fluid from the blood into the interstitial space lowers the π_{if} and raises the P_{if} .

The overall effect of these changes means that on the venous end of the capillary $P_c - \pi_c - P_{if} + \pi_{if} < 0$ which pulls fluid back into the capillary. The filtration on the arterial end (fluid pushed out of the capillary) is greater than the filtration on the venous end (fluid pulled into the capillary) so that there is a net positive filtration. This means some of the capillary fluid stays in the interstitial fluid and is picked up the lymphatic system. This more detailed picture of what is happening accounts for the dynamic filtration process going on in capillaries throughout the body. When the fluid on the arterial end of the capillary is pushed out it mixes with interstitial fluid. Then approximately 80% of fluid is pulled back into the capillary on the venous end. The end result is substantial overall mixing of vascular and interstitial ECF ensuring continuity of ECF throughout the body for homeostasis of ions, delivering nutrients, and removing waste. Additionally, since the remaining 20% enters the lymphatic system there is a screening process on the plasma passing through the lymphatics which has significant immunological benefits.

7.2.5 Factors that Prevent Extra Cellular Edema

There are a number of situations that result in changes to the overall filtration balance and result in interstitial edema. Some of these situations are minor and self limiting. Mild edema in the lower extremities after a day of being upright

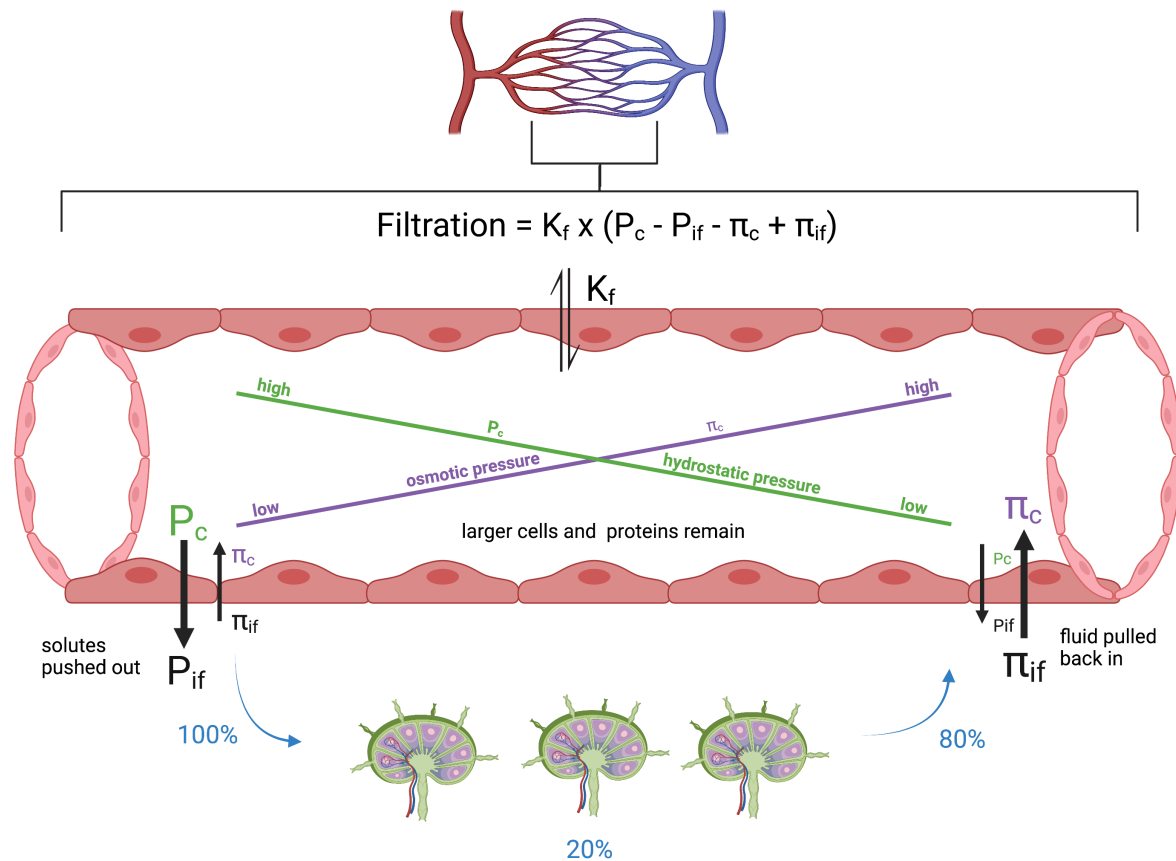


Fig. 7.6: Regulation of Micro-Circulation. At the arterial end of capillary net hydrostatic pressure pushes solute out of the capillary. However this pressure drops as solute leaves the capillary blood flows across the length of the capillary. With the movement of solute out and larger cells and proteins remaining in the capillaries, the net osmotic pressure increases and pulls most of the fluid back into the capillary. Overall, filtration results in a net accumulation in the interstitial space that is then picked up by the lymphatic system. (Created with Biorender.com by Ashley Ney, SPT, in collaboration with Brianna Rowe, SPT; classmates in the PSU DPT Class of 2025)

(mostly standing) is due to elevated venous pressures increasing P_c . There are three safety factors that prevent such situations from becoming severe edema (See Figure 7.7).

1. Low tissue compliance limits the hydrostatic pressure gradient as more fluid accumulates in the interstitial space (low compliance means higher pressure with a small change in volume, an increased P_{if}).
2. As edema develops there is a decrease in π_{if} .
3. Lymphatic flow can increase by as much as 50-fold. These three factors also interact. As lymphatic flow increases the decrease in interstitial fluid protein further decreases the π_{if} concentration in the interstitial fluid reduces the osmotic pressure.

7.2.6 Extra Cellular Edema

There are two general causes of extracellular edema. First, a higher overall filtration pressure (> 0) with a greater net flow of solute from the capillary to the interstitial fluid that is not returned to the capillary. This can be caused by anything that increases capillary permeability (K_f), alters hydrostatic pressure, or alters osmotic pressure. In these situations the increase in filtration exceeds the capacity of a normal lymphatic system. The problem is caused not by faulty lymphatics,

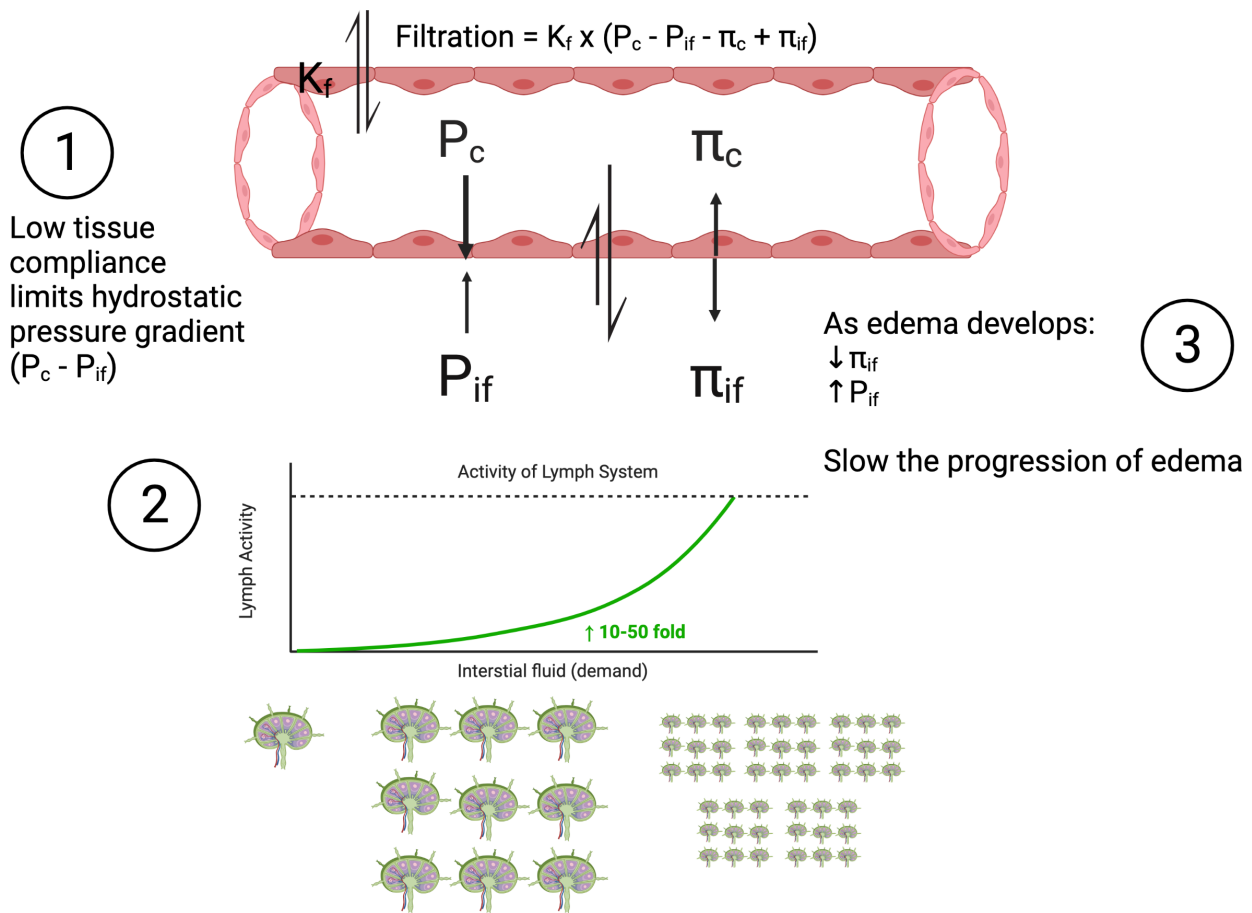


Fig. 7.7: Factors that Prevent Edema (Created with BioRender.com)

but by an increase in filtration. The second general cause of extra cellular edema is failure of the lymphatics to pick up extra fluid from the interstitial fluid. In this situation filtration is normal, but the lymphatics are not able to pick up the 20% of fluid that normal filtration does not return to the capillary. There is a faulty lymphatic system, the edema caused by this situation is called lymphedema.

Three particular situations are presented in Figure 7.8 that can cause interstitial edema. Please note that these are not mutually exclusive.

1. Inflammatory mediators (local damage) or a cytokine storm (systemic inflammation) results in increased K_f which results in a much higher than normal filtration. Since a high K_f can also include movement of larger molecules from the capillaries into the interstitial fluid such as proteins and white blood cells they can have a secondary effect of increasing the π_{if} which contributes to the edema.
2. There are two examples. To the left, heart failure increases venous pressure due to impaired flow through the heart. The increase in venous pressure increases P_c which pushes more solute out of the capillaries. If this delay in flow is on the left side of heart the edema occurs in the lungs. If it is on the right side of the heart the edema occurs in the extremities (primarily the lower extremities since it is made worse by gravity). It is important to note that the most common cause of a heart pump problem on the right side of the heart is a problem with the left side heart pump. The second example for situation 2 is a reduction in albumin production in liver disease (such as cirrhosis). The reduction in the blood protein albumin reduces π_c which lowers the pull of fluid back into the capillaries.
3. There is failure of the lymphatic flow combined with a lower P_{if} due to an increase in the compliance of the interstitial space. Without lymphatics removing solute from the interstitial fluid there is also an increase in π_{if} despite an increase in fluid.

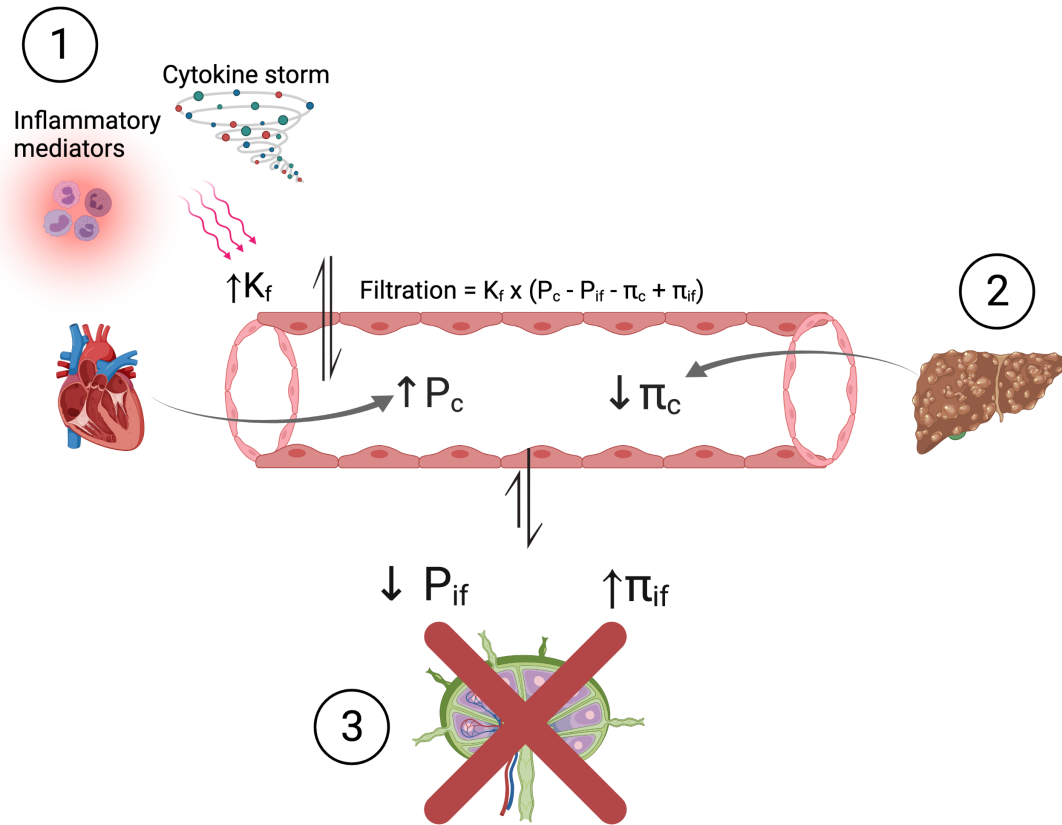


Fig. 7.8: Extra Cellular Edema (Created with BioRender.com)

As with any set of possible comorbidities that are not mutually exclusive (which is most of them), the worst case scenario includes combinations of situations. For example, someone with heart failure and liver disease getting COVID and having a systemic cytokine storm and local inflammatory mediators associated with local hypoxia.

7.3 Muscle Connections: Smooth Muscle

Smooth muscle is present in the walls of organs (i.e. stomach, intestines), passageways (arteries, veins, bronchioles). They are spindle shaped and much shorter than skeletal muscle fibers. Smooth muscles do not have striations but they do have actin and myosin. Unlike skeletal muscles the actin and myosin is not precisely arranged to form sarcomeres. Fiber arrangement is highly variable between smooth muscles and dependent on the role.

7.3.1 Smooth Muscle Activation

Unlike skeletal muscle the tension of smooth muscle fibers acts on other smooth muscle fibers for the purpose of making a compartment or passageway narrow (such as vasoconstriction or vasodilation). Actin is anchored by dense bodies (similar to the Z-discs) which are fastened to the sarcolemma. A smooth muscle contraction is activated by calcium which is supplied by the SR and directly from the extracellular fluid moving through channels in the smooth muscle sarcolemma. Calcium binds to calmodulin (smooth muscle version of troponin-tropomyosin). When activated the sliding

filament theory is similar to skeletal muscle, and the overall effect is tension that attempts, and in most cases does, shorten the fiber. Shortening occurs in most cases because the tension developed does not usually have high resistance preventing the shortening, or at least there are far less variations in the resistances that smooth muscles must compete with (other than perhaps the uterus during labor and delivery where contractions are pushing against a substantial resistance). Since smooth muscle is attached to dense bodies that are attached to the sarcolemma, in addition to shortening smooth muscle tends to also pull inward from all around itself.

7.3.2 Smooth Muscle Excitation

Smooth muscle excitation is initiated by the influx of Ca^{2+} which depolarizes the membrane, activates crossbridges and excites SR to release additional Ca^{2+} . The wave of excitation over the sarcolemma from fiber to fiber because unlike skeletal muscle when one smooth muscle fiber is activated all should be activated (single-unit organization).

7.3.3 Smooth Muscle Regulation

Smooth muscles can be organized as a single-unit (more common) or as as a multi-unit. Single unit smooth muscles have gap junctions that allow quick sharing of Ca^{2+} between cells during excitation so that all the fibers are activated together as a single-unit. Single-unit smooth muscle surrounds the visceral organs and the small blood vessels. Single-unit smooth muscle is regulated by both the autonomic nervous system and stretch. The autonomic nervous system (sympathetic nerves and hormones, parasympathetic nerves) can excite single unit smooth muscles for more or less tension using frequency summation (but not motor unit summation because there are no motor units). Autonomic nerve excitation of single-unit smooth muscle does not occur at precisely located motor end plates and neuromuscular junctions. Instead the autonomic nerves have locations along the nerve fibers called varicosities that are filled with vesicles that release a neurotransmitter that is then free to bind to any acceptable and available smooth muscle receptor (See Figure 7.9). Single-unit smooth muscle also can be excited by stretch in what could be considered a local stretch reflex. Unlike a muscle spindle stretch reflex in skeletal muscle that loops through the spinal cord, the smooth muscle stretch reflex occurs because the mechanical stretch of the fiber opens Ca^{2+} channels that both excite and activate the smooth muscle.

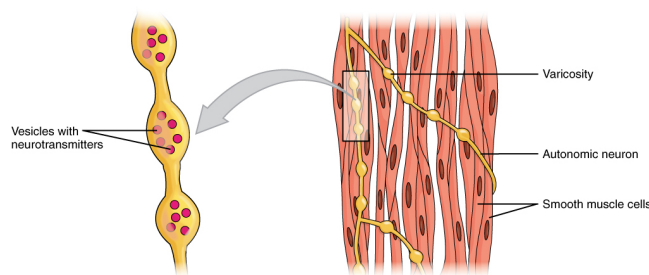


Fig. 7.9: Smooth Muscle (CCBY4.0 from Version 8.25 from the Textbook OpenStax Anatomy and Physiology)

Multi-unit smooth muscle fibers can be regulated by both frequency and fiber unit based excitation. To allow this increase in precision of tension for multi-unit smooth muscles they have precision between the autonomic nerves and the smooth muscle fibers. They also do not have gap junctions so excitation does not spread from one fiber to the next. Like skeletal muscle, excitation is confined to the fiber that was originally excited. Excitation for multi-unit smooth muscles also does not originate from stretching. The large blood vessels and the respiratory airways have multi-unit smooth muscle.

7.3.4 Smooth Muscle Tone

Activation continues until ATP-dependent calcium pumps actively transport Ca^{2+} ions back into the SR and out of the cell. Many smooth muscles maintain a low concentration of Ca^{2+} to maintain tone. In the case of blood vessels this is important because if all blood vessels were fully dilated (no smooth muscle tone) then blood pressure would drop so low that maintaining upright posture would not be possible. Abnormalities in the smooth muscle tone of blood vessels one possible cause for conditions such as orthostatic hypotension, postural orthostatic tachycardia syndrome (POTS) and vasovagal syncope.

To allow smooth muscles to maintain tone for prolonged periods without rest they can maintain contractions even as Ca^{2+} is removed and myosin is inactivated. This can happen due to a subset of crossbridges form latch-bridges. Latch bridges keep actin and myosin connected without ATP, allowing tension (as tone) in smooth muscle that lines arterioles and other visceral organs with very little energy expenditure.

7.4 Clinical Connections: Compression

Compression is an important part of edema management (RICE stands for rest, ice, compression and elevation). Elevation reduces hydrostatic pressure. Compression increases the gradient between venous pressure being compressed and venous pressure further away from compression (usually more proximal). This encourages reduction in P_c which allows more fluid to reenter the capillary during filtration. The pressure also increases the pressure gradient between the local lymphatics and the lymph vessels not being compressed which facilitates movement of edema also into the lymphatics. Compression has also now become more common for recovery to increase micro circulation to remove post exertion waste and delivery needed nutrients. It works through the same mechanisms as above, but in response to small amounts of local edema from a combination of increased cellular and interstitial osmolarity, and increased cellular and capillary permeability (K_f). The compression does not alter the permeability, or the osmolarity changes, but it does reduce P_c which shifts the balance of the filtration equation to encourage movement of edema from the area and into both the capillary and lymphatics.

Compression is also gaining popularity during endurance events, particularly running. However whether there is a benefit and what the mechanisms are remains elusive [Mot+20]. There are several factors to consider with compression during endurance events and it goes beyond what can be discussed in this already long chapter. One reason proposed for compression during an endurance event is to maintain vascular volume. With such long events may be difficult to maintain complete hydration. However, a small amount of dehydration increases π_c which facilitates return of fluid into the circulation. There is also a strong muscle pump working to maintain venous blood flow, once again, facilitating return of fluid to the capillary (keeps P_c low due to venous blood flow and would also promote lymphatic blood flow). So it is unclear whether sustaining vascular volume makes mechanistic sense. It's possible that it facilitates greater overall micro circulation and filtration which encourages greater nutrient delivery and waste product removal from the lower extremities. It is also quite possible that whether compression with endurance running provides a benefit is highly contextual (i.e. related to many interacting causal factors, complex) and is difficult to completely account for in research studies. In such situations it may be best trialed by individuals with self experimentation, which requires a slight reconsideration of what it means to gather evidence [ACR20].

7.5 Summary

The ECF supports muscle fibers by having and providing resources for muscle function. ECF requires micro-circulation. Micro-circulation between cells and the ECF requires balanced osmolarity and a functioning semi-permeable sarcolemma (cell membrane). Micro-circulation between the compartments of ECF, vascular (capillary) and interstitial spaces (fluids) requires filtration. Filtration requires a small unbalance of hydrostatic and osmotic pressures between the capillaries and the interstitial spaces as well as a semi-permeable capillary membrane. Water, ions, molecules and nutrients are constantly circulating between the vascular (capillaries and lymphatics) and interstitial compartments of the ECF. The amount of blood flowing through capillaries can be varied by local and neuroendocrine factors by altering the degree of vasodilation and vasoconstriction of arterioles, pre-capillary sphincters and venuoles. Vasoconstriction occurs

through tone (sustained contraction) of smooth muscles. All causes of intra cellular and extra cellular can ultimately be broken down into altered membrane permeability (cell or capillary), imbalances of osmolarity (and osmotic pressure), imbalances of hydrostatic pressure, and lymphatic function. Compression is an effective intervention for edema due to its facilitation of venous blood flow which reduces P_c .

7.5.1 Next Steps

The next chapter is on Renal Clearance. As you now understand the delicate balance of intra and extracellular fluid is dependent on what is in the extracellular fluid. The extracellular fluid is the entry and exit point of fluid and stuff (nutrients, protein, electrolytes) which then bath and interact with the intracellular fluid. ECF contents is largely regulated by renal function and clearance, creating a stable environment for all the cells of the body.

References

- [ACR20] Rani Lill Anjum, Samantha Copeland, and Elena Rocca. *Rethinking causality, complexity and evidence for the unique patient: a CauseHealth Resource for healthcare professionals and the clinical encounter*. Springer Nature, 2020.
- [Feh17] Joseph J Feher. *Quantitative human physiology: an introduction*. Academic press, 2017.
- [JRD06] David Jones, Joan Round, and Arnold De Haan. *Skeletal muscle: from molecules to movement*. Vol. 40. 2006.
- [Mot+20] Gustavo R Mota et al. *Effects of Wearing Compression Stockings on Exercise Performance and Associated Indicators: A Systematic Review*. en. Open Access Journal of Sports Medicine Volume 11 (Jan. 2020), pp. 29–42. ISSN: 1179-1543. DOI: [10.2147/OAJSM.S198809](https://doi.org/10.2147/OAJSM.S198809). URL: <https://www.dovepress.com/effects-of-wearing-compression-stockings-on-exercise-performance-and-a-peer-reviewed-article-OAJSM> (visited on 06/23/2022).

Sample Questions

1. Why does the osmolarity of intracellular and interstitial fluid stay equalized?
2. A large volume intravenous (IV) infusion of a mildly hypotonic solution will result in?
3. What are safety factors that help prevent severe interstitial edema?
4. Filtration at the capillary includes an exchange of fluid between the vascular and interstitial compartments of the extra-cellular fluid. This exchange includes what processes?
5. What features allow less energy utilization in smooth muscle to maintain tone for vaso-constriction?

Chapter 8

Renal Clearance

Updated on May 22, 2024

Contents

8.1	Mass Balance.....	116
8.2	Renal Function Overview	119
8.3	Filtration-Reabsorption-Secretion	123
8.4	Renal Regulation	126
8.5	Disturbances in Fluid Volume & Osmolarity	129
8.6	Muscle Connections	138
8.7	Summary	138

This chapter covers the concept of mass balance with a particular focus on water and electrolytes. The mass balance of water and electrolytes are supported by renal clearance which includes filtration, reabsorption and secretion for the formation and excretion of urine. Water and electrolyte mass balance are important processes in the support of plasma (and therefore blood) volume; as well as the content and concentration of electrolytes in the extra-cellular fluid. Alterations in the plasma volume and concentration influence the interstitial fluid and the ICF. The regulation of body fluid volume and composition is accomplished by acting directly on the plasma. The kidneys, and systems such as the gastrointestinal, metabolic, and pulmonary, participate in mass balance and the regulation of the body fluid volume and composition.

Objectives include:

1. Explain mass balance.
2. Relate mass balance to the ICF model and physical therapy practice.
3. Summarize the role of the kidneys.
4. Explain the role of the kidneys in mass balance.
5. Explain the functional anatomy of the kidney and nephron and the unique arteriole system including possible symptoms and signs of kidney conditions that should be part of the differential diagnosis of musculoskeletal back pain.
6. Explain renal clearance.
7. Explain urine formation and excretion by micturition and common causes of incontinence
8. Explain glomerular filtration, factors that influence GFR and the use of BUN and creatinine to assess GFR.
9. Explain the role of tubular reabsorption and secretion and how they influence the excretion rate.
10. Describe renal regulation of sodium, potassium, calcium, acid base balance.
11. Explain renal regulation, and integration of blood volume and osmolarity including the role of renin, angiotensin II, aldosterone and ADH.
12. Explain the response to positive and negative water balance.
13. Explain disturbances of fluid volume & osmolarity.
14. Explain rhabdomyolysis and the possible role of water homeostasis on frailty in aging.

8.1 Mass Balance

Mass balance applies the conservation of mass and energy to the analysis of physical systems. It considers the mass and energy entering (inputs) and leaving a system (outputs). Physical therapists help individuals maximize activity and participation such as activities of daily living (ADLs). ADLs are largely based on the need for water and nutrient input, and hygienic output of waste. To use the terms from the ICF model introduced in Chapter 1, the inputs of food, drink and air require activity and participation (function). Many activities of daily living (ADLs) are based on the functional activities that provide appropriate inputs and hygienic outputs required for physiological mass balance.

The inputs of food, drink and air ultimately provide the water, nutrients (macronutrients and micronutrients) and oxygen that are utilized for physiological processes. An overview of physiological mass balance is provided in Figure 8.1. How these are used, whether they can be recycled, and the ways in which they are lost influences the time scale of the need for inputs. When describing outputs the term incidental loss refers to loss that occurs without purpose. Incidental loss is contrasted with intentional loss. Intentional loss is loss for a purpose, such as the loss of water in urine to regulate water balance. An example of incidental loss is the loss of Ca^{2+} in the gastrointestinal tract due to sloughing of cells from the mucosal wall; and the loss of water with breathing. It is interesting to consider whether losses such as Na^+ in sweat is incidental or intentional. It is clear that water loss in sweat is intentional - for the purpose of thermoregulation. While Na^+ loss is not needed for thermoregulation, its loss may help sustain ECF osmolarity. There are several additional examples that go beyond what we can consider here.

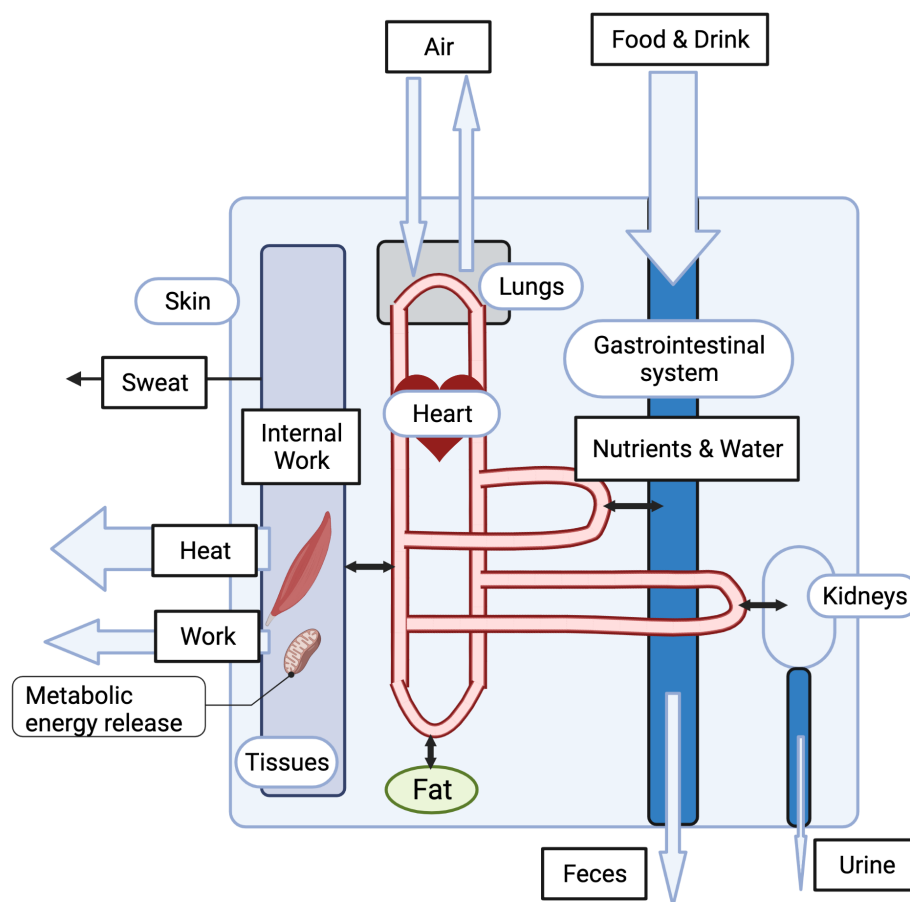


Fig. 8.1: Physiological Mass Balance (Created with BioRender.com)

8.1.1 Oxygen Mass Balance

Oxygen is used regularly and relatively quickly with very little storage (myoglobin) and no recycling, making the input of air with O_2 a regular requirement. Life is not sustainable for prolonged periods without the input of air with O_2 over the time scale of a few minutes in most individuals. At nearly the same rate, and at times at a slightly higher rate than oxygen utilization carbon dioxide is produced as a waste product of mitochondrial respiration. Input and output of air allows for both the input of oxygen, and the output of carbon dioxide.

8.1.2 Nutrient Mass Balance

When discussing mass balance of nutrients, which includes the consideration of diet and nutrition since it is about the input of nutrients, it is common to break nutrients into two main classifications. Macronutrients are larger molecules that can be utilized for energy purposes. They have an energetic (caloric) value in addition to other physiological roles they fulfill. Macronutrients include carbohydrate, fat and protein. Micronutrients are small molecules and in some cases simply elements or minerals. They do not have an energetic value but are essential to many physiological processes. For muscle physiology it has already become clear that Na^+ , K^+ and Ca^{2+} are essential micronutrients.

Semantics of Stuff

The same stuff - such as glucose, calcium, sodium and potassium - can be referred to using different words based on context. Glucose is a macronutrient, but once ingested and absorbed and present in the blood or cells it tends to be referred to as substrate. Micronutrients are dietary inputs. They can be referred to as electrolytes when considering the role they have in determining the osmolarity of a solution (e.g., Na^+). They are considered ions considering the role they have in membrane potentials (e.g., K^+). They are considered signaling molecules (or just molecules) when considering the role they play in activation (e.g., Ca^{2+}).

Macronutrients are routinely utilized for energy and structural health of cells, tissues and molecules such as hormones and neurotransmitters. They are also stored purposely. Fat is stored most abundantly as adipose tissue. Carbohydrate as glycogen. Protein is available from the breakdown of any existing proteins when in need of this substrate. Proteins can be recycled. If a protein is broken down into amino acids and used to build an enzyme that is needed; it can be broken down again and used to build another enzyme. Fat and carbohydrate utilized in structures can be recycled. Macronutrients utilized for energy lose their structure and cannot be recycled, they must be replaced. The digestion of food to extract and absorb usable substrates ultimately yields waste that remains in the gastrointestinal (GI) tract (for example, insoluble fiber) and is removed as feces.

The lining of the GI tract not only absorbs all nutrients (macro and micro), but there is some incidental loss of micro-nutrients (electrolytes and minerals) such as Ca^{2+} in the feces from the GI secretions and normal wear and tear of the cells and tissue that line the GI tract. A small amount of incidental loss of micronutrients such as Na^+ and K^+ occurs in the sweat and urine. Since the input of these micronutrients can easily exceed the incidental loss, there can be intentional (regulated) loss through the renal filtration and secretion processes to maintain Na^+ , K^+ and Ca^{2+} homeostasis.

8.1.3 Water Mass Balance

Water is the source of the fluid environment that the biochemical processes of life occur, and transport of essential nutrients flow through the circulation and micro-circulation. Without water the blood plasma, interstitial fluid and intra-cellular fluid would not be fluid. The fact that they are fluid enables nutrient movement. The amount of water in the body also influences the concentration (and therefore osmolarity) of these fluid spaces. The amount of water influences the blood volume. If blood volume is too high, blood pressure increases and edema may result. If blood volume is too low due to a loss of water, viscosity increases which increases resistance to circulation. Problems occur when there is too little

water (dehydration) or too much water (elevated blood volume, edema). The water in the body is always in use, being filtered between compartments in support of the movement of nutrients and electrolytes as well as by-products (waste). In this way it is recycled - there is no need for a complete overturn of water in the course of a day. However, the amount of water that must be taken in (input) should equal the amount of water that is lost in a day (output). Water loss occurs through four routes, two incidental and two intentional. Incidental water loss occurs with ventilation of air out of the body; and a small amount of water is routinely lost in the feces. Evaporation of water from the skin is a mechanism of cooling the body and sweating is a route of loss. Intentional loss of water in sweat is highly variable based on the temperature and activity levels which influence the need for cooling. Renal filtration and excretion requires a loss of water in urine for the removal of metabolic waste. Of these four routes for the output of water sweating is highly variable and can be partially regulated through behavioral adaptations. Renal is also highly variable and is regulated through physiological processes. The maintenance of normal water volume, and as a result plasma volume, is buffered by fluctuations in renal output in response to water input. Because there is a continual regular loss of water (breathing and sweating), and a necessary regular loss of water for waste removal (feces and urine), there is a regular need (demand) for water input. While the kidneys can make more or less concentrated urine which includes less or more water (respectively), the body must make urine regularly to filter the blood and keep it free from the toxicity of metabolic waste products.

8.1.3.1 Blood Volume Composition

Blood volume includes plasma (approximately 55%) and cells (See Figure 8.2). Most of the plasma is water. Most of the cells are red blood cells (RBCs) and are also called erythrocytes. RBCs contain hemoglobin (Hgb). RBCs are the oxygen carrying cells in blood. Hematocrit (HcT) reports on the percentage of blood volume that contains RBCs. Since other cells occur in relatively low volume, it is acceptable to consider plasma as simply 1 minus HcT when HcT is reported as a fraction (HcT of 45% is 0.45). If HcT is 0.45, then $1 - 0.45 = 0.55$, which means 55% of the blood is plasma. The volume (size) of RBCs influence HcT. If there are the same number of RBCs but each has more volume then the HcT is increased (assuming plasma volume has not changed). If the RBCs have less volume then the HcT is decreased (assuming plasma volume has not changed).

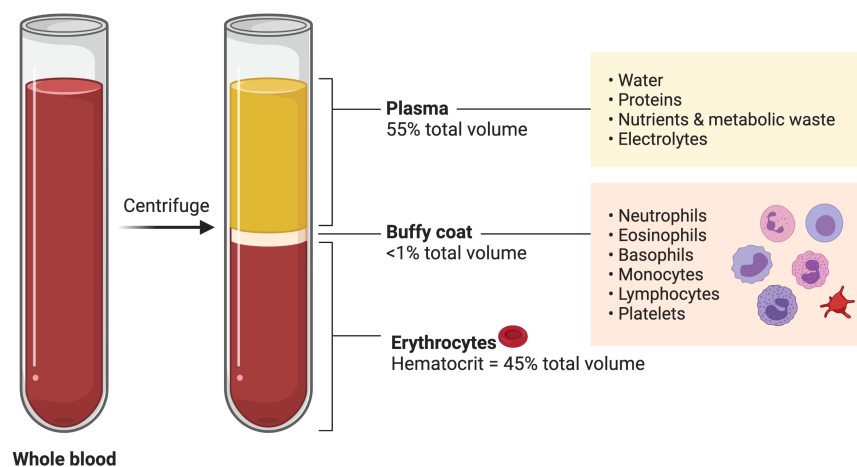


Fig. 8.2: Plasma & Hematocrit of Blood (Created with BioRender.com)

8.2 Renal Function Overview

Renal function relies on the kidneys. The kidneys have several important physiological functions summarized below and in Figure 8.3.

1. **Excretion:** The kidneys ensure that harmful substances are excreted in urine. Harmful substances can be absolutely harmful (waste products such as urea and creatinine), or relatively harmful (electrolytes when they exceed their normal ranges).
2. **Regulation & Buffer by Excretion:** The kidneys regulate and therefore buffer the mass balance of plasma, and therefore body water and electrolyte amounts. Through both water and the electrolyte content they regulate electrolyte concentration.
3. **Regulation by Excretion:** Through the regulation of plasma volume the kidneys provide long term regulation to blood pressure.
4. **Regulation by Hormone Secretion:** The kidneys synthesize and secrete three hormones:
 - a. Renin for the regulation of blood pressure
 - b. Erythropoietin for the regulation of red blood cells
 - c. 1,25-dihydroxycholecalciferol for the regulation of Ca^{2+}

Functions 1, 2 and 3 and the endocrine function of renin are covered in this chapter.

Functions 1 and 2 are accomplished through renal clearance. Renal clearance is a general concept that describes the rate that substances (including water) are removed from plasma. Renal clearance is accomplished as blood passes through the renal capillaries and undergoes filtration. Filtration in the renal capillaries greatly exceeds the filtration in systemic capillaries and therefore is also referred to as ultrafiltration. The fluid that is filtrated out of renal capillaries is referred to as ultrafiltrate. Ultrafiltrate proceeds through the renal tubule while the more selective processes of reabsorption and secretion fine tune its content until it is eventually urine and headed for clearance. In total, renal clearance ensures that harmful substances are excreted (filtered and not reabsorbed); that the correct volume of water and electrolytes remain in the plasma (filtered and then reabsorbed as necessary, or not filtered and secreted as necessary).

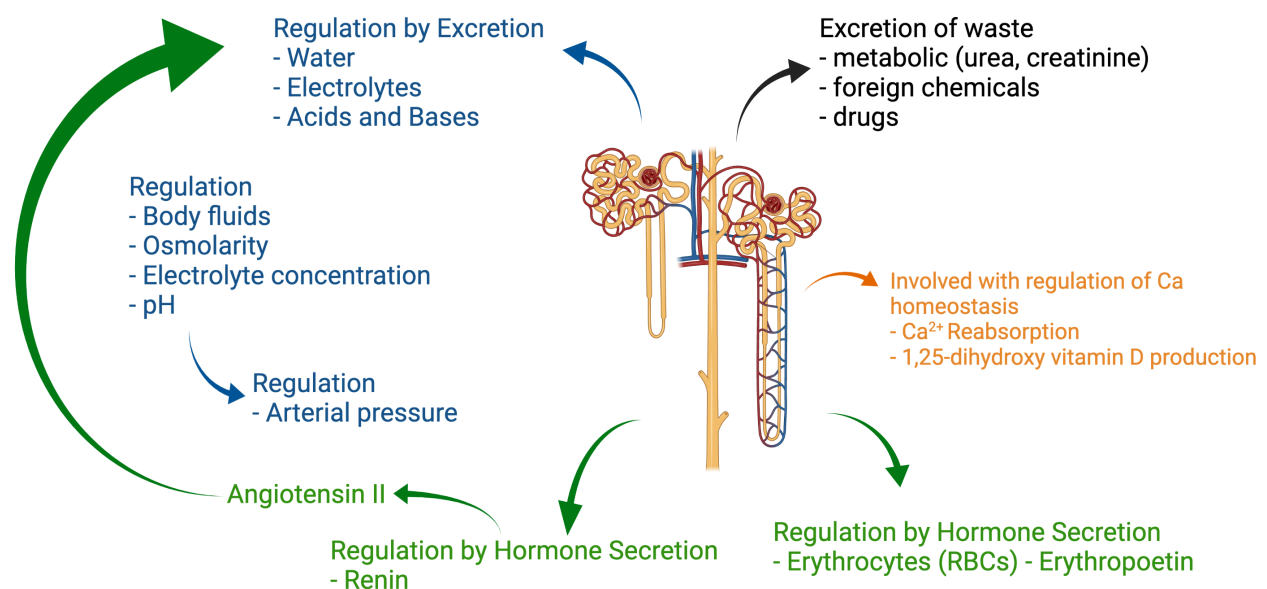


Fig. 8.3: Overview of Renal Functions (Created with BioRender.com)

8.2.1 Functional Anatomy

The kidneys are in the retroperitoneal cavity of the body. Given their location and the high prevalence of back pain in physical therapy practice it is important to consider the possibility that back pain is related to the kidneys.

Kidney Pain

Kidney pain is pain from disease or injury to a kidney. Kidney pain or discomfort may manifest as a dull, one-sided ache in the upper abdomen, side or back. Back pain that includes upper abdomen or side components and that does not respond to provocations that would typically provoke a musculoskeletal source of pain such as movements (active & passive tension) should be considered. The patient should be questioned regarding a history of the common causes of kidney diseases such as alcohol use, diabetes, high blood pressure, family history or a history of kidney stones. Fever and changes in urinary habits or urinary symptoms often accompany kidney pain and are therefore important considerations when ruling out a kidney as a source of back pain.

Each kidney is divided into two regions: the outer cortex and the inner medulla. Comprising these two regions are tube-like structures called nephrons - the functional unit of the kidney. Each kidney can contain a million individual nephrons with each nephron serving a vital role in reabsorption, secretion, and filtration of solutes. At its simplest, each nephron contains a renal corpuscle (glomerulus surrounded by Bowman's Capsule) and renal tubule. The glomerulus is a capillary network and the primary site for plasma filtration and the formation of urine. Distributed on both ends of the glomerulus are the afferent and efferent arterioles, which provide inflow of filtrates to the glomerulus and outflow of filtrates to the upper portion of the nephron - the Bowman's Capsule (or Bowman's Space).

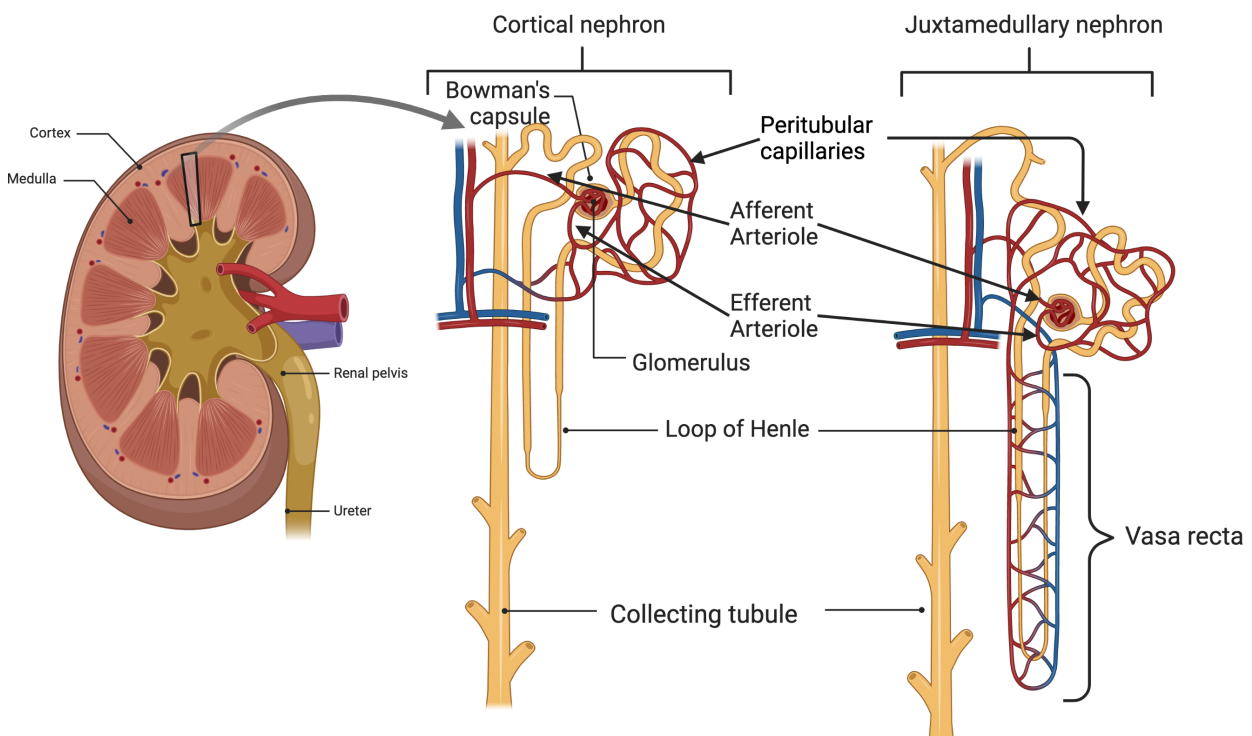


Fig. 8.4: Overview of Kidney Anatomy (Created with BioRender.com)

From this region the ultrafiltrate (which will become urine) travels through a looping tubular structure where the functions of re-absorption and secretion occurs. The segments of the nephron tubule that emerge from Bowman's capsule includes several distinct functional units (in order): the proximal convoluted tubule, the proximal straight tubule, the loop of

Henle (which contains a thin descending limb, a thin ascending limb, and a thick ascending limb), the distal convoluted tubule, and the collecting ducts. The distinct reabsorption and secretion functions across the tubule are based on a highly specialized epithelial cells lining each segment. The collecting ducts drain regionally into pouches which then drain larger pouches eventually reaching the renal pelvis which proceed into the ureter. Through this collecting duct system the urine from each kidney drains into a ureter and is transported to the bladder for storage and eventual elimination.

There are two types of nephrons, superficial cortical nephrons and juxtamedullary nephrons, which are distinguished by the location of their glomeruli and the length of their loops of Henle. The superficial cortical nephrons have their glomeruli in the outer cortex (which is more superficial) and have relatively short loops of Henle. The juxtamedullary nephrons have their glomeruli deeper in the kidney with larger glomeruli and therefore have higher glomerular filtration rates. The juxtamedullary nephrons also have long loops of Henle that go deep into the kidney that are essential for the concentration of urine.

Renal Blood Vessels

A unique aspect to renal blood vessels is that they have two sets of arterioles. Blood enters each kidney through the renal artery, which branches until the smallest arteries subdivide into the first set of arterioles, the afferent arterioles. The afferent arterioles function to deliver blood to the glomerular capillaries - a major site of filtration in the nephron. Following filtration blood travels through the second set of arterioles, the efferent arterioles, which deliver blood to the peritubular capillaries, a major site for water and solute reabsorption. Having arterioles before and after the glomerulus equips each nephron with a unique ability to regulate the hydrostatic pressure of the glomerular capillaries, resulting in changes to both renal blood flow and glomerular filtration rate. For example, constriction at either the afferent or efferent end of the glomerulus results in increased vascular resistance and reduced renal blood flow. To further assist in homeostatic regulation of water-solute balance, the juxtamedullary nephrons have specialized capillary loops called Vasa Recta. These blood vessels travel in a similar form to the Loop of Henle and act as osmotic exchanges, which regulate urine concentration.

8.2.2 Renal Clearance

Renal clearance is the volume of plasma cleared of a substance per unit time and is determined by filtration, reabsorption and secretion.

1. Filtration (indiscriminate elimination): pushing filtrate (plasma) out of the capillary into Bowman's capsule
2. Reabsorption (selective): selectively pulling substances (including water) back into the capillary
3. Secretion (selective): selectively pushing substances out of the capillary into the renal tubule

Substances with the highest clearances are both filtered and secreted and may be completely removed on a single pass of blood through the kidneys. Substances with the lowest clearances either are not filtered or are filtered and subsequently reabsorbed and may not be removed at all. Renal clearance is based on the rate substances are removed (cleared) from plasma in the kidneys. The higher the renal clearance of a substance, the more plasma that is cleared of the substance. For example, kidneys regulate the renal clearance of Na^+ to match what is absorbed so that people do not have to adjust what they ingest. In situations of reduced or impaired renal function dietary modification of Na^+ may be necessary. The general principle is that normal physiology provides individuals lots of freedom since it can make adjustments to buffer the consequences of behavior. As that physiological capacity is lost with aging and various chronic conditions, behavior must be adjusted to avoid the consequences.

Renal Clearance of Substances

Renal clearance of albumin is approximately zero because albumin is not filtered. The renal clearance of glucose is also zero. Glucose is filtered and then completely reabsorbed. Na^+ , urea, phosphate, and Cl^- have clearances that are higher than zero because they are filtered and then partially, and selectively, reabsorbed as required. Inulin, a fructose sugar, is a special case that has made it valuable in the study of kidney function. Inulin is freely and completely filtered across the

glomerular capillaries and it is neither reabsorbed nor secreted. Therefore the amount of inulin in the urine is the amount that was filtered. Its clearance measures the glomerular filtration rate (GFR) and has been utilized as the gold standard for identifying other useful measures of the GFR.

8.2.3 Urine Formation & Excretion by Micturition

Urine is formed by anything (including water) that is filtered and not reabsorbed, or not filtered and then secreted. Urine leaves the kidneys and flows through the ureters into the bladder. Peristaltic contractions of smooth muscles in the ureter, enhanced by parasympathetic activation, force urine from the kidneys toward the bladder which help maintain the flow of urine from the kidneys even as pressure rises in the bladder.

The bladder is a smooth muscle chamber with two parts. The bladder body collects urine and the neck is a funnel-shaped extension that connects with the urethra. The smooth muscle of the bladder can increase bladder pressure to 40-60 *mmHg* and play a major role in emptying the bladder during micturition.

The smooth muscle of the bladder wall is influenced by the autonomic nervous system and a spinal cord stretch reflex. Micturition (urination) is the process of emptying the bladder when it is filled. Filling of the bladder increases its pressure and therefore wall tension rises. Pelvic nerves connect with the spinal cord through the sacral plexus include sensory and motor nerve fibers. The sensory nerve fibers detect when the bladder wall tension rises above a threshold level that activates the micturition reflex to empty the bladder or, if this fails, creates a conscious desire to urinate. The motor nerves transmitted to the pelvic nerves are parasympathetic fibers.

Urinary Semantics

The terms micturition and urination are equivalent and can be used interchangeably. However, the reflex that triggers urination is the micturition reflex. Though, if you heard someone say urination reflex, I'm quite sure you'd know what they meant.

The bladder neck includes the internal sphincter muscle interlaced with a large amount of elastic tissue. The natural tone of the internal sphincter muscle keeps the bladder from emptying until the pressure in the main part of the bladder rises above a critical threshold. Beyond the neck of the bladder the urethra passes through the urogenital diaphragm, which includes a layer of muscle called the external sphincter. This muscle is a voluntary skeletal muscle and is used to prevent urination even when the micturition reflex is attempting to empty the bladder.

8.2.3.1 Urinary Incontinence

There are numerous potential causes of urinary incontinence. A common cause of all forms includes the inability to prevent bladder pressure, either through the micturition reflex or external provocations such as abdominal pressure creating bladder pressure (for example during coughing, laughing, valsalva, etc), from resulting in urination. The micturition reflex includes a complete cycle of (1) progressive and rapid increase in bladder pressure, (2) sustained increase in bladder pressure, and (3) return of the pressure to the basal tone of the bladder. The steps include [HH20]:

1. Sensory signals from the bladder wall stretch receptors are conducted to sacral segments of the spinal cord through the pelvic nerves and then reflexively back to the bladder through the parasympathetic nerves by way of the pelvic nerves.
2. Once the micturition reflex is sufficiently powerful, it causes another reflex that passes through the pudendal nerves to the external sphincter to inhibit it. If this inhibition is more potent than the voluntary constrictor signals to the external sphincter, urination occurs.
3. The micturition reflex is an autonomic spinal cord reflex, but it can be inhibited or facilitated by centers in the brain stem, mainly the pons, and several centers in the cerebral cortex that are mainly excitatory but can become inhibitory.

Not urinating includes the collaborative function of the internal and external sphincter. Both are involved in the micturition reflex above. Voluntary tone of the external sphincter must be properly timed and sufficient to prevent

urination when the micturition reflex is sufficiently powerful. The strength of the micturition reflex is influenced by bladder pressure and wall tension. Therefore bladder pressure can be increased by urine volume or increased abdominal pressure (for example, that comes on during laughing, sneezing or coughing). Much of Women's Health Physical Therapy is centered on helping post-partum women regain voluntary control over the micturition reflex. It is important to point out that the terminology "voluntary" simply means a voluntary skeletal muscle. It does not necessarily mean conscious control. For example, several voluntary skeletal muscles are involved when you breath, maintain a particular posture or walk. While being voluntary they are not necessarily conscious (you can think about your breathing, posture and walking muscle activation, but you typically do not).

8.3 Filtration-Reabsorption-Secretion

8.3.1 Filtration

Glomerular filtration is the first step in the formation of urine. The rate of glomerular filtration is appropriately called the glomerular filtration rate (GFR). As blood enters the glomerular capillaries, a portion of that blood is filtered into Bowman's capsule. The fluid that is filtered is similar to interstitial fluid and at this point is called an ultrafiltrate. Ultrafiltrate contains water and all of the small solutes of blood, but not proteins and blood cells. The pressures responsible for glomerular filtration are similar to the pressures operating for micro-circulation filtration in systemic capillaries (Chapter 7). However, the characteristics and surface area of glomerular capillaries result in a much higher GFR than filtration in systemic capillaries (glomerular capillaries have a higher filtration coefficient than systemic circulation capillaries). The ultrafiltrate of glomerular filtration reenters the peritubular capillaries through reabsorption and can be more selective. The ultrafiltrate that does not re-enter the vascular fluid does not enter lymph vessels, but rather proceeds through the renal tubules for urine formation and excretion.

8.3.1.1 Factors Influencing Glomerular Filtration

As with microcirculation there are four pressures for filtration: two hydrostatic pressures and two osmotic pressures. Applying these pressures to glomerular capillaries, there is one small modification because the osmotic pressure of Bowman's capsule is considered to be zero and therefore is not considered for GFR [Cos13]. The primary factors that influence GFR include:

1. Glomerular filtration coefficient
2. Hydrostatic pressure in glomerular capillary
3. Hydrostatic pressure in Bowman's capsule
4. Osmotic pressure in glomerular capillary
5. Net filtration pressure (nfp)

Glomerular filtration coefficient

The glomerular filtration coefficient is the permeability of the glomerular capillary wall. Two factors that contribute to filtration coefficient are the water permeability per unit of surface area and the total surface area. The coefficient for glomerular capillaries is approximately 100-fold greater than systemic capillaries (such as skeletal muscle capillaries) because of the combination of a higher total surface area and a higher intrinsic water permeability. The consequence of this extremely high coefficient is that much more fluid is filtered from glomerular capillaries than from other capillaries. This coefficient can be decreased with damage to the glomerular capillaries brought on by chronic conditions such as diabetes and high blood pressure (hypertension (HTN)).

Hydrostatic pressure in glomerular capillary (favors filtration)

The hydrostatic pressure in glomerular capillary favors filtration. Compared with systemic capillaries, it is relatively high (60 *mmHg*). In systemic capillaries, hydrostatic pressure falls along the length of the capillary; in glomerular capillaries, it remains constant along the entire length. This is possible because it is regulated with two sets of arterioles, those entering the glomerular capillaries (afferent arterioles) and those exiting the glomerular capillaries (the efferent arterioles). This regulatory system also allows for it to remain relatively constant across a wide range of blood pressures (80-200 *mmHg*). Changes in the hydrostatic pressure in glomerular capillary are produced by changes in the resistance of the afferent and efferent arterioles which subsequently change renal blood flow (RBF). Changes in GFR depend on changes to this hydrostatic pressure which depend on both the afferent and efferent arterioles.

Hydrostatic pressure in Bowman's capsule (opposes filtration)

The hydrostatic pressure in Bowman's capsule is a pressure opposing filtration. The origin of this pressure (18 *mmHg*) is the fluid present in Bowman's capsule and the tubule of the nephron. Changes in this pressure can be produced by obstructing urine flow (e.g., ureter stone or constriction). If the ureter is constricted, urine cannot flow through that ureter to the bladder, causing urine to back up in the kidney. Consequently, hydrostatic pressure in the nephrons will increase as far back as Bowman's capsule. An increase in the hydrostatic pressure in Bowman's capsule decreases the GFR.

Osmotic pressure in glomerular capillary (opposes filtration)

The osmotic pressure in glomerular capillary is a pressure opposing filtration. It is determined primarily by the protein concentration of glomerular capillary blood. It progressively increases along the capillary length as fluid (but not protein) is filtered out of the capillary. The osmotic pressure in glomerular capillary eventually increases to the point where net filtration pressure becomes zero and glomerular filtration stops (filtration equilibrium). Changes in the osmotic pressure in glomerular capillary are produced by changes in plasma protein concentration. Increases in plasma protein concentration produce increases, which decrease the GFR. Decreases in plasma protein concentration produce decreases in osmotic pressure in glomerular capillary which increase GFR.

Net filtration pressure

The net filtration pressure (nfp) is the simple sum of of the previous three pressures with those favoring filtration as positive numbers, and those opposing filtration as negative numbers. The nfp in the glomerular capillaries is positive, and thus favors filtration. The direction of fluid movement is always out of the capillaries. The greater the nfp, the higher the GFR (they are directly related).

8.3.1.2 Clinical Estimation of GFR

The clinical estimation of GFR provides an overall assessment of renal function. Inulin is freely and completely filtered across the glomerular capillaries and it is neither reabsorbed nor secreted. The amount of inulin in the urine is the amount that was filtered and its clearance measures the GFR. Inulin is the gold standard for identifying other useful measures of the GFR. The use of inulin to measure GFR requires a prolonged highly controlled situation that includes set infusion of inulin and collection of urine for hours to record its clearance. Therefore, inulin is rarely utilized clinically to measure GFR.

BUN and creatinine

The closest endogenous¹ substance to inulin for estimation of GFR is creatinine. It is freely filtered across the glomerular capillaries but is also secreted to a small extent. Therefore, clearance of creatinine slightly overestimates the GFR. However, since creatinine is an endogenous substance it does not need to be infused in order to estimate GFR. Urea is another endogenous substance that does not need to be infused that is filtered across the glomerular capillaries. Together, blood urea nitrogen (BUN) and serum creatinine concentration are used to estimate GFR because each substance depends on the filtration step in order to be excreted in urine; urea is reabsorbed in small quantities, and creatinine is secreted in small quantities.

With a decrease in GFR due to renal conditions (such as renal failure), BUN and serum creatinine both increase because they are not adequately filtered. But with lower blood volume (hypovolemia, due to dehydration) or reduced RBF (perfusion, due to heart failure) there is also decreased GFR and both BUN and serum creatinine are increased. However, because urea is reabsorbed and creatinine is not, BUN increases more than serum creatinine. Therefore an indicator of hypovolemia or heart failure caused renal insufficiency is an increased ratio of BUN/creatinine to more than 20. While renal failure due to renal causes produces an increase in both BUN and serum creatinine, it does not produce an increase in the BUN/creatinine ratio [HH20].

	BUN	Creatinine	BUN/Creatinine Ratio
Renal Disease (renal failure)	↑	↑	Normal
↓ Renal Blood Flow (hypovolemia, heart failure)	↑↑	↑	↑

Table 8.1: Summary of BUN/Creatinine Ratio

8.3.2 Tubular Reabsorption

Glomerular filtration results in the production of large quantities of ultrafiltrate each day (approximately 180 L/day). If all of this ultrafiltrate were excreted as urine the following quantities would be lost each day: 180 L of water; 25,200 mmol of Na^+ ; 19,800 mmol of Cl^- ; 4320 mmol of HCO_3^- ; and 14,400 mg of glucose. Each of these losses is 10 times more than the amount present in the entire ECF. This is clearly not a sustainable approach to mass balance.

There are a set of reabsorption mechanisms in the epithelial cells lining the renal tubule that return these (and other) substances to the peritubular capillaries. The details of these mechanisms are beyond the scope of this text. Overall, reabsorption includes the return to capillaries through several different transport proteins that either selectively and directly transport substances, or through the transport of substances which then manipulate the osmolarity between the renal tubule and the peritubular capillaries. Water and many solutes (Na^+ , Cl^- , HCO_3^- , glucose, amino acids, urea, Ca^{2+} , Mg^{2+} , phosphate, lactate, and citrate) are reabsorbed from ultrafiltrate back into the peritubular capillaries. If reabsorption did not occur, most of these constituents of ECF would be rapidly lost in the urine.

This approach, the generation of a large quantity of ultrafiltrate and the selective reabsorption back into the capillaries, is a highly efficient way to filter and clear waste products (urea, creatinine that have limited reabsorption), and regulate the blood and ECF content of other substances. For example, even though glucose is filtered and then reabsorbed, in certain hyperglycemic situations with diabetes, glucose cannot be fully reabsorbed which results in glucose in the urine (glycosuria). Glycosuria contributes to an upper limit of blood glucose and its acutely deleterious effects, but it is not a healthy long term homeostatic mechanism for glucose control.

Reabsorption is adjusted based on ECF fluid volume. Increased ECF volume inhibits tubule reabsorption which results in loss of ECF fluid volume, and decreased ECF volume stimulates tubule reabsorption which results in an increase in ECF. Different diuretics (drugs that promote water removal in urine) work at different points along the renal tubule to reduce reabsorption of water.

¹ Originating or produced within an organism, tissue, or cell.

8.3.3 Tubular Secretion

Secretion mechanisms in the epithelial cells can remove selected and specific substances from the peritubular capillary blood and add them to urine. Organic acids, organic bases, and K^+ are secreted from peritubular capillary blood into tubular fluid as needed. In addition to filtration, secretion provides a mechanism for excreting substances in the urine. Compared to filtration secretion is selective like reabsorption. The secretion mechanisms involve transporters in the membranes of the epithelial cells lining the renal tubule.

8.3.4 Excretion Rate: putting it all together

Excretion rate refers to the amount of a substance excreted per unit time and is related to the clearance rate. However the excretion rate is in terms of the amount of a substance excreted per time frame ($mmol/day$) whereas the clearance rate is in terms of the volume of plasma cleared of the substances ($mmol/L$). The overall reabsorption or secretion rate of a substance is the difference between the amount filtered and its excretion rate. The excretion rate reflects the overall result of glomerular filtration, tubular reabsorption, and tubular secretion. It can be compared with the amount filtered to determine whether a substance has been reabsorbed or secreted. The difference between the amount filtered and the excretion rate is the rate of net reabsorption or net secretion. When the amount filtered is greater than the excretion rate, there has been reabsorption of the substance. If the amount filtered is less than the excretion rate, there has been secretion of the substance.

8.4 Renal Regulation

Renal regulation includes the regulation of renal function, and the renal regulation of several critical characteristics of ECF. These two aspects of regulation are integrated and coordinated. Renal function regulates blood volume and osmolarity by adjusting both water content and electrolytes (primarily Na^+). Regulation of Na^+ content is balanced with regulation of Na^+ concentration. There are limits to what the kidneys do to adjust water volume and osmolarity because they are also making sure the Na^+ concentration stays within fairly tight limits. The regulation of renal function is similarly based on alterations in blood volume that influence glomerular filtration, tubular reabsorption and tubular secretion. Figure 8.5 depicts the bidirectional relationship between renal blood flow (RBF), GFR and renal function (excretion rates and other functions). Renal function influences circulation and arterial pressure which then influences RBF. Arterial pressure is also regulated and is integrated with renal regulation via the hormones renin, angiotensin II, aldosterone and anti-diuretic hormone (ADH).

8.4.1 Sodium

The kidneys maintain a normal body Na^+ content by ensuring that Na^+ excretion equals Na^+ intake, a matching process called Na^+ balance. Na^+ balance is achieved by variations in the reabsorption of Na^+ . If Na^+ excretion is less than Na^+ intake, then there is positive Na^+ balance. Extra Na^+ is retained, primarily in the ECF. When the Na^+ content of ECF is increased, there is increased ECF osmolarity and thus volume; blood volume and arterial pressure also increase, and there may be edema. If Na^+ excretion is greater than Na^+ intake, then a person is in negative Na^+ balance. When excess Na^+ is lost from the body, there is a decreased Na^+ content of ECF, decreased ECF osmolarity and volume, increased ICF volume, and decreased blood volume and arterial pressure.

$$\text{RBF} = (\text{renal artery pressure} - \text{renal venous pressure}) / \text{renal vascular resistance}$$

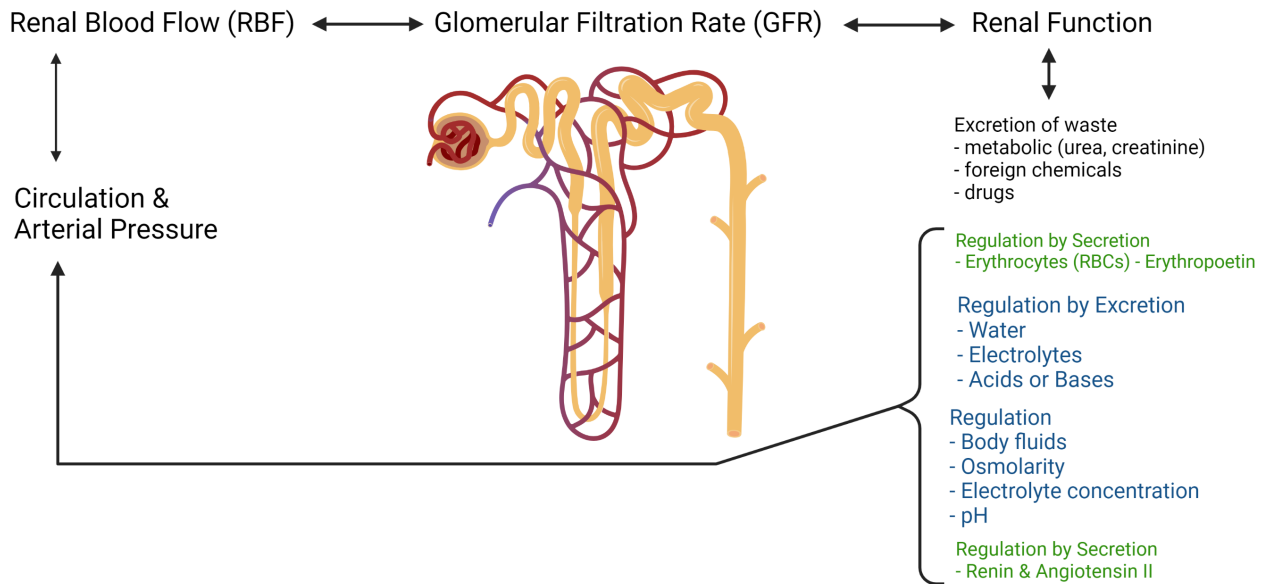


Fig. 8.5: Overview of Renal Regulation (Created with BioRender.com)

8.4.2 Potassium

K^+ balance is maintained by shifts of K^+ across cell membranes and by renal regulation. The renal mechanisms for K^+ balance include filtration, reabsorption, and secretion. Secretion is influenced by dietary K^+ , aldosterone, acid-base balance, and flow rate. With low K^+ intake, more K^+ is reabsorbed and less is secreted; with high K^+ intake less is reabsorbed and more is secreted. The secretion of K^+ in the distal tubule and collecting ducts allows the fine-tuning of K^+ excretion to maintain K^+ balance.

Acid-base disturbances can effect blood K^+ concentration due to alterations in K^+ secretion. Alkalosis increases K^+ secretion, and acidosis decreases K^+ secretion. Therefore, conditions such as chronic pulmonary disease that increase blood carbon dioxide and decrease blood pH (respiratory acidosis) can be accompanied by hyperkalemia.

Commonly used diuretics, the loop diuretics and the thiazide diuretics, cause increased K^+ excretion. Therefore, an important side effect of diuretic therapy is hypokalemia. The K^+ -sparing diuretics (e.g., spironolactone, amiloride, triamterene) do not cause increased K^+ excretion because they inhibit all of the actions of aldosterone and inhibit K^+ secretion. The most common use of K^+ -sparing diuretics is in combination with the loop or thiazide diuretics to offset hypokalemia produced by those drugs.

8.4.3 Calcium

Calcium regulation involves kidney reabsorption in response to parathyroid hormone. The overall coordinated process of calcium regulation involves variations in intestinal absorption and release from bones. Therefore, it is best considered with an understanding of intestinal absorption.

8.4.4 Acid Base Balance

There are three ways the kidneys influence acid base balance. First, the secretion of H^+ ; second the reabsorption of HCO_3^- ; and third, the production of new HCO_3^- . It is best to consider kidney (metabolic) causes and compensations of acid-base balance along with respiratory causes and compensations of acid-base balance. Therefore, clinical interpretation of acid-base imbalance is in the chapter on respiration.

8.4.5 Blood Volume & Osmolarity

Blood volume is regulated based on two primary factors: 1. impact on BP, RBF and subsequently GFR; 2. impact of water in plasma on ECF osmolarity. These are not completely separate regulatory pathways (there is overlap). The overlap is clear when considering the integration of the hormonal mechanisms underlying the regulation of blood pressure and osmolarity. The primary hormones to be considered include: renin, angiotensin II, anti-diuretic hormone (ADH) and aldosterone. The influence of these three hormones on renal function is summarized in Table 8.2. How these hormones and renal function is involved in long term blood pressure regulation is covered in Chapter 9 on Circulation.

Hormone	Activated by	Effects
Renin	↓ GFR	↑ Angiotensin Converting Enzyme (ACE)
Angiotensin II	Angiotensin Converting Enzyme (ACE)	↑ Aldosterone; ↑ $NaCl$ reabsorption, ↑ GFR
Aldosterone	Angiotensin II & ↑ plasma K^+ concentration	↑ K^+ secretion, ↑ $NaCl$ reabsorption
Anti-Diuretic Hormone (ADH)	↑ plasma osmolarity	↑ H_2O reabsorption

Table 8.2: Hormonal Regulation of Renal Function

8.4.5.1 Water Balance

The intentional and incidental loss of water requires regular ingestion and absorption of water. Situations with more water loss than intake is a negative water balance. Situations with more water intake than loss is a positive water balance.

$$WaterBalance = WaterIntake - WaterLoss \quad (8.1)$$

Monitoring water balance occurs through the monitoring of plasma osmolarity by the anterior hypothalamus (osmoreceptors). The response to changes in osmolarity involves ADH which has multiple end organ effects, including those that influence perception (thirst) and behavior (drinking).

Response to Negative Water Balance

1. Water is continuously lost from the body in sweat and in water vapor from the mouth and nose, as well as in urine (the kidneys must continue making urine to for waste excretion). If water is not replaced by drinking water, then plasma osmolarity increases.
2. Increased plasma osmolarity stimulates osmoreceptors in the anterior hypothalamus, which are stimulated by changes in osmolarity of less than 1 mOsm/L. Stimulation of the hypothalamic osmoreceptors has two effects. It stimulates thirst, which drives drinking behavior. It also stimulates secretion of ADH from the posterior pituitary gland.
3. The posterior pituitary gland secretes ADH. ADH circulates in the blood to the kidneys, where it produces increased water reabsorption. As more water is reabsorbed urine osmolarity increases and urine volume decreases (concentrated urine).
4. Increased water reabsorption means that more water is returned to the body fluids. Coupled with increased thirst and drinking behavior, plasma osmolarity is decreased.

The overall response is an excellent example of coordinated homeostatic negative feedback. The original disturbance (increased plasma osmolarity) causes a set of feedback responses (thirst, secretion of ADH and increased water reabsorption) that decrease plasma osmolarity [HH20]. Negative water balance results in concentrated urine.

Response to Positive Water Balance

1. Ingested water is distributed throughout the body fluids. Because the amount of solute in the body is unchanged, the added water dilutes the body fluids and cause a decrease in plasma osmolarity.
2. The decrease in plasma osmolarity inhibits osmoreceptors in the anterior hypothalamus.
3. Inhibition of the osmoreceptors has two effects. It decreases thirst and suppresses water drinking behavior. It also inhibits secretion of ADH from the posterior pituitary gland.
4. When ADH secretion is inhibited, circulating levels of ADH are reduced and less ADH is delivered to the kidneys. Lower ADH levels decreases water reabsorption and water is excreted, decreasing urine osmolarity and increasing urine volume (diluted urine).
5. Because less water is reabsorbed, less water is returned to the circulation. Coupled with the inhibition of thirst and the suppression of water drinking, plasma osmolarity increases back toward the normal value.

The overall response is an excellent example of coordinated homeostatic negative feedback. The original disturbance (decreased plasma osmolarity) causes a set of feedback responses (decreased thirst, less secretion of ADH and decreased water reabsorption) that increase plasma osmolarity [HH20]. Positive water balance results in diluted urine.

Electrolyte Induced Changes to Osmolarity

Water balance is not the only change that alters osmolarity and trigger the osmoreceptor - ADH response. Changes to osmolarity due to changes in solute concentrations also trigger the osmoreceptor - ADH response. The response coordinated by ADH primarily changes water content to adjust osmolarity, but also facilitates correction of solute concentrations, though not solute content. For example, if Na^+ content increases with no change in water then Na^+ concentration increases along with plasma osmolarity. This hypernatremia induced increase in osmolarity causes a shift in water out of the ICF, which partially reduces but does not normalize the ECF osmolarity. The still higher than normal osmolarity activates osmoreceptors that result in the release of ADH. ADH results in thirst, increased intake of water, and reduced loss of water. Together this response dilutes the plasma so that Na^+ concentration is reduced along with osmolarity. A consequence is higher plasma volume and therefore higher blood volume, which may result in higher blood pressure. Ideally the kidneys continue to correct this problem with less Na^+ reabsorption that returns plasma volume to a normal range.

8.5 Disturbances in Fluid Volume & Osmolarity

It is a useful exercise to understand and reason through disturbances in fluid volume (2) and osmolarity (3). Altogether six disturbances. It is not intended that readers would memorize the sequences below, but rather understand them and, if needed, recreate them from thinking through the processes involved from an understanding of the principles. The descriptions below do not include all of the possible trajectories or all of the physiological responses. One goal is to start considering the impact of ECF volume on hematocrit, and the fact that the volume of a RBC is intracellular.

Osmolarity is the concentration of osmotically active particles, expressed as milliosmoles per liter ($mOsm/L$). The normal value for osmolarity of the body fluids is approximately $290 mOsm/L$ in both ECF and ICF since water will shift between these spaces until osmolarity equilibrates [Cos13].

Plasma osmolarity ($mOsm/L$) can be estimated from the plasma Na^+ concentration ($mmol/L$), plasma glucose concentration (mg/dL), and blood urea nitrogen (BUN) (mg/dL), since these are the major solutes of ECF and plasma.

$$PlasmaOsmolarity(mOsm/L) = 2 \times [Na^+] + \frac{[Glucose]}{18} + \frac{[BUN]}{2.8} \quad (8.2)$$

The Na^+ concentration is multiplied by 2 because Na^+ is balanced by an equal concentration of negative ions (in plasma, these anions are Cl^- and HCO_3^-).² The glucose concentration in mg/dL is converted to mOsm/L when it is divided by 18. The BUN in mg/dL is converted to mOsm/L when it is divided by 2.8. Calculating Equation 8.2 with the values of 140 mmol/L of Na^+ , 80 mg/dL of glucose, and 15 mg/dL for BUN results in 289.79 mOsm/L.

In the following sections volume contraction means a decrease in ECF volume. Volume expansion means an increase in ECF volume. In place of the more general terms isotonic, hypertonic and hypotonic the more specific terms isosmotic, hyperosmotic, and hyposmotic refer to the osmolarity of the ECF for the osmolarity disturbances. Consistent with Chapter 7, an isosmotic disturbance means that there is no change in ECF osmolarity; a hyperosmotic disturbance means that there has been an increase in ECF osmolarity; and a hyposmotic disturbance means that there has been a decrease in ECF osmolarity.

Three-Step Approach to Reasoning Through Disturbances in Fluid Volume & Osmolarity

To understand these disturbances, a three-step approach is recommended ([Cos13]):

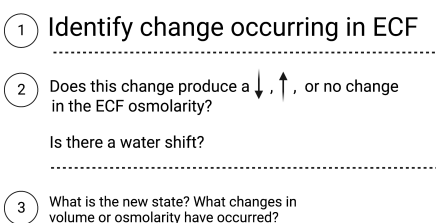


Fig. 8.6: Three Step Approach (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

Six Disturbances: Volume x Osmolarity

The following six pages each includes a description of a volume osmolarity disturbance. Each page also includes a graphic depiction of how to examine each using the three step approach.

Osmolarity	Volume Contraction	Volume Expansion
Isosmotic	Diarrhea	Infusion of NaCl
Hyperosmotic	Sweating Lactate	High NaCl Intake
Hyposmotic	Adrenal Insufficiency	SIADH

Table 8.3: Volume & Osmotic Disturbances

² Despite common assumptions that the outside of the cell is positive and the inside of the cell is negative, there is electroneutrality in both ECF and ICF fluid. Recall from Chapter 4 that membrane potentials are generated from the movement of the ions, not the concentration of the ions.

8.5.1 Isosmotic Volume Contraction: Diarrhea

Diarrhea results in a large volume loss of fluid from the gastrointestinal tract. The osmolarity of the fluid lost is approximately equal to that of the ECF (isosmotic). ECF volume decreases, but there is no accompanying change in ECF osmolarity. Therefore there is no need for a fluid shift across cell membranes and ICF volume remains unchanged. In the new steady state, ECF volume decreases and the osmolarities of ECF and ICF are unchanged. The decrease in ECF volume means that blood volume (a component of ECF) also is reduced, which produces a decrease in arterial pressure. Other consequences of diarrhea include increased hematocrit and increased plasma protein concentration.

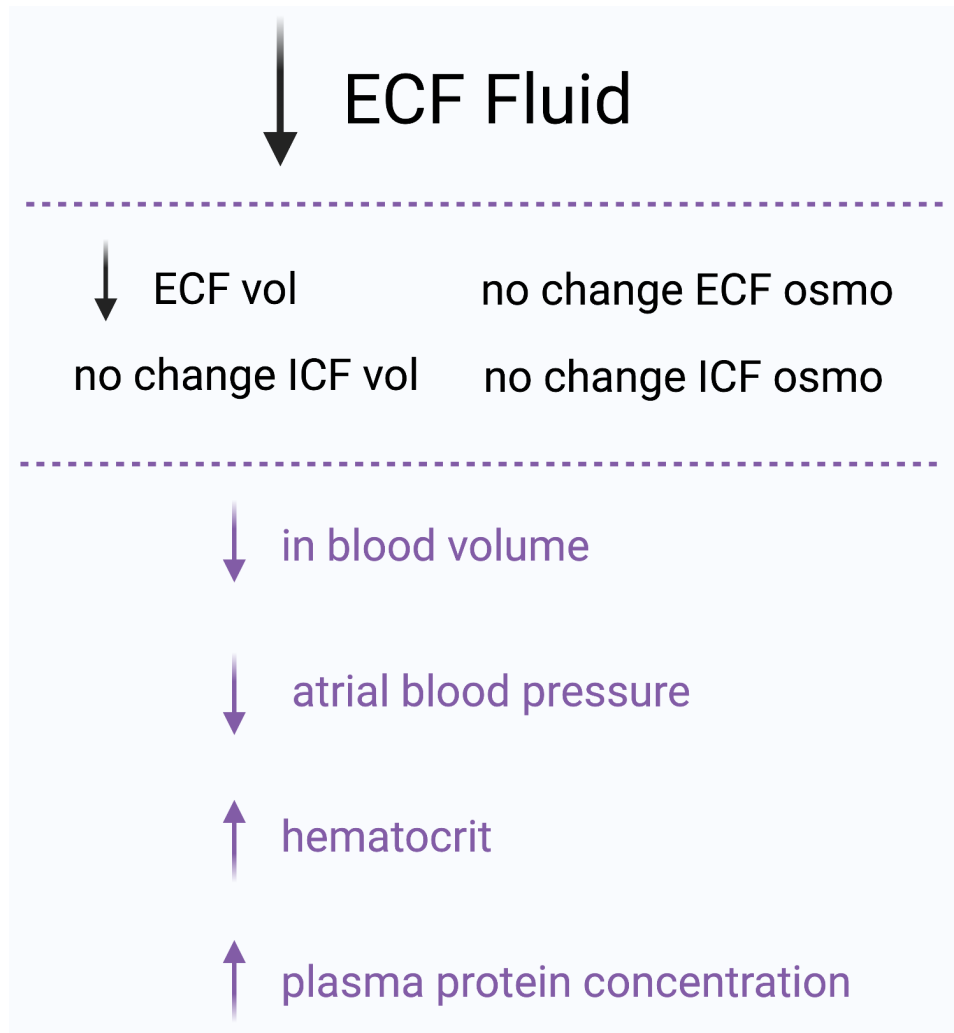


Fig. 8.7: Isosmotic Volume Contraction (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.2 Hyperosmotic Volume Contraction: Sweating

Water deprivation (negative mass balance of water) in situations with sweating (hot environments and/or high metabolism) results in net losses to both NaCl and water in sweat. Sweat is hyposmotic relative to ECF. Sweat contains relatively more water than solute. When sweating hyposmotic fluid is lost from the ECF, ECF volume decreases and ECF osmolarity increases. ECF osmolarity is transiently higher than ICF osmolarity, and this difference in osmolarity causes water to shift from ICF into ECF until ICF and ECF osmolarity equalizes. In the new steady state, both ECF and ICF volumes are decreased and ECF and ICF osmolarities increased. In hyperosmotic volume contraction, the plasma protein concentration is increased but the HcT is unchanged. The explanation for the increase in plasma protein concentration is straightforward: Fluid is lost from ECF, and the plasma protein remaining behind becomes concentrated. It is less obvious, however, why the hematocrit is unchanged. Loss of fluid from ECF alone would cause an increase in the concentration of red blood cells and an increase in HcT. However, there also is a fluid shift in this disturbance: Water moves from ICF to ECF. Because RBCs are cells, water shifts out of them, decreasing their volume. Thus, the concentration of red blood cells increases, but red blood cell volume decreases. The two effects offset each other, and HcT is unchanged. The shift of water from ICF to ECF offsets the loss of ECF from water loss.

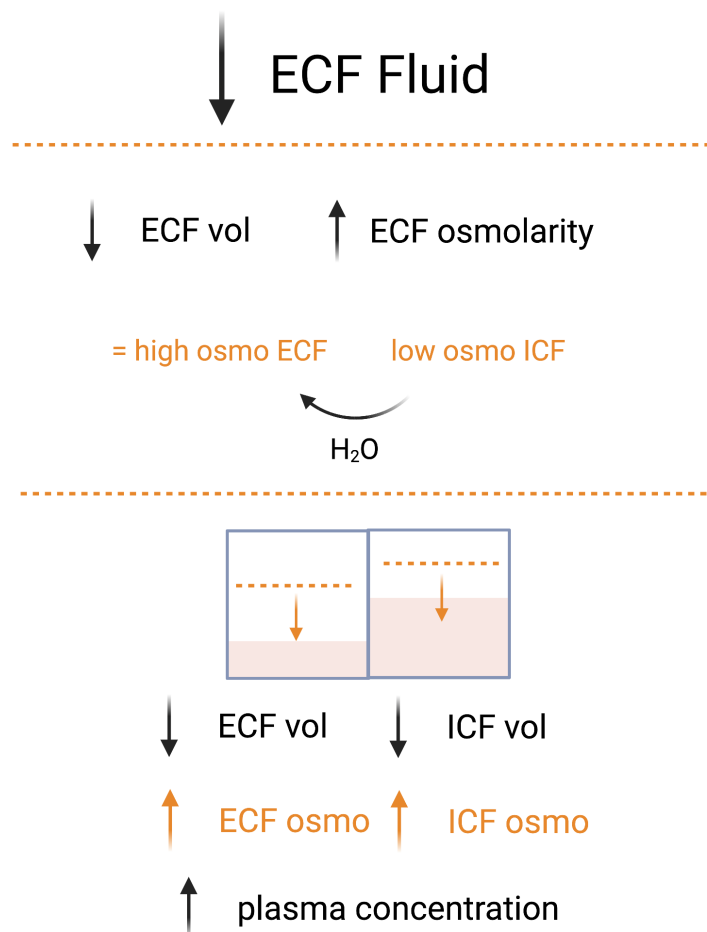


Fig. 8.8: Hyperosmotic Volume Contraction (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.3 Hyposmotic Volume Contraction: Adrenal Insufficiency

Adrenal insufficiency includes a deficiency in aldosterone, a hormone that promotes Na^+ reabsorption. Aldosterone deficiency results in excess NaCl excreted in the urine. Because NaCl is an ECF solute, ECF osmolarity decreases. ECF osmolarity is less than ICF osmolarity and causes water to shift from ECF to ICF until osmolarity equilibrium. In the new steady state, both ECF and ICF osmolarities are lower than normal. The shift of water results in decreased ECF volume and increased ICF volume. In hypotonic volume contraction, both plasma protein concentration and Hct will be increased because of the decrease in ECF volume. Hct also increases because of the shift of water into red blood cells, increasing cell volume.

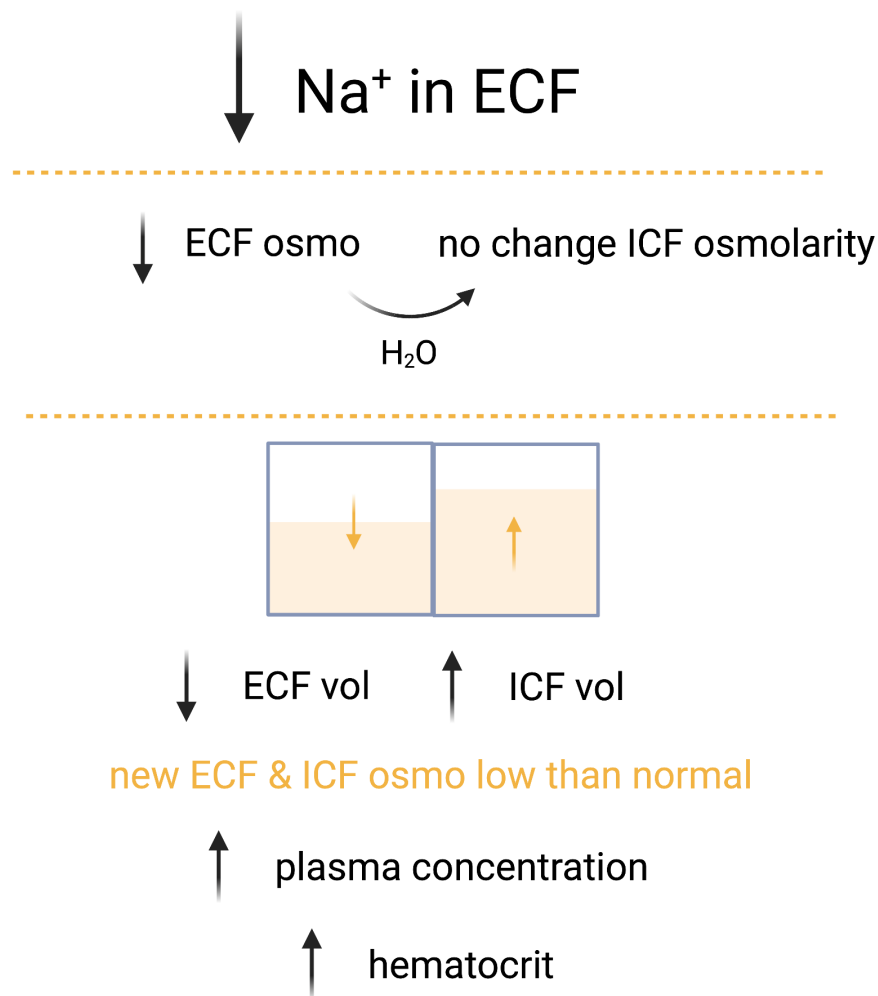


Fig. 8.9: Hypotonic Volume Contraction (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.4 Isosmotic Volume Expansion: Infusion of NaCl

An infusion of isotonic NaCl presents the opposite clinical picture of losing isotonic fluid through diarrhea. Because NaCl is an extracellular solute, all isotonic NaCl solution is added to the ECF, causing an increase in ECF volume but no change in ECF osmolarity. There is no shift of water between ICF and ECF because there is no difference in osmolarity between the two compartments. Both plasma protein concentration and HcT decrease because of the increase in ECF volume.

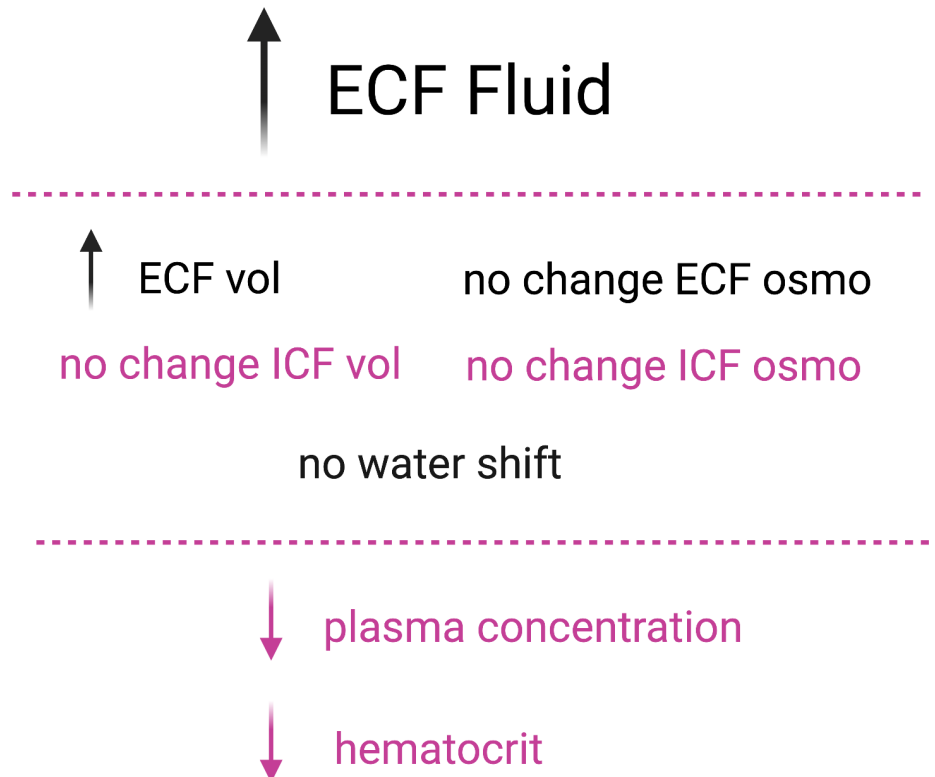


Fig. 8.10: Isosmotic Volume Expansion (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.5 Hyperosmotic Volume Expansion: High NaCl Intake

Ingesting dry NaCl (for example, eating a salty snack) increases the total amount of Na^+ in the ECF and osmolarity increases transiently as the higher ECF osmolarity causes water to shift from ICF to ECF, decreasing ICF volume and increasing ECF volume. In the new steady state, both ECF and ICF osmolarities are higher than normal and equal to each other. Because of the shift of water out of cells, ICF volume will decrease and ECF volume will increase. In hyperosmotic volume expansion, both plasma protein concentration and hematocrit will decrease due to the increase in ECF volume. Hematocrit also will be decreased because of the water shift out of the red blood cells.

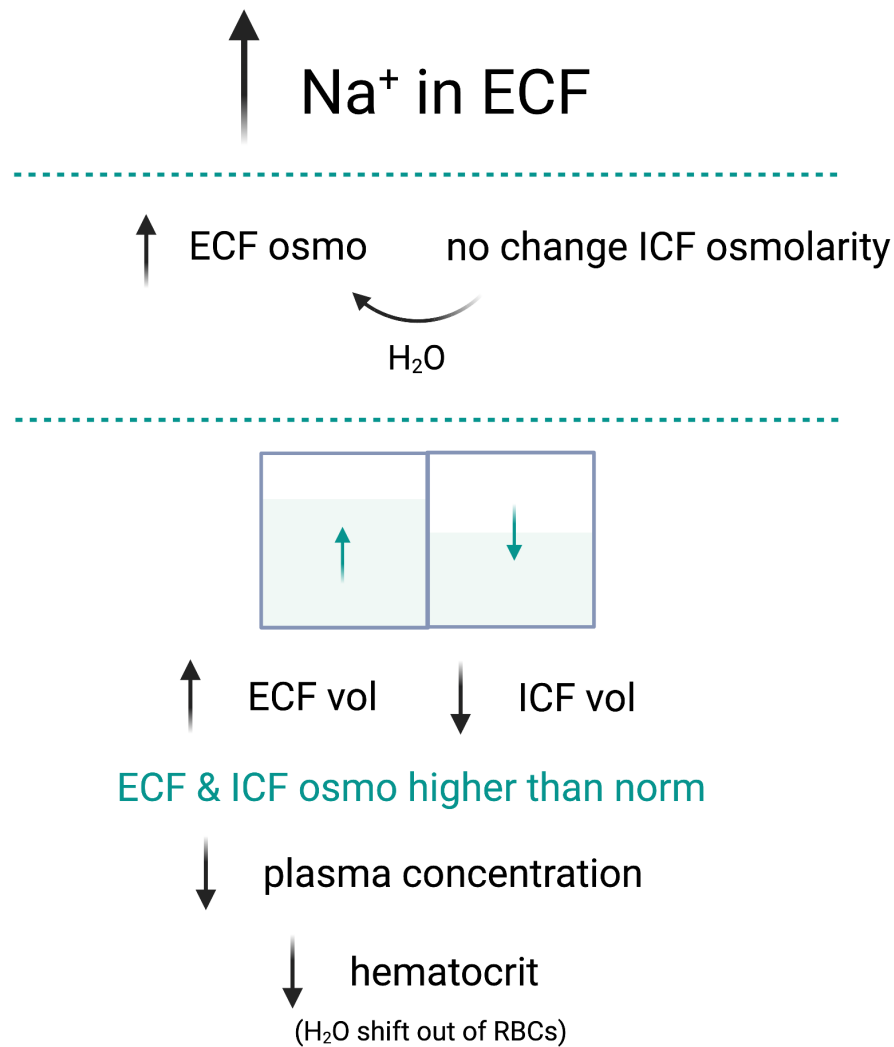


Fig. 8.11: Hyperosmotic Volume Expansion (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.6 Hyposmotic Volume Expansion: SIADH

Syndrome of inappropriate antidiuretic hormone (SIADH) secretes inappropriately high levels of ADH, which promotes too much water reabsorption. The excess water is retained and distributed throughout the total body water. ECF is diluted and transiently has a lower osmolarity. Water shifts into the ICF since it has higher osmolarity. When compared with the normal state, ECF and ICF volumes are both increased and ECF and ICF osmolarities will be decreased. In hypotonic volume expansion, plasma protein concentration is decreased by dilution. The hematocrit is unchanged as a result of two offsetting effects: The concentration of red blood cells decreases because of dilution, but RBC volume increases because water shifts into the cells. But this increase in RBC volume (due to water) does not necessarily increase oxygen carrying capacity of HbG so maintaining HcT in the situation of hypotonic volume expansion does not necessarily mean oxygen carrying capacity is normal.

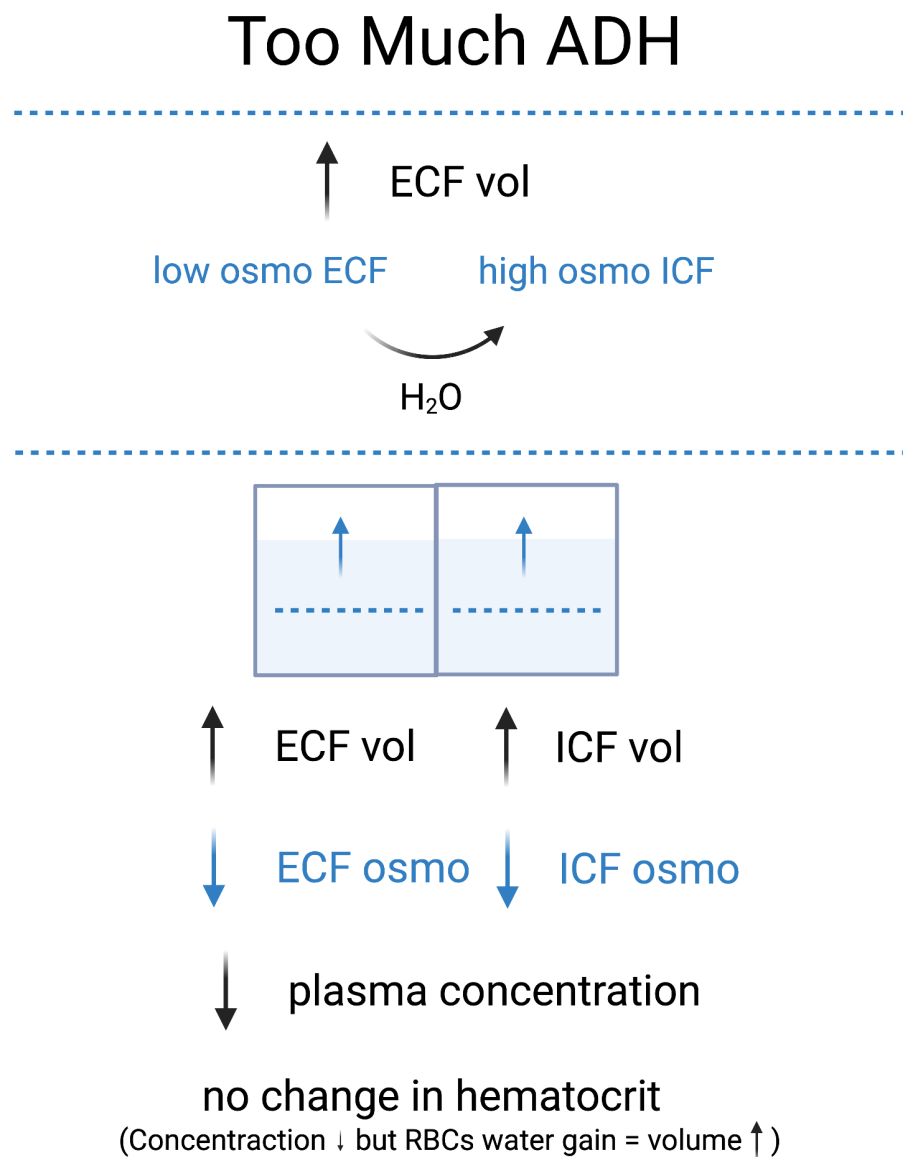
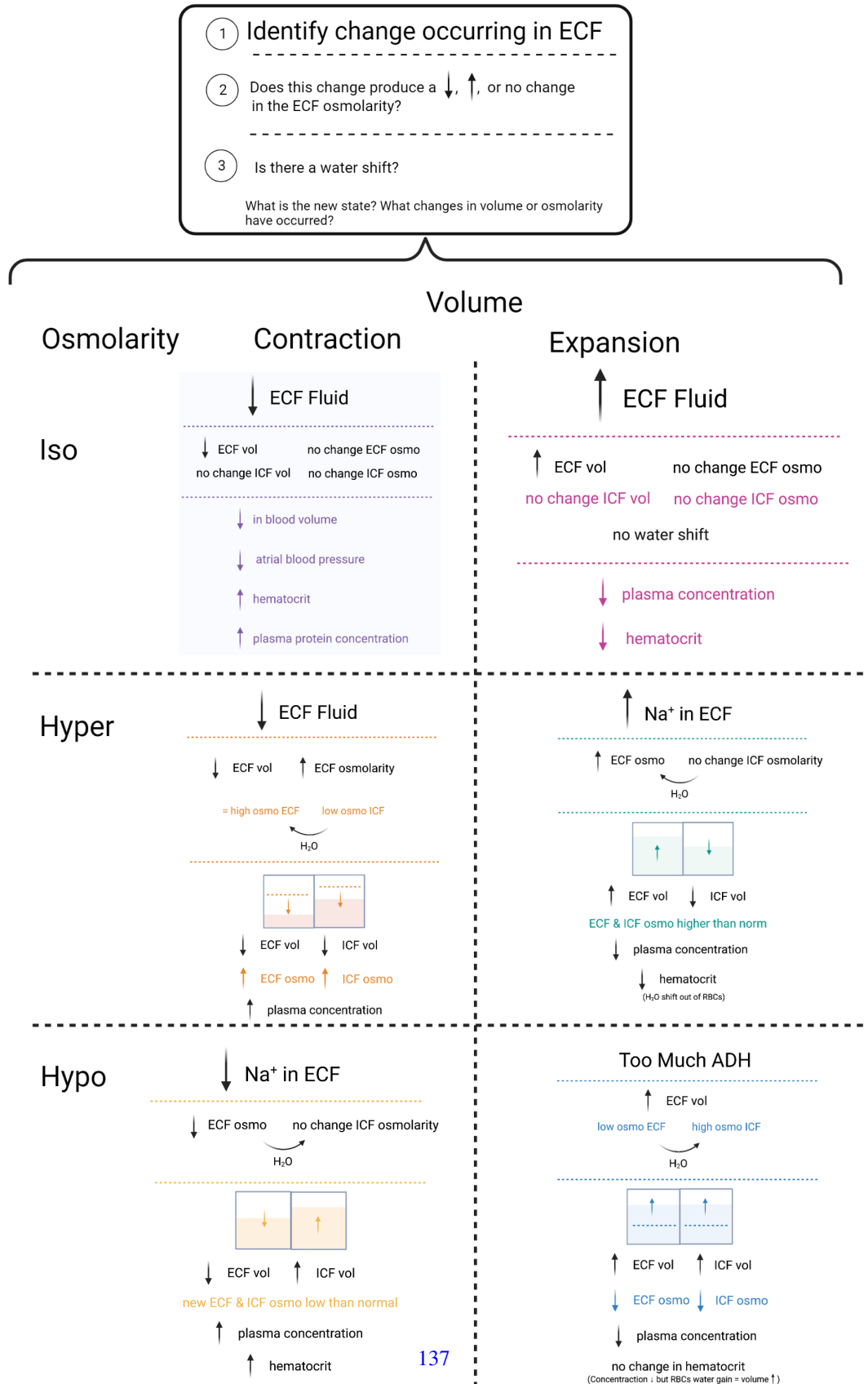


Fig. 8.12: Hypotonic Volume Expansion (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

8.5.7 Summary of Disturbances in Fluid Volume & Osmolarity



8.6 Muscle Connections

8.6.1 Speculation on Water Homeostasis in Muscle Function & Frailty

Older adults have reduced thirst sensation and the ability to concentrate urine, together this results in increased ECF osmolarity (hyperosmotic stress) [LSY19]. Hyperosmotic stress leads to cell dehydration (reduced ICF), which has consequences for the intracellular protein structure and function. In severe situations this can result in cell damage. There is also some evidence to suggest that cell volume may act as a signal, with cell swelling acting as an anabolic signal (promotes hypertrophy) and cell shrinkage acting as a catabolic signal (promotes atrophy). One possible reason, or contributing factor, for the age related progressive loss in muscle mass and strength may be impaired ECF - ICF water homeostasis that originates from mild kidney insufficiency and a loss of thirst. Intracellular water content in lean mass has been related to muscle strength, functional capacity, and frailty risk, and has been proposed as an indicator of muscle quality and cell hydration [LSY19]. Muscle weakness can promote further decreases in fluid (and nutrient) intake due to challenges in ADLs. This can occur directly with difficulty attaining the necessary muscle tension for performing ADLs (either basic ADLs or instrumental ADLs), or with the accumulation of objective fatigue (difficulty sustaining ADLs)[CD15]. Several chronic conditions that include renal insufficiency such as diabetes, high blood pressure, chronic heart failure, obstructive pulmonary disease and renal failure all share a common feature of muscle catabolism, atrophy and weakness. With these conditions the atrophy and weakness seem to go beyond disuse atrophy [AV06]. A common, and plausible, feature of these conditions may be water homeostasis.

8.6.2 Rhabdomyolysis

Rhabdomyolysis is a condition in which damaged skeletal muscle breaks down rapidly. Symptoms include muscle pain; signs include weakness, muscle stiffness and edema. Signs may also include tea-colored urine, an irregular heartbeat, vomiting, and confusion. A positive urine myoglobin test provides supportive evidence. Myoglobin is harmful to nephrons so if rhabdomyolysis is severe it may lead to kidney failure. Common causes of muscle damage include crush injury, strenuous exercise, medications (including statins for lowering blood cholesterol), or substance abuse. Less common causes include infections, electrical injury, heat stroke, prolonged immobilization, lack of blood flow to a limb, or snake bites.

8.7 Summary

Mass balance is an essential concept for physical therapy practice. The need for intake and the intentional or incidental output is central to physiology and wellness. The need for intake frames our activities of daily living. Having the mass balance to sustain an environment for muscle function (ECF) requires muscles to function (performing activities). Renal function, including clearance, filtration, reabsorption, secretion and excretion are essential to mass balance. Kidneys ensure not only that waste products are eliminated but that extra micronutrients (or even nutrients in extreme cases) are eliminated; or that when there are not extra micronutrients that their elimination is reduced. It is much easier to ensure mass "balance" by adjusting what we eliminate than by adjusting what is ingested. The kidneys allow that freedom. In doing this primary role that kidneys influence, and regulate, blood volume and therefore blood pressure. In turn, the kidneys are regulated based on blood volume and blood pressure. There is an exquisite set of integrated renal, blood pressure and hormonal responses that keeps fluid volume, blood volume, blood osmolarity, blood pressure and micronutrient content and concentration balanced. With aging or certain chronic conditions the physiological range of integrated renal regulation may be reduced and behavior must change to allow for easier regulation (i.e., regulating water and Na^+ intake through diet makes it easier for the kidneys to regulate water and Na^+ levels by varying water and Na^+ output).

References

- [AV06] Gregory R. Adams and Nosratola D. Vaziri. *Skeletal muscle dysfunction in chronic renal failure: effects of exercise*. en. American Journal of Physiology-Renal Physiology 290.4 (Apr. 2006), F753–F761. ISSN: 1931-857X, 1522-1466. DOI: [10.1152/ajprenal.00296.2005](https://doi.org/10.1152/ajprenal.00296.2005). URL: <https://www.physiology.org/doi/10.1152/ajprenal.00296.2005> (visited on 07/09/2022).
- [CD15] Sean Collins and Konrad Dias. *Heart Failure: Stability, Decompensation, and Readmission*. en. Cardiopulmonary Physical Therapy Journal 26.3 (Sept. 2015), pp. 58–72. ISSN: 1541-7891. DOI: [10.1097/CPT.000000000000011](https://doi.org/10.1097/CPT.000000000000011). URL: <http://Insights.ovid.com/crossref?an=01823246-201509000-00003> (visited on 01/17/2020).
- [Cos13] Linda S Costanzo. *Physiology*. Elsevier Health Sciences, 2013.
- [HH20] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences, 2020.
- [LSY19] Lorenzo, Serra-Prat, and Yébenes. *The Role of Water Homeostasis in Muscle Function and Frailty: A Review*. en. Nutrients 11.8 (Aug. 2019), p. 1857. ISSN: 2072-6643. DOI: [10.3390/nu11081857](https://doi.org/10.3390/nu11081857). URL: <https://www.mdpi.com/2072-6643/11/8/1857> (visited on 07/09/2022).

Sample Questions

1. What are the functions of the kidneys for maintaining acid base balance?
2. Diuretics work by increasing urine formation in the kidney's. Some diuretics (Aldosterone antagonists and Sodium channel blockers) are called potassium sparing diuretics because they decrease K⁺ secretion. What are possible consequences of diuretics that are NOT potassium sparing?
3. What happens during Hyposmotic Volume Expansion?
4. What happens during negative water balance (Water Balance = Water Intake - Water Loss)?
5. What happens without any tubular reabsorption to the mass balance of essential nutrients such as glucose and electrolytes such as sodium?
6. What are possible causes of a large increase in the BUN but only a small increase in Creatinine so that the BUN/Creatinine ratio increases?

Chapter 9

Circulation

Updated on May 22, 2024

Contents

9.1	Structures of Circulation	142
9.2	Cardiac Output	143
9.3	Regulation of Cardiac Output	145
9.4	Blood Pressure Regulation	147
9.5	Cardiovascular Vital Signs	152
9.6	Summary	154

This chapter covers the circulation support of micro-circulation, which supports the ECF, which supports skeletal muscle function. Circulation includes two vascular circuits that circulates blood through the systemic and pulmonary circulation. The cardiac pump that supports circulation is covered in the next chapter.

The primary goal of this chapter is an understanding of circulation in its supportive role for skeletal muscle. Circulation is critical for sustaining the micro-circulation of the ECF for cellular functions throughout the body. Without circulation cells with a high resting metabolism such as the brain and heart can only survive for minutes. The lack of O_2 delivery, and lack of CO_2 removal being the first two problems that threaten cell life. Problems with circulation can limit the metabolic function of muscles.

The importance of circulation for life and function makes it a highly studied area of physiology. Since it has been highly studied it has several useful models for understanding circulation. After reviewing the important structures for circulation the chapter introduces several of the useful models (in the form of equations that demonstrate relationships).

Objectives include:

1. Describe how the anatomy of the heart and blood vessels supports circulation.
2. Identify cardiac anatomy structures.
3. Explain and Interpret models (equations) of cardiac output, blood pressure and resistance.
4. Apply equations of cardiac output, blood pressure and resistance (including Poiseuille's Law) to explain and evaluate the function and possible dysfunction of circulation.
5. Explain circulation regulation based on blood pressure regulation.
6. Explain circulation distribution and cardiac index.
7. Explain blood pressure regulation.
8. Interpret the events of vasovagal syncope.
9. Perform and interpret cardiovascular tests and measures including: Heart rate and Blood pressure

9.1 Structures of Circulation

The heart and the roots of the great vessels occupy the pericardium, which is located in the mediastinum. The sternum, the costal cartilages, and the medial ends of the third to fifth ribs on the left side of the thorax create the anterior border of the mediastinum. It is bordered inferiorly by the diaphragm, posteriorly by the vertebral column and ribs, and laterally by the pleural cavity (contains lungs). Specific cardiac structures and vessels are depicted in Figure 9.1.

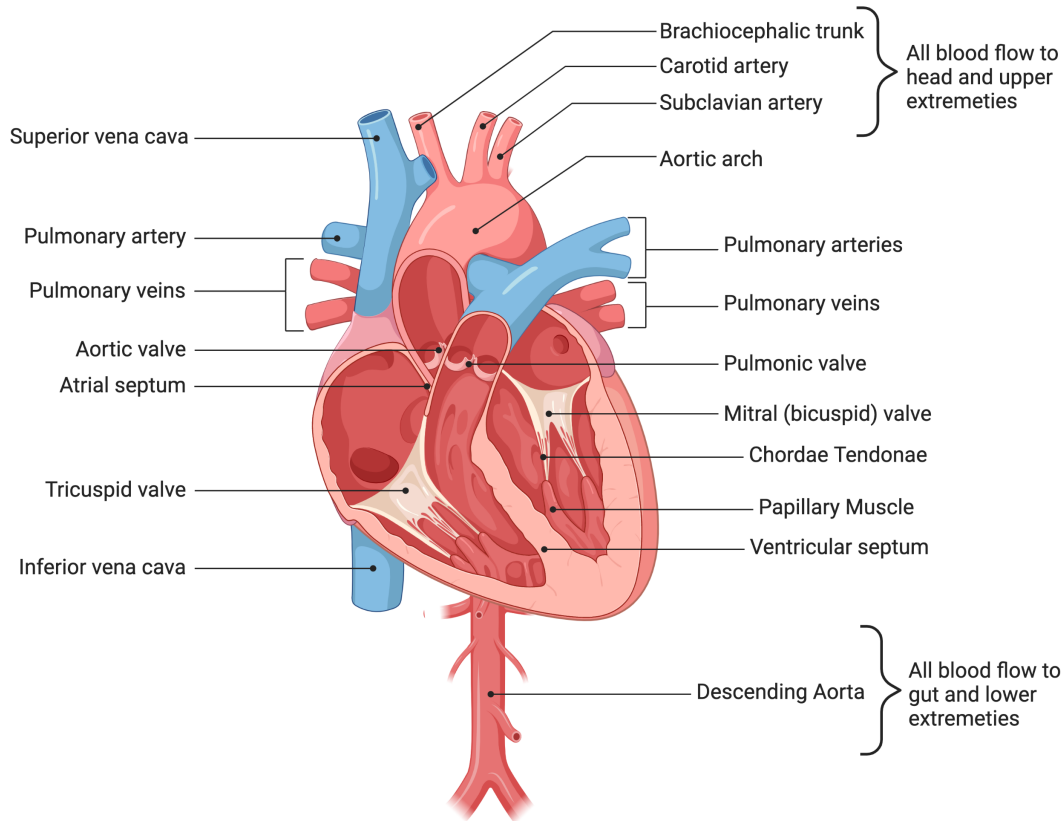


Fig. 9.1: Cardiac Anatomy (Created with BioRender.com)

Beyond being familiar with the structure in Figure 9.1 that are utilized throughout the chapter, there are several important structural takeaways.

- Structurally there are four chambers, right and left atria (RA and LA); right and left ventricle (RV and LV) (the anatomical view).
- The four chambers form two primer-pumps for two circulations (right and left), RA and RV for the pulmonary circulation; and LA and LV for the systemic circulation (which includes all skeletal muscles and the heart itself) (the pump view).
- Cardiac muscle (myocardium) is thicker in the ventricles than the atria; and on the left side than the right side.
- The four chambers are divided into two excitation regions, RA and LA for atrial excitation; and RV and LV for ventricular excitation (the excitation view).
- The amount of tension and excitation is related to the thickness of the myocardium; which is based on how much pressure that chamber must generate for circulation.
- Circulation between the right and left primer-pumps is prevented by a septum (septal wall), also myocardium, that is thinner than the outer myocardial wall; and also thicker in the ventricles than the atria.

- Circulation direction is guided by four one-way valves. Circulation occurs due to pressure gradients - there is nothing inherently directional about pressure gradients other than high pressure to low pressure - therefore the valves are necessary to keep blood flowing from a chamber to the next chamber and now "backwards".
- All circulation from the gut and lower extremity passes through the descending aorta; and comes to the heart through the inferior vena cava.
- All circulation from the upper extremities and head passes through the great vessels (first three branches off of the aorta); and comes to the heart from the superior vena cava.
- For blood to flow through the circulation, venous pressure must be lower than arterial pressure
- Venous pressure (systemic in the RA; pulmonary in the LA) refers to the pressure that is the final end point for all venous circulation returning to the heart which must, for circulation to occur, have the lowest venous blood pressure of the respective circulations ($P_{RA} = 0$ when the RA is not generating tension).
- Ventricular systolic pressure refers to the pressure in the ventricles during active tensioning (contraction), for circulation to occur it must have the highest arterial pressure of the respective circulations. The right sided ventricular systolic pressure is lower than the left sided ventricular systolic pressure, the respective circulations are considered low and high pressure.
- Because the ventricular systolic pressures must be the highest through the circulation, the tricuspid valve (between the RA and RV) and the bicuspid (mitral) valve (between the LA and LV) have additional support supplied by the papillary muscles and chordae tendoneae (the support does not help them open, it prevents them from allowing backflow while closed).
- Because the LV has greater pressure than the RR, the mitral valve papillary muscle and chordae tendoneae are more developed
- The aortic valve, which is not supported by papillary muscle, is under more pressure than the tricuspid valve because the aortic pressure is the second highest pressure in the circulation, however aortic peak pressure is short lived because of how rapidly blood flows out of the aorta.

9.2 Cardiac Output

Circulation is measured and assessed based on the cardiac output ($\dot{Q}$) in liters (or mL) per minute. $\dot{Q}$ is the flow of blood through the circuit, hence the circulation. $\dot{Q}$ is directly proportional to the overall pressure gradient and inversely proportional to resistance to circulation. A conceptual understanding of circulation based on ($\dot{Q}$) and the major determinants of, and factors influencing ($\dot{Q}$) facilitates an understanding of both normal and abnormal circulation.

$$\dot{Q} = \frac{P_a - P_{RA}}{TPR} \quad (9.1)$$

In Equation 9.1:

- P_a is pressure in the arteries (arterial pressure, which is highest in the aorta)
- P_{RA} is central venous pressure, commonly taken as the pressure in the right atrium (RA)
- TPR is the total peripheral resistance of the vascular circuit

Equation 9.1 is based on Ohm's Law and points attention to the influence of the driving forces for flow (the pressure gradient) and the resistance to flow. Based on the equation it is clear that the pressure gradient is directly proportional to flow (numerator); and resistance is inversely proportional to flow (denominator).

P_{RA} normally approximates to 0. $P_a - P_{RA}$ is equivalent to the term ΔP (change in pressure, pressure gradient) and when $P_{RA} = 0$, then $\Delta P = P_a$. However, in various clinical scenarios P_{RA} does not approximate to 0 so it is useful to keep central venous pressure in mind. If there is an increase in P_{RA} , without a rise in P_a or a drop in TPR then $\dot{Q}$ would be reduced.

An equation that represents the cardiac pump viewpoint of ($\dot{Q}$) is:

$$\dot{Q} = HR \times SV \quad (9.2)$$

In Equation 9.2:

- HR is heart rate in beats per minute (bpm)
- SV is stroke volume (the amount of blood pumped during systole) in mL

Equation 9.2 is also easier to confirm, by unit analysis, that these values equate to cardiac output $\dot{Q}$ in mL/min.

Combining Equations 9.1 & 9.2 results in:

$$HR \times SV = \frac{P_a - P_{RA}}{TPR} \quad (9.3)$$

The left side of Equation 9.3 is based on the output of the heart (cardiac) and the right side is based on the pressure gradient generated by the closed vascular circuit and the resistance encountered by the circulating blood. To keep the two sides of Equation 9.3 equal (balanced) an increase in HR must be balanced with resultant changes in at least one of the other parameters.

? Conceptual Understanding

What changes could occur if HR increases to keep Equation 9.3 balanced? What changes could occur if TPR decreased to keep Equation 9.3 balanced?

Equation 9.3 is a model and is useful to a point. A challenge with this equation, really all of the equations used in physiology, is that the variables are not independent of one another. They are useful for reasoning about the concepts, but not in an exhaustive way. For example, you may have just figured out that to keep Equation 9.3 balanced when HR increases, P_a could increase. It turns out that P_a is not independent of HR . If HR increases then P_a does tend to increase based on the relationship between arterial pressure and cardiac output.

9.2.1 Arterial Pressure (P_a)

Equation 9.1 simplified under the condition of $P_{RA} = 0$ can be manipulated to provide an equation for arterial pressure:

$$P_a = \dot{Q} \times TPR \quad (9.4)$$

Arterial pressure is blood pressure ($P_a = BP$). Equation 9.10 shows that blood pressure is directly proportional to cardiac output and total peripheral resistance (TPR). It is hoped that seeing the centrality of blood pressure in the equations for $\dot{Q}$ will inspire future clinicians to respect the potency and value of BP measurement in clinical practice. All measured changes in blood pressure (arterial pressure, P_a) can be inferred to represent changes in either cardiac output, or total peripheral resistance, or some combination.

Since the heart has two phases - a filling phase known as diastole, and a pumping phase known as systole - it creates pulsatile flow. Systole occurs when the SV is being ejected from the heart. Pulsatile flow creates a rhythmic pulse in BP with a peak value referred to as systolic BP (SBP) and a low value referred to as diastolic BP (DBP). It is useful to consider SBP and DBP separately, as well the mean arterial pressure (MAP):

Mean Arterial Pressure

$$MAP = DBP + \frac{1}{3} \times (SBP - DBP) \quad (9.5)$$

MAP is a weighted average blood pressure (average arterial pressure) based on the measured values SBP and DBP. DBP has a stronger influence on MAP than SBP because the time period that circulation is in systole is much shorter than the time they are in diastole. With an increase in heart rate (HR) the time in systole becomes proportionally more time but the systolic time interval is relatively unchanged. The change in overall time in systole is based on there being less time in diastole. Despite this there is no modification to the equation for MAP during periods of higher HR .

Pulse Pressure

Pulse pressure (PP) is often considered clinically as the driving pressure for blood flow. PP is calculated as the difference between SBP and DBP ($PP = SBP - DBP$).

Therefore Equation 9.5 for MAP, can be simplified to:

$$MAP = DBP + \frac{1}{3} \times PP \quad (9.6)$$

9.2.2 Poiseuille's Law

Poiseuille's Law offers deeper insight into $\dot{Q}$ and Equation 9.1. The benefit of this insight is the ability to break down the factors that determine TPR .

$$\dot{Q} = \frac{\Delta P \pi r^4}{\eta L} \quad (9.7)$$

In Equation 9.7:

- r^4 is vessel radius to the fourth power
- η the viscosity of the fluid (blood)
- L length of the vessel
- π is the mathematical constant that is the ratio of a circle's circumference to its diameter, approximately equal to 3.14159

In Equation 9.7 $\dot{Q}$ is directly proportional to the pressure gradient (ΔP) as in Equation 9.1 as well as to vessel radius to the fourth power (r^4); it is inversely proportional to the viscosity of the fluid (η) and the length of the vessel (L).

Poiseuille's Law includes all the components of Equation 9.1. ΔP is the pressure gradient between arteries (primarily aorta) and the right atria. Everything else in Equation 9.7 further elucidates components of TPR as resistance (R):

$$R = \frac{\eta L}{\pi r^4} \quad (9.8)$$

The physiologically relevant variables of Equation 9.7 for circulation are ΔP , r^4 and η . π is a constant and therefore does not contribute to our understanding of how circulation changes. And for all intents and purposes L is a constant since variation in vessel length is not utilized to regulate circulation.

9.3 Regulation of Cardiac Output

In this section we consider the factors that influence, and thus can be used to regulate, cardiac output ($\dot{Q}$) as a measure of circulation. Both the quantity (how many liters / minute) of blood being circulated as well as the distribution of that circulation throughout the body. Under normal circumstances all capillaries receive circulation. However, the proportion of circulation received in areas of the body is adjusted to meet local needs. The regulation of $\dot{Q}$ quantity and distribution is coordinated and includes overlapping regulated variables. In general, the quantity of cardiac output and its distribution is adjusted to meet the metabolic demands of the body. While the circulation serves many other purposes, it is the moment to moment need to deliver oxygen and remove carbon dioxide (which are metabolically determined) that predominates the need for circulation and cardiac output. The maximum value of $\dot{Q}$ represents the maximal functional capacity of the circulatory system to meet the demands of physical activity.

! Example of Quantity vs. Distribution: Vasovagal Syncope

Vasovagal syncope is passing out, losing consciousness, due to a low blood pressure due to an emotional trigger. It is an example of a situation where there is not appropriate coordination between the regulation of $\dot{Q}$ and its distribution. With

vasovagal syncope many areas of the body are attempting to get more circulation all at once. This results in widespread vasodilation that reduces TPR despite no increase in $\dot{Q}$. The BP drops enough that circulation to the brain is reduced and there is a temporary loss of consciousness. The loss of consciousness usually results in a movement (falling) and a position (on the ground) that increases $\dot{Q}$ by increasing the amount of blood returning to the heart. Once unconscious the emotional trigger is no longer present and the vagal outflow that triggered a drop in TPR is removed, and reflexes that increase TPR in response to a drop in BP kick in to increase TPR . The combination of increased $\dot{Q}$ and TPR increase BP and restore brain circulation (See Equation 9.10).

9.3.1 Cardiac Output: Quantity

Equipped with the concepts from the previous section the various ways in which the quantity of cardiac output, as a measure of circulation, is regulated can be addressed. Normal resting $\dot{Q}$ is approximately 4 to 8 liters per minute (L/min).

> Cardiac Index

$\dot{Q}$ also can be described relative to body mass as the cardiac index (CI), the amount of blood pumped per minute per square meter of body mass. Normal CI is between 2.5 and 4.2 $L/min/m^2$. This wide normal range makes it possible for $\dot{Q}$ to decline by almost 40% and still remain within the normal limits. Although several factors interrupt a direct correlation between CI and functional aerobic capacity, CI below 2.5 $L/min/m^2$ represents a marked disturbance in cardiovascular performance and is always clinically relevant.

Substituting Equation 9.7 of Poiseuille's law into Equation 9.3 for $\dot{Q}$:

$$HR \times SV = \frac{\Delta P \pi r^4}{\eta L} \quad (9.9)$$

Equation 9.9 provides all of the variables for consideration in the regulation of cardiac output.

9.3.1.1 Heart Rate times Stroke Volume ($HR \times SV$)

Heart rate and stroke volume (the left side of Equation 9.9 has previously been referred to as the viewpoint from the cardiac pump. The factors influencing these variables is covered briefly here and in more depth in Chapter 10. Normal resting $\dot{Q}$ is approximately 4 to 8 liters per minute (L/min), with a resting HR of 70 beats per minute (bpm); resting SV is approximately 71 mL/beat.

Heart rate is adjusted by the coordinated regulation of the autonomic nervous system (ANS). Both the sympathetic (adrenergic) and the parasympathetic (cholinergic) branches of the ANS influence the heart rate, as well as endocrine influences from the sympathetic ANS (epinephrine and norepinephrine). In general, the sympathetic nervous system increases heart rate (positive chronotropy) and the parasympathetic nervous system decreases heart rate (negative chronotropy).

Stroke volume is adjusted by changes in cardiac contractility, venous return, referred to as **preload**, and resistance to ejection from the heart, referred to as **afterload**. Cardiac contractility increases with sympathetic nervous system activity and through an increase in passive tension with increased preload (increased cardiac filling). Afterload varies directly proportionally with arterial pressure.

The factors that influence heart rate and stroke volume are discussed with more depth in Chapter 10 on the Cardiac Pump.

9.3.1.2 Pressure Gradient (ΔP)

The pressure gradient for cardiac output is regulated through the regulation of arterial pressure. The central venous pressure is kept at approximately 0 due to the **Frank-Starling law**. The Frank-Starling law is that the volume of blood

leaving the heart equals the volume of blood returning to the heart. Under that condition the pressure at the right atrium is kept near 0. The Frank-Starling law also contributes to the arterial pressure (systolic). If more blood returns to the heart (higher preload) then there is more contractility and a higher stroke volume which result in a higher systolic arterial pressure. The combination of an elevated systolic arterial pressure with a central venous pressure of 0 assures greater cardiac output (greater circulation flow). As we will see, the regulation of arterial pressure is the primarily monitored (sensed) variable for the homeostatic regulation of circulation.

9.3.1.3 Vessel Radius to the fourth power (r^4)

The radius of the circulation is a strong contributor to, and is inversely proportional to total peripheral resistance (TPR). The small arteries, arterioles and pre-capillary sphincters have the greatest capability of adjusting their radius. And, since these vessels make up the largest total cross sectional area of the vascular circuit, changing the radius of many arterioles and pre-capillary sphincters has a large effect on TPR (recall the events Vasovagal syncope described above). During muscular activity that requires an increased cardiac output for the delivery of more oxygen and removal of more waste products, an increase in the radius of local arterioles and pre-capillary sphincters ensures that the ΔP is maintained.

9.3.1.4 Blood Viscosity (η)

Viscosity (η) is directly proportional to resistance. An elevated hematocrit (larger proportion of blood is red blood cells and less is plasma) can increase blood viscosity. For example, an isosmotic volume contraction that increases hematocrit and plasma protein concentration would result in thicker blood which would create additional resistance to flow. In such circumstances, greater pressure is required to propel the higher viscosity blood.

9.3.2 Cardiac Output: Distribution

Regional demands for tissue perfusion (based on local metabolic needs) compete for circulation, and total quantity of $\dot{Q}$ adjusts to meet these demands however local blood vessels determine distribution. Distribution simply follows the path of least resistance. That path is determined by arterial vasomotor tone and the degree of vasodilation or vasoconstriction throughout all of the arterioles and pre-capillary sphincters. When the r^4 of the active muscle arterioles and pre-capillary sphincters increases, the resistance decreases. This both helps maintain ΔP and helps to distribute circulation to the area needing it the most.

Increased metabolic rate increases levels of CO_2 , decreases levels of O_2 , decreases pH, and increases levels of lactic acid at the tissue level dilate local blood vessels and therefore increase $\dot{Q}$ distribution to that area. The number of arterioles dilating due to local input is related to the drop in peripheral resistance and influences both venous return and the BP response, which then influences cardiac pump function.

9.4 Blood Pressure Regulation

With all of this talk about cardiac output as a measured variable of circulation the reader may come to think that the body has sensors that analyze the cardiac output. That is not the case. If cardiac output is not sensed, then it cannot be directly controlled. A homeostatic system requires that a variable be sensed to be regulated. Cardiac output is not directly sensed. There is no sensor in the aorta that keeps track of the mL/min of blood leaving the heart. Regulation of cardiac output occurs primarily through the regulation and adjustment of blood pressure.

From a regulation perspective Equation 9.10 is a starting point for understanding how circulation regulation is achieved and how blood pressure is the primarily monitored variable. Blood pressure is monitored directly by baroreceptors (mechanoreceptors) located in the heart, great vessels, and intrathoracic and cervical blood vessels; and indirectly in the kidneys. Monitoring of blood pressure in the kidneys occurs by monitoring the glomerular filtration rate (GFR) which is related to renal circulation.

Expanding Equation 9.10 by adding the variables that determine $\dot{Q}$ and TPR we can derive a useful equation for thinking about circulation regulation through blood pressure regulation.

$$BP = (HR \times SV) \times \frac{\eta L}{\pi r^4} \quad (9.10)$$

SV is the end diastolic volume (EDV) multiplied by the fraction of EDV ejected (EF) during systole:

$$BP = (HR \times EDV \times EF) \times \frac{\eta L}{\pi r^4} \quad (9.11)$$

Since η (viscosity) is relatively stable, L is very stable, and π is a constant:

$$BP = \frac{HR \times EDV \times EF}{r^4} \quad (9.12)$$

Vessel Compliance

The above reasoning has left out - for the sake of simplicity - the concept of compliance. The regulation of vessel radius occurs through changes in vascular tone (smooth muscle in vessel walls, primarily arterioles to a lesser extent venules). Increased vascular tone decreases the radius and decreases its compliance. Therefore the increase in pressure goes beyond the reduction in cross-sectional area associated with the change in the radius. Decreased vascular tone increases the radius and increases its compliance. Therefore the decrease in pressure goes beyond the increase in cross-sectional area associated with the change in the radius.

9.4.1 Neural Regulation of Blood Pressure

9.4.1.1 Autonomic Neuroendocrine Input

The autonomic neuroendocrine system influences the heart and blood vessels via direct neural connections and through circulating hormones. The parasympathetic (cholinergic) system has direct neural input through the cardiac branch of the vagal nerve (Cranial Nerve X). Vagal nerve activity generally decelerates cardiac function, thus decreasing HR , speed of the conduction wave of excitation and to a lesser extent contractility (rest and digest). The parasympathetic system does not have an endocrine component (no hormones).

Sympathetic (adrenergic) system neural input is through the thoracolumbar sympathetic system and increases HR and contractility, thus increasing cardiac output (fight or flight). The sympathetic nervous system also has connections to the peripheral vascular (arteries, arterioles, veins, venules). Therefore the sympathetic nervous system can change peripheral resistance, which changes vascular pressure and can alter venous return. The sympathetic system also activates the adrenal gland which then releases adrenaline and noradrenaline into the blood for a longer lasting endocrine (hormonal) response.

Semantics of the Autonomic Nervous System

The autonomic nervous system (ANS) is thus named because it is "self governing" - meaning, involuntary. Autonomic can sometimes be confused with automatic which is only true depending on what you mean by automatic. Automatic means capable of operating without external control. The autonomic nervous system is not really capable of operating without external control because it exists *in situ*.

The sympathetic branch of the ANS includes the hormones adrenaline and noradrenaline released from the adrenal glands (epinephrine and norepinephrine) and can also be referred to as the adrenergic nervous system. The receptors the adrenergic nervous system acts on are called adrenergic receptors.

The parasympathetic branch acts using acetylcholine (neurotransmitter, not a hormone) and is also called the cholinergic nervous system. The receptors it acts on are called cholinergic. Since the predominant parasympathetic nerve is the vagus nerve, this system is also occasionally referred to as the vagal nervous system - particularly in cardiac physiology.

? Benefit of Sympathetic Endocrine Response?

Can you explain what the benefit of the sympathetic endocrine response is, as opposed to just a nervous system component?

9.4.1.2 Baroreflex

The baroreflex is activated through a group of mechanoreceptors located in the heart, great vessels, and intrathoracic and cervical blood vessels. These mechanoreceptors are most plentiful in the walls of the internal carotid arteries. Mechanoreceptors are sensory receptors that are sensitive to mechanical changes, such as pressure and stretch. Increased activity in the mechanoreceptors by high pressures increases vagal output and decreases sympathetic output. This chain of events results in vasodilation, decreased HR , and decreased contractility. The net effect is lower HR and SV , which reduce $\dot{Q}$ and therefore BP . Reduction in activity in the mechanoreceptors by low pressure results decreases vagal output and increases sympathetic output. This chain of events results in vasoconstriction, increased HR , and increased contractility. The net effect is higher HR , SV (as compared to lower sympathetic activity), which increase $\dot{Q}$ and therefore BP . The baroreflex is a quintessential homeostatic negative feedback system - elevated BP promotes reduced BP , and reduced BP promotes elevated BP . It is the fast acting regulator of BP during body position changes that prevents syncope (loss of consciousness) during orthostatic challenges (moving from supine to sit or standing). Dysfunction of any part of the baroreflex can result in orthostatic intolerance (hypotension with position changes) that can include syncope. The baroreflex may also be involved in the poorly coordinated response to orthostatic challenges displayed in postural orthostatic tachycardia syndrome (POTS) where BP does not drop with position changes, but HR increases and is sustained at a high level for longer, which when utilized alone to regulate BP is not the most efficient or effective mechanism.

9.4.2 Renal Regulation of Blood Pressure

Renal regulation of blood pressure includes the maintenance of blood volume and a cascade of hormonal regulators that influence both blood volume as well as vessel radius. Blood volume has a direct impact on blood pressure. At the extreme, no blood in the vascular circuit, no pressure. Hypervolemia in the vascular circuit causes elevated pressure (unless, vasodilation can create enough space to avoid it). The role of the kidneys in BP regulation is a slower responding, and longer acting, regulatory system.

9.4.2.1 Renin Angiotensin II - Aldosterone System (RAAS)

The RAAS system (depicted in Figure 9.2) is initiated in the kidneys in response to reduced fluid volume that decreases BP , RBF and GFR. The kidneys produce and secrete the hormone renin. Renin activates angiotensin I which will be converted to angiotensin II by angiotensin converting enzyme (ACE). Angiotensin II directly influences kidney function by promoting both Na^+ and water reabsorption. It also activates aldosterone and ADH and to promote water and Na^+ reabsorption, and decrease tubular secretion. Angiotensin II also directly promotes vasoconstriction of vascular smooth muscle. The overall effect of these events is increased blood pressure through vascular tone and increased blood volume.

! RAAS Gone Wrong - Heart Failure & ACE-Inhibition

The events of RAAS depicted in Figure 9.2 can go wrong. With chronic heart failure (HF) there can be reduced BP , or even with preserved BP but lower cardiac output, reduced RBF and GFR, that promotes the secretion of renin by the kidneys. In the context of HF, RAAS increases return of blood to the heart (preload) due to an increased vascular volume. The increased preload adds additional burden to an already failing heart. RAAS also increases the load that the heart must pump against (afterload) due to an increased blood pressure and systemic vasoconstriction. The increased afterload also adds additional burden to an already failing heart. The result tends to be a further reduction in cardiac output and GFR

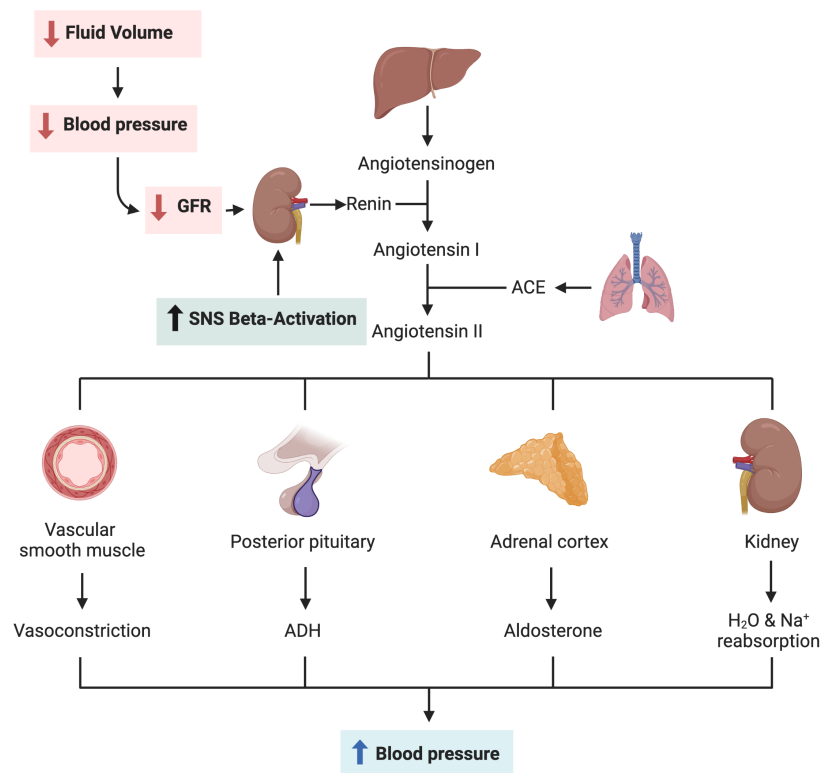


Fig. 9.2: Renin Angiotensin - Aldosterone System (RAAS) (Created with BioRender.com)

which promotes more renin secretion. This positive feedback loop is far from positive for the health of the individual with HF. In such situations the best course of action for restoring stability is a class of medication referred to as ACE-Inhibitors (ACE-I, i.e. captopril, lisinopril, enalapril) that block the enzyme that converts angiotensin I to angiotensin II. ACE-I are also a preferred treatment for HTN (high blood pressure). Rather than forcing diuresis (as a diuretic does, which can create hypokalemia), they prevent anti-diuresis. Anti-diuresis is a gentle and physiologically stable approach to diuresis that prevents one of the several deleterious responses to reduced GFR due to HF. The therapeutic benefit of ACE-I goes beyond preventing anti-diuresis and includes reducing the release of aldosterone, and preventing angiotensin II from directly vasoconstricting blood vessels (increasing total peripheral resistance (TPR) and diastolic blood pressure (DBP)), and directly resulting in water and Na^+ reabsorption.

9.4.2.2 Overall Integration of Renal Mechanisms and Arterial Pressure

Figure 9.3 is a directed graph that depicts the closed loop integration of renal and circulatory systems through the renal regulation of blood volume, and the arterial pressure influence over renal excretion.

The variables in Figure 9.3 are familiar other than the mean circulatory filling pressure (P_{mcf}). P_{mcf} is a variable created by Guyton in a model of renal blood pressure regulation [Rot93]. P_{mcf} is defined as the mean vascular pressure that exists with a stop in cardiac output and distribution of blood until all pressures are the same throughout the system. It is based on the fullness of the circulatory system. A change in P_{mcf} is a useful index of a change in overall venous smooth muscle tone if the blood volume is not changed. P_{mcf} estimates the pressure in the small veins and venules, which

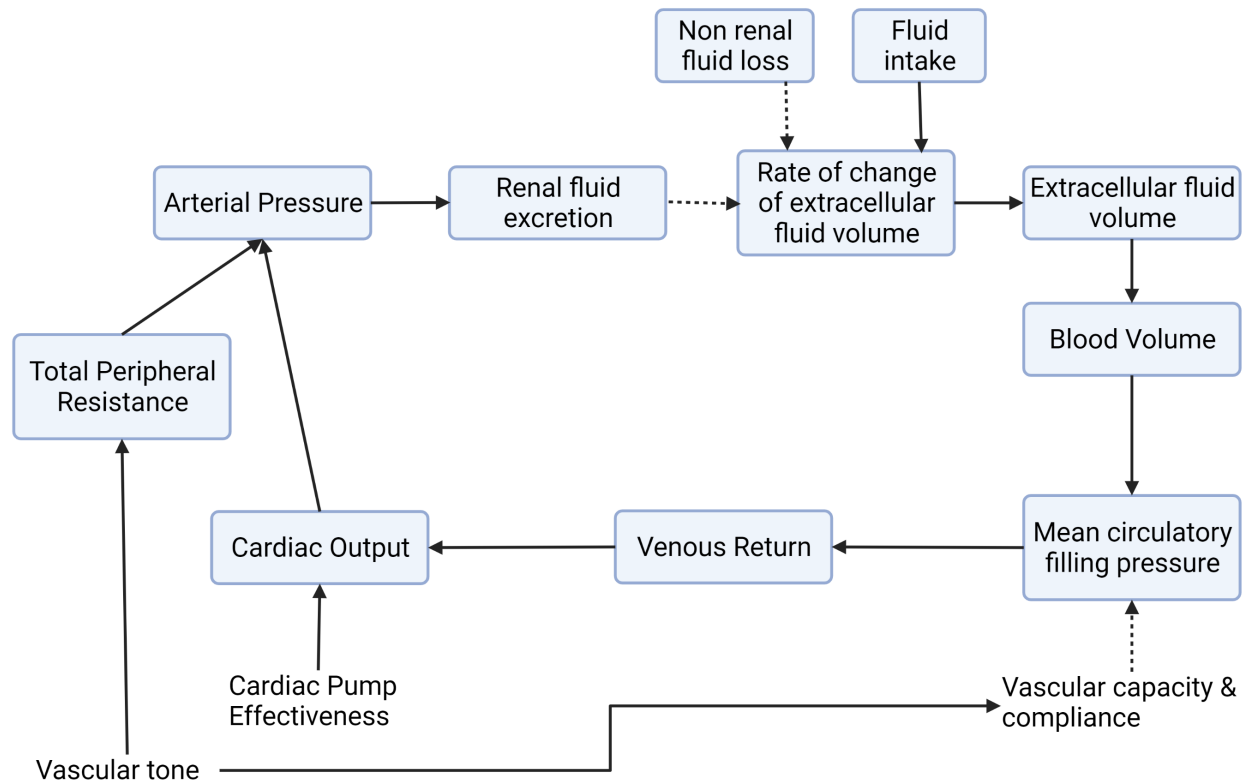


Fig. 9.3: Integration of Renal Mechanisms with Arterial Pressure (Modified from [HH20] with BioRender.com)

contain most of the blood and account for most of the compliance. In total, P_{mcf} provides an estimate of the pressure that determines the rate of flow returning to the heart, independent of cardiac output.

9.5 Cardiovascular Vital Signs

9.5.0.1 Palpation of Pulses

Palpation of pulses is an important component of a physical examination and can be used to evaluate and identify the following:

- Circulation quality
- Pulse rate and rhythm

Circulation Quality

Pulse are pressure waves through the circulation that are related to both blood pressure and local circulation. Weak pulses can indicate either low blood pressure (overall) or poor circulation that particular region of the body. For example, a weak pulse in the right radial artery but not the left radial artery most likely indicates a local circulation problem in the left UE but not the right. Whereas a weak pulse in all extremities is most likely a low blood pressure overall. A strong, or bounding, pulse can indicate a high blood pressure overall, or a high blood pressure in a particular area. Bounding pulses can also indicate vascular abnormalities such as aneurysms (for example a bounding abdominal pulse is a sign of a descending aortic aneurysm).

Heart Rate & Rhythm

Pulses are correlated with heart rate (cardiac cycle) and rhythm. When palpating a pulse to obtain *HR*, counting the pulse rate for 15 seconds and multiplying by 4 is sufficient with normal rates and rhythms. If rates are faster than 100 bpm or slower than 60 bpm, palpate the pulse for 60 seconds. If the rhythm is irregularly irregular (e.g., during AFib) or regularly irregular (e.g., PACs or PVCs), perform auscultation of heart sounds to identify the apical *HR* for a full minute. In these cases, palpation of pulse cannot substitute for ECG analysis to monitor the patient's rhythm, but it may alert the therapist to the onset of these abnormalities.

Use caution in palpating pulses because manual pressure on the carotid sinus may cause baroreflex drops in heart rate (*HR*), blood pressure (*BP*), or both. In all pulse palpation situations the examiner should start with light pressure and gradually increase pressure until a pulse is felt, and to not exert more pressure than required to feel the pulse.

Some people can feel a pulse in their thumb, therefore it is generally advised to not use the thumb to palpate a pulse.

HR is the primary means of determining the exercise intensity level for patients who are not taking beta-blockers or who have rate-responsive pacemakers. *HR* combined with blood pressure (*RPP*) is the most appropriate approach to determine exercise intensity in any patient with a known ischemic threshold (even if they are on beta-blockers).

- A linear relationship exists between *HR* and work.
- In general, a 20- to 30-beat increase from the resting value during activity is a safe intensity level in which a patient can exercise.
- If a patient has undergone an exercise stress test during the hospital stay, a percentage (e.g., 60%–80%) of the maximum *HR* achieved during the test can be calculated to determine the exercise intensity.
- An example of a disproportionate *HR* response to low-level activity (bed or seated exercises or ambulation in room) is an *HR* of more than 120bpm or less than 50bpm.
- Heart rate recovery (*HRR*), provides an indication of reduced parasympathetic activity and an indicator of all-cause mortality, can be used to document improvement of tolerance to functional demands. *HRR* is the absolute difference between peak *HR* achieved with exercise minus the *HR* at 60 seconds after the completion of exercise (*HRR* 60 sec). An abnormal *HRR* at 1 minute, after a treadmill test, is reported to be a decrease of 12 bpm or less with a cool-down period and less than 18 bpm without a cool-down period [CD19].
- When prescribing activity intensity for a patient taking beta-blockers, consider that *HR* should not exceed 20 beats above the resting *HR*. If *RPP* for the patient's ischemic threshold is known, then the *RPP* should not be allowed to increase to the ischemic threshold.
- If prescribing an activity intensity, with use of *HR*, for patients with an automatic implantable cardiac defibrillator (AICD), remember that the exercise target *HR* should be 20 to 30 beats below the threshold rate on the defibrillator.

- *HR* should not be used to prescribe exercise status post heart transplantation secondary to denervation of the heart during transplantation.
- Baseline *HR* and recent changes in medications always should be considered before beginning an exercise session.

9.5.0.2 Blood Pressure

Blood pressure is critical to circulation and a primary variable utilized to help regulate cardiac output. Low blood pressure (and low PP) is problematic because it compromises circulation and may indicate hemodynamic compromise due to cardiac (reduced cardiac output), vascular (reduced peripheral resistance) or blood volume problems. High blood pressure can create immediate health care emergencies by placing blood vessels under too much pressure; or cardiac emergencies by causing ischemia (high RPP and myocardial O_2 demand), or heart failure (increased afterload). Long term increased high blood pressure (hypertension) leads to long term changes in the heart (left ventricular hypertrophy, fibrosis of the atria and ventricles that makes the heart more susceptible to arrhythmias, or valve dysfunction). Hypertension is a silent killer since many individuals have no idea they have it - which is one of the reasons for the APTA Academy of Cardiovascular and Pulmonary Physical Therapy's "Vitals are Vital" campaign to encourage all physical therapists in all settings to make sure they routinely examine vitals signs in all clients.

Measurement of *BP* with a sphygmomanometer (cuff) and auscultation is an indirect, noninvasive measurement of the force exerted against the arterial walls during ventricular systole (SBP) and during ventricular diastole (DBP). *BP* is affected by total peripheral resistance (blood volume, radius and compliance of arterial walls) and $\dot{Q}$. Table 9.1 includes normal *BP* ranges.

Ranges	Systolic (mmHg)	Diastolic (mmHg)
Age 8 years	85-114	52-85
Age 12 years	95-135	58-88
Adult	< 120	< 80
Elevated (Adult)	120-129	< 80
Stage 1 Hypertension	130-139	80-89
State 2 Hypertension	≥ 140	≥ 90

Table 9.1: Blood Pressure Ranges, Modified from [CD19]

Occasionally, *BP* measurements can be performed only on certain limbs secondary to the presence of such conditions as a percutaneously inserted central catheter, arteriovenous fistula for hemodialysis, blood clots, or lymphedema (e.g., status post mastectomy).

BP of the upper extremity should be measured in the following manner [CD19]:

1. Check for posted signs, if any, at the bedside that indicate which arm should be used in taking *BP*. *BP* variations of 5 to 10 mmHg between the right and left upper extremity are considered normal. Patients with arterial compression or obstruction may have differences of more than 10 to 15 mmHg.
2. Use a properly fitting cuff. The inflatable bladder should have a width of approximately 40% and length of approximately 80% of the upper arm circumference.
3. Position the cuff 2.5 cm above the antecubital crease.
4. Rest the relaxed arm at the level of the heart.
5. To determine how high to inflate the cuff, palpate the radial pulse, inflate until no longer palpable, and note the cuff inflation pressure. Deflate the cuff. This is an estimate of the systolic pressure. But this method cannot be utilized to obtain a diastolic pressure.
6. Place the bell of the stethoscope gently over the brachial artery.
7. Reinflate the cuff to 30 to 40 mmHg greater than the value in step 5. Then slowly deflate the cuff. Cuff deflation should occur at approximately 2 to 3 mmHg per second.¹⁵
8. Listen for the onset of tapping sounds, which represents circulation returning to the brachial artery. This is the SBP.
9. As the pressure approaches diastolic pressure, the sounds will become muffled and in 5 to 10mmHg will be completely absent. These sounds are referred to as Korotkoff sounds (See Table 9.2).

Phase	Korotkoff Sound	Indicates
1	First sound, faint tapping	Systolic pressure (circulation through compressed artery)
2	Blowing or swishing sound	Blood is increasing
3	Distinct tapping	Circulation is increasing
4	Muffled	Diastolic pressure in certain situations ^a
5	Disappearance	Diastolic pressure in adults

Table 9.2: Blood Pressure Ranges (^a Phase 4 represents DBP in children, and adults that are exercising, pregnant or with hyperthyroid conditions)(Modified from [BS12])

a

Systolic Blood Pressure / Palp

In situations when it is difficult to auscultate or to obtain a distinct diastolic blood pressure reading (DBP), the palpation method may be utilized and blood pressure is noted as systolic blood pressure (SBP)/Palp. This is the same procedure utilized in Step 5 to determine how high to pump the cuff for a full measurement of blood pressure without having to over inflate the cuff.

9.5.1 Physical Therapy Considerations

- Recording pre exertion, para exertion, and post exertion *BP* is important for identification of *BP* responses to activity. During recovery from exercise, blood vessels dilate to allow for greater circulation to muscles. In cardiac-compromised or very deconditioned individuals, total $\dot{Q}$ may be unable to support this increased flow to the muscles and may lead to decreased output to vital areas, such as the brain.
- If you are unable to obtain *BP* on the arm, the thigh is an appropriate alternative, with auscultation at the popliteal artery.
- Falsely high readings occur if the cuff is too small or applied loosely or if the brachial artery is lower than the heart level.
- Evaluation of *BP* and *HR* in different postures can be used to monitor orthostatic hypotension with repeat measurements on the same arm 1 to 5 minutes after position changes.
- The same extremity should be used when serial *BP* recordings will be compared for an evaluation of hemodynamic response.
- A *BP* record is kept on the patient's vital sign flow sheet. This is a good place to check for *BP* trends throughout the day and, depending on your hospital's policy, to document *BP* changes during the therapy session.
- An auscultatory gap is the disappearance of sounds between phase 1 and phase 2 and is common in patients with high *BP*, venous distention, and severe aortic stenosis. Its presence can create falsely low SBPs if the cuff is not inflated enough (prevented by palpating for the disappearance of the pulse before measurement), or falsely high DBPs if the therapist stops measurement during the gap (prevented by listening for the phase 3 to phase 5 transitions).
- In general there is a lower SBP, higher DBP, and lower workload at any given submaximal *HR* during arm exercise compared with leg exercise [Dia+22]
- A clearly disproportionate response to exercise includes SBP decrease of 10 mmHg below the resting value, a hypertensive systolic response of greater than 200 mmHg, or a hypertensive diastolic response of greater than 110 mmHg. A normotensive systolic blood response should increase 5 to 12 mmHg per increase in metabolic equivalents (METs). One MET is approximately 3.5 ml/kg/min of oxygen consumption.
- If the patient is on a pacemaker that does not have heart rate modulation, *BP* response can be used to gauge intensity.

9.6 Summary

This chapter has provided introduced circulation from the perspective of cardiac output. The focus has been on the systemic circulation, what leaves the left ventricle and returns to the right atria without much consideration of what the

heart is doing in its role as a pump. We simply took it for granted that the heart pumps a volume of blood out with each beat (SV) and it does so a certain number of times per minute (HR). Circulation is measured by us as cardiac output but is sensed physiologically by the indirect means of BP . Homeostatic mechanisms are in place that allow the monitoring of BP to adjust of HR (chronotropy), SV through cardiac contractility (inotropy) and venous return (vascular tone through neuroendocrine mechanisms), blood volume (renal and renal endocrine mechanisms) and TPR . Monitoring BP and HR provide physical therapists with non-invasive insights. When combined with an understanding of circulatory physiology these insights provide useful information when working with clients both with and without conditions that directly affect the circulation.

Next Step - The Cardiac Pump

In Chapter 10 on the Cardiac Pump we fill in the gap of circulation between the blood entering the right atrium and the blood leaving the left ventricle. This includes the factors the influence heart rate, stroke volume; properties of the cardiac muscle; cardiac circulation and the electrocardiograph to assess cardiac function.

References

- [BS12] Lynn Bickley and Peter G Szilagyi. *Bates' guide to physical examination and history-taking*. Lippincott Williams & Wilkins, 2012.
- [CD19] Sean M Collins and Konrad J. Dias. *Cardiac System*. Acute care handbook for physical therapists. Ed. by Jaime Paz and Michele West. Elsevier Health Sciences, 2019.
- [Dia+22] Konrad J. Dias et al. *Differences in Blood Pressure and Workload Intensity During Arm and Leg Ergometry at Submaximal Target Heart Rates in College-Aged Individuals*. en. *Cardiopulmonary Physical Therapy Journal* 33.1 (2022), pp. 31–39. ISSN: 1541-7891. DOI: [10.1097/CPT.000000000000181](https://doi.org/10.1097/CPT.000000000000181). URL: <https://journals.lww.com/10.1097/CPT.000000000000181> (visited on 10/06/2021).
- [HH20] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences, 2020.
- [Rot93] C. F. Rothe. *Mean circulatory filling pressure: its meaning and measurement*. en. *Journal of Applied Physiology* 74.2 (Feb. 1993), pp. 499–509. ISSN: 8750-7587, 1522-1601. DOI: [10.1152/jappl.1993.74.2.499](https://doi.org/10.1152/jappl.1993.74.2.499). URL: <https://www.physiology.org/doi/10.1152/jappl.1993.74.2.499> (visited on 07/12/2022).

Sample Questions

1. Based on Poiseuille's Law which factor has both a strong and a highly variable (physiologically adjustable) determinant of peripheral resistance?
2. How is central venous (right atrial) pressure kept at approximately 0?
3. What factors allow the distribution of cardiac output to respond to metabolic demands?
4. What are the implications associated with changes in vascular tone based on altered sympathetic nerve activity change vessel radius and compliance in the same direction (both increase or both decrease)?
5. What is the sequence of events based on the baroreceptor response to standing up from supine?

Chapter 10

Cardiac Pump

Updated on May 22, 2024

Contents

10.1 Cardiac Muscle	157
10.2 Cardiac Cycle	167
10.3 ECG Rhythm Analysis	169
10.4 Summary	171

This chapter covers the cardiac pump in its role to support circulation. Circulation is completely dependent on the pumping action of the cardiac muscle to create the pressure gradient for both systemic, and a bit less so, the pulmonary circulation.

The relationship between circulation and cardiac muscle is similar to the relationship between movement and skeletal muscle. Concepts from skeletal muscle such as tension, excitation, regulation and energetics are applicable to cardiac muscle and are included. It is useful to compare and contrast cardiac muscle with skeletal muscle across these concepts while focusing on how the differences enable the cardiac muscle to meet the unique needs of circulation. The primary goal of this chapter is an understanding of cardiac pump in its supportive role for circulation.

Objectives include:

1. Compare and contrast cardiac and skeletal muscle based on tension, excitation, regulation and energetics.
2. Explain how the unique characteristics of cardiac muscle tension, excitation (automaticity, rhythmicity, conductivity, contractility), regulation and energetics support circulation.
3. Explain factors that influence cardiac output (including preload, afterload, inotropy, chronotropy).
4. Explain the relationship between RPP and myocardial oxygen demand.
5. Explain cardiac circulation.
6. Explain the cardiac cycle (to the point of drawing and explaining all aspects of Wiggers Diagram).
7. Perform an ECG rhythm analysis including the identification of arrhythmias and myocardial circulation abnormalities.

10.1 Cardiac Muscle

The role of the heart as a pump is based on its ability to regularly provide the pressure necessary for circulation of blood through arterial vessels (both systemic arteries (which includes the cardiac circulation) and pulmonary arteries). The function of the heart as a pump depends on several characteristics of cardiac muscle:

- Automaticity: The ability to initiate its own excitation

- **Rhythmicity:** The ability to repeat the its own excitation cycle in synchrony with regularity
- **Excitability:** The ability to respond to excitation with activation
- **Conductivity:** The ability to transmit excitations from cell to cell within the heart
- **Contractility:** The ability to generate a full contractile (shortening) cycle with each excitation

The next four sections parallel the chapters on skeletal muscle function of Part II: Tension, Excitation, Regulation and Energetics. They, in turn, compare and contrast cardiac and skeletal muscle in these key areas and also how the cardiac muscle and pump are thus uniquely capable of supporting circulation. As a reminder Chapter 9 on Circulation focused on circulation leaving the left ventricle and returning to the right atrium. This chapter is focused on the circulation through the heart itself as it is pumped and travels from the right atrium to the left ventricle. There are, however, two caveats. First, this chapter includes details about the cardiac circulation within the section on energetics (since cardiac muscle energetics is very dependent on the regular delivery of oxygen via blood flow). Second, this chapter does not include details on the pulmonary circulation (leaving the right ventricle and returning to the left atrium). The pulmonary circulation is covered in the next chapter on Respiration.

10.1.1 Tension

It is reasonable to use the term contraction when describing activation and active tension for the cardiac muscle because contraction (shortening) is what occurs (there is no eccentric or isometric activation of cardiac muscle). A major difference between cardiac and skeletal muscle is that cardiac muscle does not have a tendon attachment. Cardiac muscle attaches to itself to create a chamber that has a volume. When cardiac muscle contracts, it makes the chamber smaller (less volume). In skeletal muscle the length of fibers translates into the length of the entire muscle and this relationship has a functional interpretation. In cardiac muscle the length of fibers directly translates into the volume of chambers. It is the volume of these chambers that has a functional interpretation.

The heart has four chambers that function as four pumps (right and left atria; right and left ventricles). From a muscular perspective there are two chambers, each separated into a right and a left side with a relatively thin muscular barrier between them. This is quite noticeable in conditions that result in abnormal timing of the right and left ventricle excitation. If the ventricles are not excited at the same time, off even by 50 milliseconds, the overall pump effectiveness of both ventricles is impaired (mechanical cardiac asynchrony).

Active Tension

Active tension is generated in cardiac muscle with the same sliding filament model of skeletal muscle described in Chapter 3 on Tension. Active tension generates tension that shortens cardiac muscle fibers and makes the cardiac chambers smaller (and less compliant).

Passive Tension

Passive tension is generated by titin and other connective tissue structures that resist lengthening of sarcomeres, which in turn resists enlarging the chamber. Passive tension also attempts to shorten cardiac muscle fibers when they are stretched (lengthened fibers when the chamber has a larger volume). Additional passive tension is developed by layers of connective tissue surrounding the heart including the epicardium (visceral layer of the serous pericardium), the parietal layer of the serous pericardium, and the fibrous pericardium (See Figure 10.1). The pericardial cavity is filled with pericardial fluid which reduces friction during the cardiac cycle of contraction - relaxation. All of these structures, including the amount of fluid in the pericardial fluid, impact the compliance of the heart chambers.

Volume-Tension Curve

Cardiac muscle fibers *in vitro* have a length-tension curve that is similar to skeletal muscle fibers. The primary difference is that the passive component is shifted to the left and is a steeper exponential curve. The active component displays a

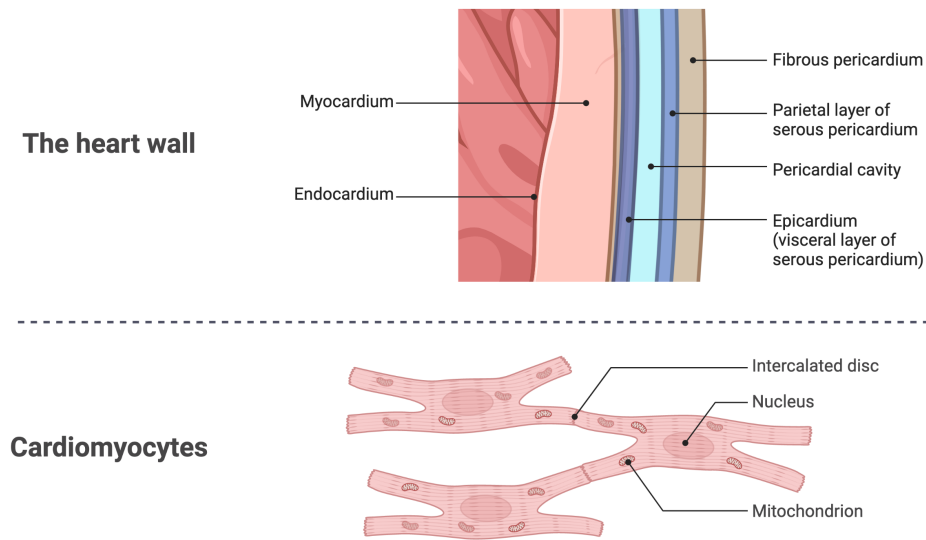


Fig. 10.1: Cardiac Muscle (Created with BioRender.com)

similar curve based on the overlap of actin and myosin. The passive component displays that cardiac muscle has a stiffer form of titin that contributes to tension with a smaller increase in length.

Cardiac muscle *in situ* has a **volume-tension curve** that is synonymous with the length-tension curve of whole skeletal muscle. The cardiac volume-tension curve is similar looking to its length-tension curve, the primary difference again being the passive component is shifted further to the left and is steeper. The volume-tension curve is influenced by the length-tension of the cardiac muscles and the passive structures surrounding the heart that resist changes in the chamber size directly. Passive tension starts to develop as the cardiac chamber fills. The functional implication is that under normal resting conditions passive tension contributes to cardiac contraction. And with larger volumes of blood in the chambers the contribution of passive tension rises exponentially.

The passive component of the volume-tension curve is related to cardiac compliance. As described in Chapter 7 there is a relationship between volume, compliance and pressure. The cardiac muscles vary the volume of chambers while having a particular compliance (that varies with the passive and active tension) for the purpose of generating pressure. This is a fundamental difference between cardiac and skeletal muscle tension. Skeletal muscle tension changes length based directly on tension for the purposes of generating torque. Cardiac muscle tension changes length for the purpose of changing volume and compliance, for the purpose of changing pressure. It is the pressure generated in the cardiac chambers that result in circulation.

Understanding cardiac function requires a familiarity with the equations relating compliance, volume and pressure introduced in Chapter 7, repeated here specific for the heart:

$$\Delta P_{chamber} = \frac{\Delta V_{chamber}}{C_{chamber}} \quad (10.1)$$

In Equation 10.1 the dependency of pressure on volume and compliance is clear. Since circulation is a cyclic variation in pressure gradients through the cardiac pump (and vascular vessels), the entirety of cyclic circulation through the heart can be considered from the perspective of cyclic changes to volume and compliance that develop pressure gradients to create flow through and out of the heart. Consideration of volume and compliance for the function of the heart as a pump correctly orients the discussion for both active and passive tension since they both contribute to cardiac pump function.

10.1.1.1 Preload

Preload is the term given for the pressure in a chamber when the cardiac muscle is filled (circulation into the chamber is approaching 0) but is not yet activated. Preload is the passive tension in the chamber right before it develops active

tension. Preload for cardiac muscle is analogous to stretch on skeletal muscle. The pressure of preload is related to the volume and the compliance. The compliance is at its highest when the myocardium is empty and not activated. As the chamber fills with blood, passive tension starts to develop which gradually reduces compliance while increasing the volume of blood, which when combined increases pressure. Under normal circumstances this increase in pressure (higher blood volume and lower compliance) is not great enough to fully impede circulation into the chamber. The passive tension developed during preload is subsequently utilized to assist active tension. Which develops the tension necessary to change the volume and compliance of the chamber for the purpose of increasing pressure to facilitate circulation out of the chamber. The increased pressure, alone, simply directs circulation out of the chamber from an area of high to low pressure. The direction of actual circulation through the heart is dependent on properly functioning valves that limit flow "backwards" and allow circulation "forwards" based on cardiac anatomy (Figure 9.1).

The preload is directly related to end-diastolic pressure (EDP), which is directly related to end-diastolic volume (assuming the compliance is normal and relatively low during diastole). The volume that results in preload is from circulation. Where the circulation is coming from for preload depends on the chamber. The right atrium preload comes from the systemic veins; right ventricle preload comes from the right atrium through the tricuspid valve; left atrium preload comes from the pulmonary veins; left ventricle preload comes from the left atrium through the mitral valve.

It is most common to discuss preload from the perspective of the left ventricle (LV). Therefore, unless otherwise noted, EDV and EDP mean LVEDV and LVEDP (respectively). EDP is related to the **Frank-Starling law** which states that the amount of venous return influences the performance of the left ventricle (assumes all the venous return moves through the right heart circulation (RA, RV, pulmonary arteries, pulmonary capillaries, pulmonary veins to the LA and finally the LV). The Frank-Starling mechanism is based on the relationship between EDP and total tension developed from active and passive tension.

10.1.1.2 Afterload

Afterload is the term given for the pressure that a chamber must overcome to cause circulation out of the chamber. Afterload for cardiac muscle is analogous with resistance (or load) for skeletal muscle. It is the resistance that the cardiac chamber must exert its tension in order to reduce the volume of the chamber. The pressure of afterload depends on which chamber must overcome that pressure. For the RA it is the pressure exerted on the tricuspid valve by the RV; for the RV it is the pressure exerted on the pulmonic valve by the pulmonary arteries; for the LA it is the pressure exerted on the mitral valve by the LV; for the LV it is the pressure exerted on the aortic valve by the aorta (first of the systemic arteries). It is most common to discuss afterload from the perspective of the LV. Therefore, unless otherwise noted, afterload refers to aortic pressure and is best estimated as mean arterial pressure (MAP) (See Equation 9.6). MAP is related to vascular compliance and resistance. As afterload, MAP affects aortic valve opening and is the most obvious load encountered by the ejecting ventricle.

! HFpEF

Long term high blood pressure (hypertension) results in left ventricular hypertrophy (LVH) that is associated with a reduced size and reduced compliance of the LV chamber. LVH is a common cause of diastolic (preserved ejection fraction (pEF)) heart failure (HFpEF).

Ejection Fraction

The amount of blood ejected from the LV during systole is called stroke volume (SV). It should be clear that SV is impacted by both preload and afterload. SV is simply the EDV minus the end systolic volume (ESV) ($EDV - ESV$). The percentage of blood in the chamber (EDV) that is ejected during systole is the ejection fraction:

$$EF = \frac{SV}{EDV} \quad (10.2)$$

A useful equation to remember for SV is therefore:

$$SV = EDV \times EF \quad (10.3)$$

Equation 10.3 highlights how the two phases of the cardiac cycle (systole and diastole) contribute to stroke volume. Diastole contributes with a volume of blood (EDV). Systole contributes the tension to eject a fraction of that blood (LVEF). As discussed above, LVH can cause something called diastolic or preserved EF heart failure (pEF). That is because LVH reduces SV by reducing the size of the LV chamber. It reduces EDV, but does not reduce LVEF. There are also causes of HF that do reduce LVEF and they are called systolic or reduced EF heart failure (rEF). There are also situations - most often due to comorbidities (having more than one condition) - that cause HF with both diastolic (reduced EDV) and systolic (reduced LVEF) components.

10.1.2 Excitation

To initiate a contraction cardiac muscle must undergo excitation. Excitation of cardiac muscle includes several similarities, but also differences, with skeletal muscle excitation.

10.1.2.1 Automaticity & Rhythmicity

Cardiac muscle is self excitatory (automaticity). Self excitations are rhythmic, occurring at rest approximately every 2/3 of a second (650 ms) with very little variability when not influenced by the neuroendocrine system (HR of approximately 90 bpm with low heart rate variability (HRV)). Or every 60 seconds with more variability when influenced by the neuroendocrine system at rest (HR of approximately 60 bpm with a HRV of approximately 50-100 ms when measured as a standard deviation with primarily parasympathetic influence).

In normal conditions automaticity occurs in a cluster of myocardial cells in the right atrium known as the **sino atrial (SA) node** (pacemaker cells). SA node cells have a higher resting permeability to Na^+ than axonal or skeletal muscle membranes. Based on this higher permeability to Na^+ these cells do not have a sustained (stable) resting membrane potential (See Figure 10.2). As soon as their membrane is repolarized it begins a slow journey back to the threshold potential due to a slow influx of Na^+ . How long it takes to reach the threshold (pacemaker potential) determines the heart rate. Under normal conditions the SA node cells have the greatest pacemaker potential.

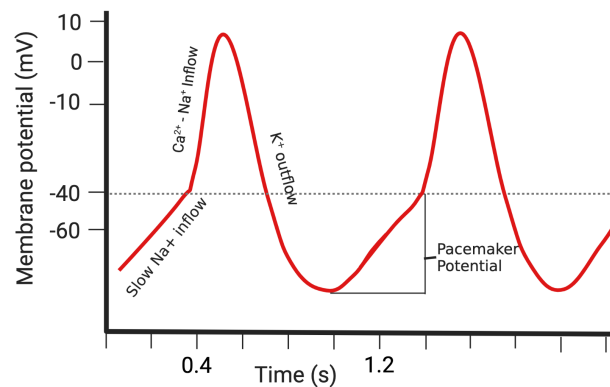


Fig. 10.2: Cardiac Sino Atrial Node Excitation (Created with BioRender.com)

10.1.2.2 Conductivity

An excitation of the SA node results in full excitation of all the cardiac muscle cells, first in the atria and after a delay, in the ventricles. There are three features of the myocardium that lead to this pattern of conductivity.

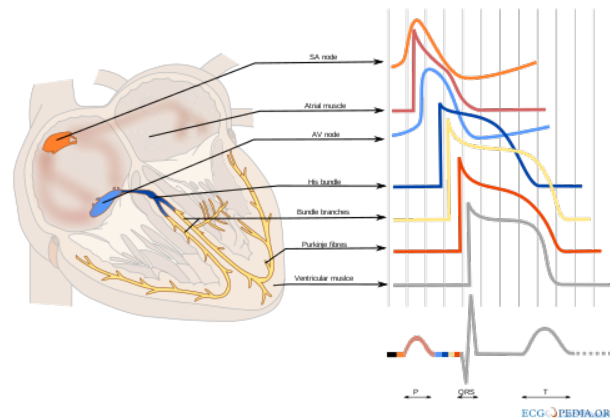


Fig. 10.3: Cardiac Conduction & Shapes of Excitation (Created with BioRender.com)

1. There are specialized cardiac cells that form a conduction network (See Figure 10.3. The conduction network starts with the SA node sending a wave of excitation through the atria, facilitated by intra-nodal fibers (not shown in Figure 10.3) that can rapidly carry an excitation from the SA node to the atrio-ventricular (AV) node. The excitation through the atria results in the P wave of the electrocardiograph (ECG). The excitation of the AV node and its distal conduction component, the bundle of His, is slower and delays the excitation of the ventricles (the PR interval on the ECG). Once through the AV node and bundle of His the excitation travels rapidly down the right and left bundle branches to the purkinje fibers. The branching conduction network in the ventricles is critically important to rapid and synchronous excitation of the ventricles (the QRS complex in the ECG). Atrial repolarization is a minor event and is easily lost in the waves of the QRS. Ventricular repolarization is a more substantial event because the ventricles are more substantial muscles and results first in a return to baseline (ST segment on the ECG), and then the T wave on the ECG. It should be noted in Figure 10.3 that the AV node fibers have a pacemaker potential but it is smaller than the SA node pacemaker potential. The SA node is the pacemaker under normal circumstances because it has a larger pacemaker potential and reaches the excitation threshold earlier than the AV node. However, if the SA node fails, the AV node reaches the excitation threshold and initiates ventricular excitation.
2. There is a fibrous connective tissue boundary between the atria and the ventricles that does not transmit an excitation (does not have an excitable membrane). This helps ensure that ventricular excitation occurs only through excitation of the AV node (and no other way). The AV node ensures that the ventricular contraction occurs after a delay so that there is time for atrial contraction to push the final 20-30% of blood into the ventricles for the final stretch of the ventricles which facilitates more tension, thus more pressure for circulation. The "atrial kick" as this is referred facilitates greater tension and thus pressure during ventricular contraction.
3. Each excitation travels from cardiac muscle fiber to the surrounding cardiac muscle fibers due to the presence of intercalated discs that have gap junctions between cardiac muscle fibers (See Figure 10.4. The presence of these gap junctions not only allows, but greatly expedites, the excitation of one fiber to the next. This is fundamentally different from skeletal muscle where each muscle fiber is insulated from nearby muscle fibers by the endomysium. In skeletal muscle this fiber to fiber excitation would be deleterious to controlled, coordinated movement. But in cardiac muscle the fiber to fiber wave of excitation facilitates a synchronized contraction of all atrial, and then ventricular fibers. Such "all together" contraction facilitates the change in atrial and ventricular chamber size which facilitates the change in pressure necessary for circulation.

Electrocardiogram (ECG)

The wave of excitation through the cardiac conduction pathway and cardiac muscle fibers generates enough electrical potential to be measured by electrodes on the skin surface. There are several possible electrode placement arrangements, the most popular being referred to as a 12-lead ECG. The 12-lead ECG requires 10 electrodes. Four electrodes are on

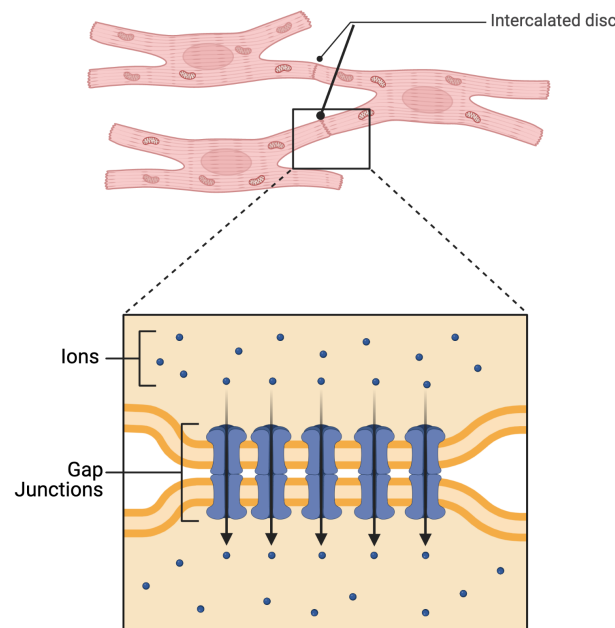


Fig. 10.4: Intercalated Discs with Gap Junctions (Created with BioRender.com)

(or near) the limbs (six limb leads), and six electrodes are on the thorax and surround the heart (six precordial leads). Lead II arises from electrodes on the right arm and left leg and tracks the wave of excitation in the frontal plane from the SA node (origin) through the left ventricle. The ECG wave from Lead II is the most commonly depicted ECG wave. Precordial ECG waveforms that are similar to Lead II include V4, V5 and V6. Lead II, V4, V5 and V6 are commonly utilized for routine ECG monitoring for arrhythmias.

A normal ECG is said to display sinus rhythm (sinus referring to the SA node). Figure 10.5 displays the sinus rhythm waves from a normal ECG from a single cardiac cycle from the view of Lead II. Readers are encouraged to view videos of ECG in real time as a preferred approach to becoming familiar with rhythm analysis. A video of normal sinus rhythm is available at this [link](#). Another option for practicing ECG rhythm analysis is available on iOS and Android devices is the [Simpl-Simulated Patient Monitor](#).

Table 10.1 refers to Figure 10.5 and provides further insight into the excitation event and clinical significance of each wave, segment and interval. ECG abnormalities that must be recognized during real time, single lead rhythm analysis are discussed later in the chapter under practice considerations.

ECG	Excitation Event	Significance
P Wave	Atrial Depolarization	SA node, atria
PR Segment	AV node delay	AV Node
PR Interval	SA & AV node	SA & AV node
QRS Complex	Ventricular Depolarization	Ventricles, RR Interval inversely proportional to HR
ST Segment	Ventricular Repolarization	Ischemia, Injury
QT Interval	Ventricular De + Repolarization	Long QT syndrome
T Wave	Ventricular Repolarization	Infarction

Table 10.1: ECG Waves, Segments, Intervals and their excitation events and significance

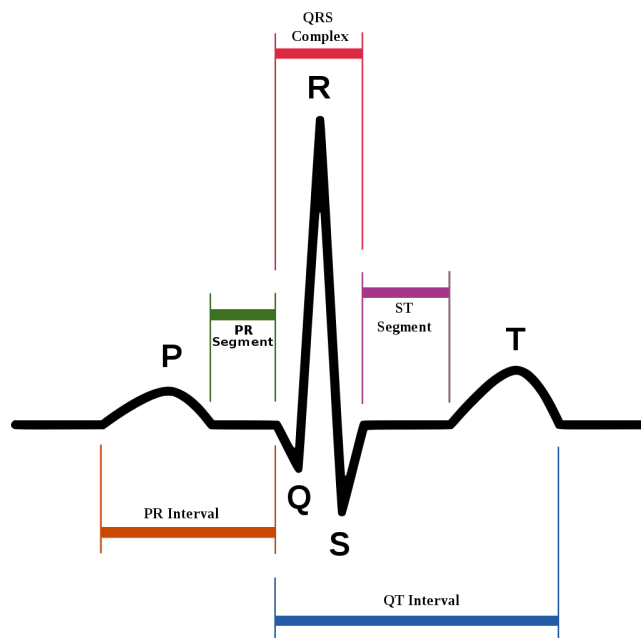


Fig. 10.5: ECG Waves (Lead II) (From Wikipedia [Sinus Rhythm Labels](#), CC BY-SA 4.0)

10.1.2.3 Contractility

An excitation results in full tetany of the myocardium and generates enough tension to generate enough pressure for circulation. For a single excitation to result in full contraction of the atria and then the ventricles requires enough Ca^{2+} to be released with one excitation for tetany (not just a twitch). There are two related features that allow for this to occur in cardiac muscle.

1. The cardiac muscle excitation (action potential) is prolonged (See Figure 10.6). It is prolonged due to specialized voltage gated Ca^{2+} channels that allow an inward movement of Ca^{2+} beyond the initiation of the excitation by Na^+ . The prolonged excitation occurs all the way down to the SR which continues to release Ca^{2+} .
2. Prolonged SR release of Ca^{2+} into the cell also contributes to the prolonged excitation of the membrane itself (a positive feedback system).

The primary take away message is that an excitation of a cardiac muscle fiber is prolonged due to inward movement of Ca^{2+} across the sarcolemma, which leads to prolonged SR release of Ca^{2+} for activation that also promotes prolonged excitation.

! Cardiac Ca^{2+} Pharmacological Targets

Based on the importance of Ca^{2+} in both prolonged excitation and activation of cardiac muscle fibers it is not surprising that medications influencing Ca^{2+} make up two major classes of cardiovascular medications. Ca^{2+} -channel blockers decrease the influx from the membrane (anti-arrhythmic) or the SR (anti-hypertensive); and digitalis, which increases the presence of Ca^{2+} in the cardiac muscle fiber.

10.1.3 Regulation

The regulation of cardiac muscle tension, as part of cardiac pump function (or just cardiac function) is involved in the regulation of $\dot{Q}$. Regulatory mechanisms can vary cardiac pump rate (HR) and vary is contractility (amount of tension and therefore pressure).

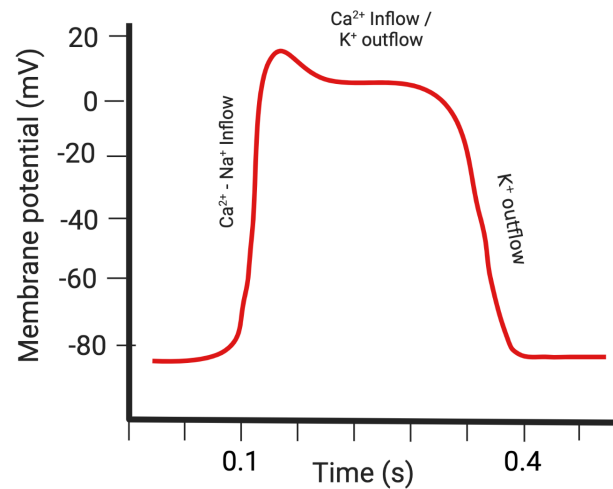


Fig. 10.6: Ventricular Excitation (Created with BioRender.com)

Equations 9.2 and 10.3 are useful models when connecting the regulation of cardiac muscle pump function to $\dot{Q}$. Combining these equations by substituting SV with its components from Equation 10.3 results in:

$$\dot{Q} = HR \times EDV \times LVEF \quad (10.4)$$

? If interested let me know

I have a note to myself to make a graphic on the regulation of cardiac function that integrates the past few pages but I don't have the time to do this for this draft. But if anyone does this while studying and is interested in payment for your artwork (if accepted of course), let me know. You'd also get credit for your work in the caption of the figure.

Based on Equation 10.4 there are two ways for cardiac pump function to regulate $\dot{Q}$. As previously discussed the regulation of $\dot{Q}$ results in changes to BP which is regularly monitored to provide feedback.

1. Regulate heart rate (HR) - referred to as chronotropy
2. Regulate tension (contractility) (LVEF) - referred to as inotropy

Equation 10.4 also includes the possibility of using venous return to regulate cardiac pump function. However, venous return is only minimally regulated. Venous return and its EDV are a consequence of venous circulation pressures which are largely determined by skeletal muscle activity. There are changes to vascular tone, but due to the small amount of muscle in the walls of the veins there is much less capacity to regulate venous tone than arterial tone.

10.1.3.1 Neuroendocrine Influence of the Cardiac Pump

Neural Input

The autonomic neuroendocrine system influences the heart via direct neural connections and through circulating hormones. It is important to keep in mind that the SA node is the pacemaker. The neuroendocrine input influences, but does not pace, the heart. The multitude of competing external influences to the pacing of HR is one of the reasons for the phenomenon known as HR variability (HRV).

The parasympathetic (cholinergic) system has direct neural input through the cardiac branch of the vagal nerve (Cranial Nerve X). Vagal nerve activity generally decelerates cardiac function, thus decreasing HR, speed of the conduction wave

of excitation and to a lesser extent contractility. The right vagus nerve stimulates primarily the SA node and affects rate, whereas the left vagus nerve stimulates primarily the AV node and affects AV conduction. Note that parasympathetic neural input does not influence EDV. However, EDV can influence parasympathetic neural input (see the Bainbridge reflex below).

Sympathetic (adrenergic) system neural input is through the thoracolumbar sympathetic system and increases HR and augments contractility, thus increasing cardiac function.

Endocrine Input

In response to physical activity or stress, the sympathetic nervous system releases catecholamines (epinephrine and norepinephrine) from the adrenal cortex of the adrenal glands which increases HR and contractility. The endocrine influence of the sympathetic nervous system provides a more efficient mechanism for fight/flight signaling since hormones, once circulating in the blood, continue to exert their influence.

10.1.3.2 Cardiac Reflexes

Cardiac reflexes influence HR and contractility and can be divided into four general categories: baroreflex (pressure) which was discussed in Chapter 9, Bainbridge reflex (stretch), chemoreflex (chemical reflex), ergoreflex (ergoreceptors).

Bainbridge Reflex

Mechanoreceptors located in the right atrial myocardium respond to stretch induced by venous return (EDV). An increased volume in the right atrium results in an increased stretch of the atrial wall. This reflex, known as the Bainbridge reflex, stimulates the vasomotor center of the medulla, which, in turn, increases sympathetic input, and decreases parasympathetic input for an overall increased HR and contractility. Respiratory sinus arrhythmia, an increased HR during inspiration and decreased HR during expiration, may be facilitated by changes in venous return and SV caused by changes in thoracic pressure induced by the respiratory cycle. At the beginning of inspiration when thoracic pressure is decreased, venous return and EDV is greater; therefore a greater stretch is exerted on the atrial wall.

Chemoreflexes

Chemoreceptors located on the carotid and aortic bodies have a primary effect on increasing rate and depth of ventilation in response to CO_2 levels, but they also have a cardiac effect. Increased CO_2 tends to promote a higher HR and decreased CO_2 a lower HR. However, compared to other reflexes and competing influences over HR the chemoreflexes exert a relatively minor influence.

Ergoreflexes

Ergoreceptors and the ergoreflex regulates circulation through activation of mechanosensitive afferents to inhibit the sustained vagal effects on the heart caused by an increase blood pressure triggering the baroreflex during physical activity that would otherwise work against the overall need for more circulation.

10.1.4 Energetics

Cardiac muscle is exceptionally well developed for aerobic metabolism, even more so than slow oxidative (SO) skeletal muscle fibers. Even during high intensity exercise cardiac muscle is capable of taking in lactate that is circulating in the blood and facilitating its transformation back into pyruvate and acetyl-CoA for entry into the citric acid cycle (TCA).

Only when there is a lack of O_2 available for electron transport (ETC) do the aerobic energetic pathways not keep up with glycolysis in cardiac muscle cells (there is no normal anaerobic threshold for cardiac muscle).

Cardiac muscle is so developed for aerobic metabolism that even at rest the myocardium utilizes approximately 80% of the O_2 delivered the blood. Skeletal muscle, at rest, only utilizes approximately 20% of the O_2 that is delivered in the blood. Though much of this difference is based on the difference in resting metabolic need. This makes cardiac muscle critically dependent on increases in circulation to provide the additional O_2 needed in response to increased myocardial oxygen demand (RPP).

! Rate Pressure Product

The rate pressure product (RPP), also referred to as the double product, is described by the equation: $RPP = HR \times SBP$, where HR is heart rate and SBP is systolic blood pressure. It is an indication of cardiac muscle (myocardium) oxygen demand. This experimental observation has useful clinical application. It also demonstrates the fact that pressure generated by the myocardium includes active tension which requires oxygen.

In situations with increased afterload the myocardium must develop more pressure for circulation. This greater pressure requires greater tension. This greater tension requires more ATP. More ATP requires more O_2 . The more frequently the tension is generated (HR) the more ATP and subsequently O_2 is required. The experimental association between RPP and myocardial O_2 demand is a rather simple consequence of the underlying relationship between pressure (resistance), active tension and ATP.

If a patient undergoes maximal exercise testing and has myocardial ischemia, RPP can be calculated at the point when ischemia is occurring to establish the patient's ischemic threshold. RPP at the ischemic threshold can then be used during exercise to provide a safe guideline of exercise intensity. This value is useful even if the patient is on a beta-adrenergic antagonist (-blocker). Even though the -blocker reduces the HR response by blocking the action of epinephrine and norepinephrine on the heart, it is this mechanism that reduces the risk of reaching ischemia. While on a -blocker, HR is not a good linear indication of overall aerobic workload (it loses its linear relationship with $\dot{V}O_2$), the RPP is a good discrete indicator of whether a patient is above or below their ischemic threshold. A patient that starts taking a -blocker (adrenergic antagonist) will have a harder time exceeding the ischemic threshold, which is the point of being on a -blocker for people with ischemic cardiac disease.

10.1.4.1 Cardiac Circulation

The major coronary arteries include the left coronary (LCA) which quickly branches into the left anterior descending (LAD) and left circumflex (LCx); and the right coronary artery RCA) (See Figure 10.7). LAD distributes the anterior wall of both ventricles (L & R) including the septum. LCx the left lateral and posterior walls. LAD and LCx join for distribution in the inferior myocardium. RCA distributes to the anterolateral, lateral, posterior and inferior RV. The RCA and LCA are the first two vessels off of the aorta. Blood is pumped to these large superficial coronary arteries during ventricular systole. At this time, myocardial contraction limits the flow of blood to the myocardium; therefore myocardial tissue is perfused during diastole.

10.2 Cardiac Cycle

The events of the cardiac cycle are displayed in Figure 10.8.

Circulation throughout the cardiac cycle depends on circulatory and cardiac pressure gradients. The right side of the heart is a low-pressure system with little vascular resistance in the pulmonary arteries, whereas the left side of the heart is a high-pressure system with high vascular resistance from the systemic circulation. The cardiac cycle is the period from the beginning of one contraction, starting with SA node depolarization, to the beginning of the next contraction. Systole is the period of contraction, and diastole is the period of relaxation. Systole and diastole can also be categorized into atrial and ventricular components.

- Atrial diastole is the period of atrial filling. The flow of blood is directed by the higher pressure in the venous circulatory system.

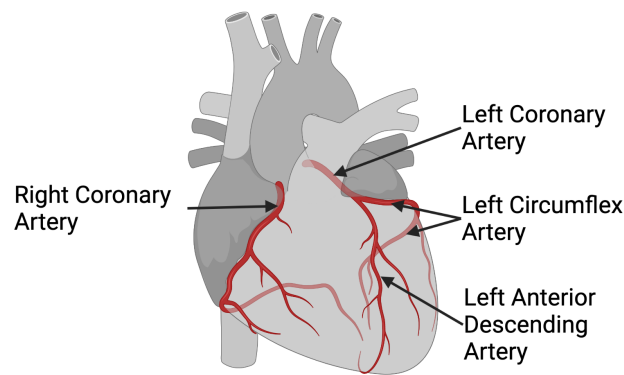
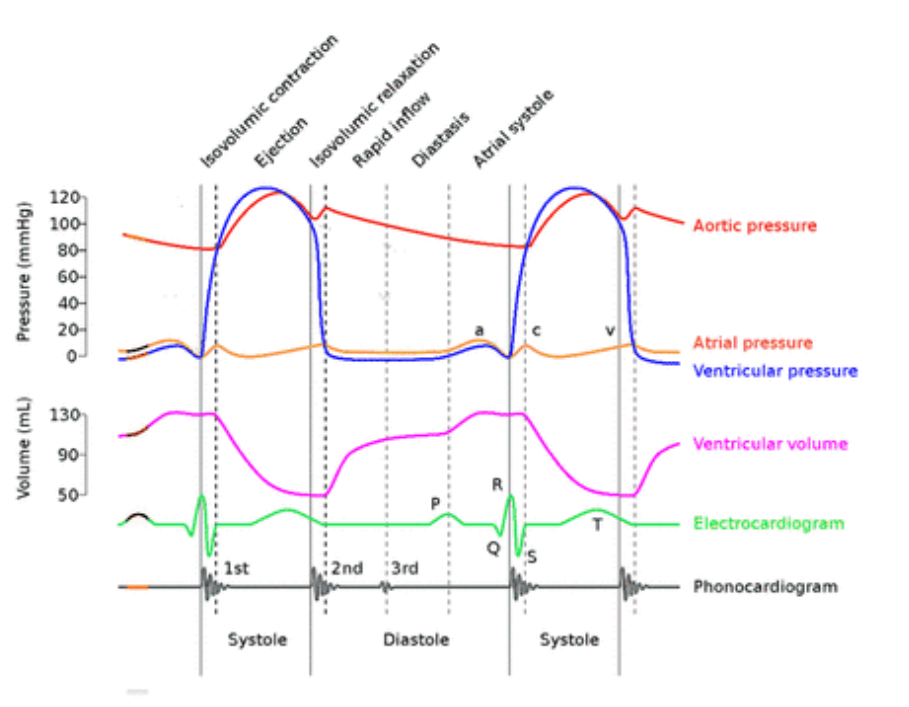


Fig. 10.7: Cardiac Circulation (Created with BioRender.com)

Fig. 10.8: Wiggers Diagram of the Cardiac Cycle (From Wikipedia [Wiggers Diagram Entry](#), CC BY-SA 4.0)

- Atrial systole is the period of atrial emptying and contraction. Initial emptying of approximately 70% of blood occurs as a result of the initial pressure gradient between the atria and the ventricles. Atrial contraction then follows, squeezing out the remaining 30%. This is commonly referred to as the atrial kick.
- Ventricular diastole is the period of ventricular filling. It initially occurs with ease; then, as the ventricle is filled, atrial contraction is necessary to squeeze the remaining blood volume into the ventricle. The amount of stretch placed on the ventricular walls during diastole, referred to as left ventricular end diastolic pressure (LVEDP), influences the force of contraction during systole.
- Ventricular systole is the period of ventricular contraction. The initial contraction is isovolumic (i.e., it does not eject blood), which generates the pressure necessary to serve as the catalyst for rapid ejection of ventricular blood. The left ventricular ejection fraction (EF) represents the percent of end diastolic volume ejected during systole and is normally approximately 60%.

Understanding Wiggers Diagram

Understanding the cyclic events of Wiggers Diagram in Figure 10.8 requires an understanding of pressure changes due to volume and volume changes due to pressure. Readers should start from the left and track the changes in each wave, using the ECG as a guide to the excitation events. Once the reader understands each wave they should proceed to observing the coherence between two, then three and finally all of the waves through the two cardiac cycles depicted. An understanding of cardiac function and the cardiac cycle can be confirmed when Figure 10.8 can be reproduced based on an understanding of the events (not by memorizing how to reproduce it). Though, some individuals find that learning to reproduce it by memorizing it ultimately leads to an understanding of cardiac function.

10.3 ECG Rhythm Analysis

ECG rhythm analysis allows identification of minor through major cardiac arrhythmias (conduction abnormalities) as well as deviations in wave forms that are associated with ischemia, injury or infarction of the myocardium (myocardial circulation abnormalities).

Rhythm analysis includes the observation of one ECG lead in real time via a hard wired ECG monitor or a telemetry (wireless) ECG monitor. The first question when observing an ECG is whether it appears normal, that is with all of its waves and segments in order and spaced as expected. A general approach when observing the ECG in real time is to ask:

- Is every P wave followed by a QRS?
- Is every QRS preceded by a P wave?
- Is every QRS followed by a T wave?
- Is every T wave preceded by a QRS?
- Is every T wave followed by a P wave?
- Is every P wave preceded by a T wave?
- Is the PR segment normal width?
- Is the QRS complex normal width?
- Does the ST segment quickly return to baseline?
- Are the QRS complexes reasonably and regularly spaced (RR Intervals)?

10.3.0.1 Arrhythmias

Examination for arrhythmias can occur with any ECG lead (any one of the standard 12, or other less commonly used electrode and lead configurations). Every lead provides the same information about conduction. For example, if a P wave is missing or a PR interval is long in Lead II, it will be equally missing or as long in Lead aVL or V3 (or any of the other leads).

Tachyarrhythmias (Ectopy)

Tachyarrhythmias refer to arrhythmias associated with more excitation than usual during cardiac conduction. Such aberrant excitation usually results in a higher heart rate. Tachycardia is a HR higher than 100 bpm and is not technically an arrhythmia - though if it occurs without provocation (physical, mental or emotional stress) then it may be indicative cardiac or autonomic dysfunction. When tachycardia is unprovoked it is often referred to as supra-ventricular tachycardia (SVT) or atrial tachycardia. The most common type of SVT is paroxysmal SVT (PSVT). Paroxysmal simply means it comes and goes without any currently understood provocation (cause). There are a group of tachyarrhythmias caused by ectopic action potentials that cause ectopic excitations and beats. Ectopic simply means these excitations are not originating in the SA node and following the normal excitation conduction pathway. Ectopy originates from an ectopic focus (myocardial cells causing an action potential) or if more than one focus, as ectopic foci. An ectopic focus (or foci) can occur due to myocardial cell irritation (producing an excitation when it should not, i.e. caffeine or other stimulants), or due to myocardial cell depression (delay or block in conduction caused by ischemia, injury or scar tissue) that produces a recurrent loop of self perpetuating excitation within the myocardium).

Table 10.2 organizes six different tachyarrhythmias based on where they occur (atria or ventricles) and the regularity of the ectopic focus (or foci) excitations.

	Atria	Ventricle
Occasional Ectopic Focus	Premature Atrial Contr. (PACs)	Premature Ventricular Contr. (PVCs) [†]
Occasional Ectopic Foci	PACs	Multifocal PVCs
Non Stop Ectopic Focus	Atrial Flutter (AFlutter)	Ventricular Tachycardia (VTach)
Non Stop Ectopic Foci	Atrial Fibrillation (AFib)	Ventricular Fibrillation (VFib)

Table 10.2: Types of Ectopic Tachyarrhythmia

The severity of all conduction abnormalities is based on the impact they have on cardiac output and blood pressure (hemodynamics). In general the impact a tachyarrhythmia has on hemodynamics is greater proceeding down the rows; and is always greater in the ventricular column than the atrial column. For example, PACs have little, if any, impact on hemodynamics; whereas VFib does not produce any circulation at all so there is no cardiac output and no blood pressure (and therefore no pulse).

Since the tachyarrhythmias exist on a spectrum, it is required that a physical therapist can identify which tachyarrhythmia is present should one arise (or be part of a patient's chronic condition). For example, it is not appropriate to stop activity due to AFib in a patient with chronic AFib, but it is important to monitor the ventricular rate and BP for signs of worsening hemodynamics in AFib. Nor is it required to stop activity due to PACs or PVCs in all situations.

A YouTube playlist of tachyarrhythmias is available at this [link](#).

Bradyarrhythmias & Blocks

Bradycardia refers to any HR below 50 bpm. However, like tachycardia, bradycardia does not always indicate there is a conduction problem. Whether bradycardia is a problem comes down to whether it is creating problems with hemodynamics. Generally, if there is an abnormality causing bradycardia it will provoke problematic changes in hemodynamics, for example, sick sinus syndrome or third degree AV block. A normal cause of bradycardia that does not provoke problematic changes in hemodynamics is being well conditioning from exercise training.

Bradyarrhythmias can be caused by AV blocks, though not all instances of AV blocks create bradycardia. It is just common to classify blocks as a bradyarrhythmia. There are three degrees of AV Block, and four types (there are two types of second degree AV block).

- First Degree AV Block: Prolonged PR Interval
- Second Degree AV Block - Mobitz 1 (Wenkebach) - progressive prolongation of the PR interval until finally there is a P wave without a QRS indicating the conduction was blocked which is then followed by a normal PR interval that begins progressive prolongation (this pattern repeats)
- Second Degree AV Block - Mobitz 2 - normal PR interval with an occasional P wave with no QRS
- Third Degree AV Block - complete AV block (also referred to as complete heart block) P waves and aberrant looking QRS complexes are completely dissociated (though may occasionally align so identification requires watching the ECG for several cycles)

A YouTube playlist of bradyarrhythmias and AV blocks is available at this [link](#).

The severity of AV blocks is related to the impact they have hemodynamics. As can be expected, the severity increases from First Degree to Third Degree. Mobitz 2 is considered more severe to Mobitz 1, not due to hemodynamic differences, but because it is more likely to deteriorate into Third Degree AV block.

10.3.1 Myocardial Circulation Abnormalities

Abnormalities in myocardial circulation (perfusion) such as ischemia lead to hypoxic conditions in the myocardial cells. If the ischemia persists cell death will occur. Prior to cell death an early indication of ischemia includes changes to the

ST segment of the ECG. In a very general way, ST changes are indicative of ischemia and/or injury. This more general fact is sometimes made more specific, with ST depression taken as a sign of myocardial ischemia and ST elevation taken as a sign of myocardial injury. Myocardium that is injured may be reversibly injured (will recover) or irreversibly injured (will die, infarction).

A YouTube playlist of ST segment changes is available at this [link](#).

One of the problems with the more specific claims of ST changes is that it assumes the ECG lead being observed is a lead that has the best view of the myocardium that has ischemia. Unlike the the arrhythmias which are seen equally well in all ECG leads, a circulation abnormality resulting in ischemia is local to the region of the myocardium with the circulation abnormality (rarely the entire myocardium). For example, aVR tends to have a reversed polarity to Lead II (what goes up in Lead II, goes down in aVR; what goes down in Lead II, goes up in aVR). If ischemia is in an area that Lead II shows depression, it may appear as elevation in aVR. During the diagnostic process (a cardiac stress test) a 12 lead ECG is utilized and additional testing strategies are utilized to confirm that what was observed was in fact ischemia. A cardiologist can get a good idea of where the ischemia is occurring from the 12-lead ECG, which is then confirmed with imaging (nuclear scans or catheterization). When it is known which location (and therefore lead) best pin points the area with a propensity for ischemia then that lead (or a couple of leads) can be utilized for rhythm monitoring. In such situations it is mostly true that ischemia appears as ST depression, and injury appears as ST elevation.

Since the physical therapist using rhythm analysis is primarily interested in identifying abnormalities that warrant the cessation of activity and alerting members of the medical team of such abnormalities (not in making a diagnosis), simply identifying ST changes is sufficient. Both ischemia and injury are sufficient reason to stop activity and alert members of the medical team. In fact, the risk associated with myocardial ischemia and injury is sufficient to warrant extra caution and stopping activity if ST changes are suspected while observing the ECG rhythm. In other words, if the ST segment appears above or below the baseline its safer to assume it is than to assume it isn't. Cease activity and continue monitoring. As with any such signs, noting other signs and symptoms is warranted (shortness of breath, diaphoresis, chest pain, blood pressure, respiratory rate).

10.4 Summary

The cardiac pump supports circulation. Cardiac muscle includes several characteristics required for its unique role in circulation. Cardiac muscle tension, excitation and regulation support circulation and pressure regulation. Excitation of cardiac muscle can be recorded with the ECG and offers insights into several abnormalities of conduction as well as myocardial ischemia. Monitoring the ECG provide physical therapists with non-invasive insights, that when combined with an understanding of circulatory physiology, provide useful information when working with clients both with and without conditions that directly affect the circulation.

10.4.1 Next Step

With an understanding of mass balance and circulation we can now proceed to the mass balance of oxygen and carbon dioxide in the next two chapters. First we consider Respiration which focuses on gas exchange between the alveoli and blood. Then we turn to Ventilation which focuses on air exchange between the environment and the alveoli (breathing).

Sample Questions

1. Normal end diastolic volume (EDV) is directly related to normal end diastolic pressure (EDP) and normal ventricular compliance. If a pathology results in reduced ventricular compliance during diastole what would occur?
2. What myocardium features lead to the normal pattern of myocardial conductivity?
3. Why is cardiac muscle excitation (action potential) prolonged and what changes to the action potential allow this to occur?
4. How are options for determining and monitoring ischemic threshold during exercise?

5. Which cardiac arrhythmia results in an irregularly irregular pulse rhythm?

Chapter 11

Respiration

Updated on May 22, 2024

Contents

11.1	Respiration Overview	174
11.2	Cellular & Internal Respiration	174
11.3	Blood Transport	179
11.4	Pulmonary Circulation	184
11.5	External (Alveolar) Respiration	186
11.6	Practice Connections	188
11.7	Summary	190

This chapter covers support of the extracellular fluid and cellular function by having the correct blood gases through respiration. Respiration is a microcirculation process that requires concentration and pressure gradients as well as blood transport and circulation. Cellular and internal respiration occurs between the capillaries and muscle fibers. External (alveolar) respiration requires pulmonary circulation to the alveoli. Just as internal respiration includes a micro-circulatory exchange with muscle fibers, external respiration includes a micro-circulatory exchange with the the alveoli (air sacs) in the lung. Alveolar ventilation (covered in the Chapter 12) ensures that the alveoli maintain the required concentration (pressure) of both CO_2 and O_2 to support external (alveolar) respiration.

Objectives include:

1. Explain cellular, internal and external respiration.
2. Explain and compute the partial pressure of O_2 given the fraction of O_2 (FO_2) and the atmospheric pressure.
3. Explain partial pressure driven diffusion.
4. Explain the factors that influence diffusion.
5. Explain the diffusion of O_2 uptake and CO_2 release by the tissues (cells) including the RER.
6. Explain the life cycle of red blood cells, anemia and ABO blood types.
7. Explain blood transport of CO_2 and O_2 including the oxyhemoglobin curve.
8. Explain pulmonary circulation.
9. Explain external (alveolar) respiration.
10. Explain the impact of P_aCO_2 on brain blood flow.
11. Explain arterial blood gases and interpret acid - base balance and disorders.
12. Explain DLCO and pulsed oximetry in the examination of respiration.

11.1 Respiration Overview

Muscle function requires energy in the form of ATP. Sustained regeneration of ATP in the mitochondria continuously produces CO_2 that must be removed from the muscle fiber; and requires delivery of O_2 to the muscle fiber (**cellular respiration**). The microcirculation removes and provides what the muscle fibers need, including the removal of CO_2 and delivery of O_2 from the capillary blood (**internal respiration**). Internal respiration of the blood gases from and into muscle fibers requires a concentration gradient that requires a continuous circulation of blood to the muscle capillary network (**circulation**). Blood provides a transport mechanism for both CO_2 and O_2 (**blood transport**). The circulation of blood from the muscles through veins to the right side of the heart delivers all blood (**venous return**) to the pulmonary circulation. The pulmonary circulation delivers venous blood through pulmonary arteries to the interstitial space surrounding alveoli in the lungs (**pulmonary circulation**). Microcirculation (primarily gas diffusion) results in the equilibration of blood gases with the alveoli gases (**external respiration**). Pulmonary circulation then returns arterial blood through pulmonary veins to the left side of the heart for systemic circulation.

11.2 Cellular & Internal Respiration

Cellular respiration occurs inside the mitochondria with aerobic metabolism. It is aerobic metabolism that utilizes O_2 and produces CO_2 in sufficient quantities to require continuous exchange with the sarcoplasm, interstitial space, circulating blood and ultimately the environment (gas exchange). Cellular respiration establishes the need for internal respiration, the gas exchange of O_2 and CO_2 between the intracellular and extracellular (both interstitial and vascular compartments).

11.2.1 Partial Pressure Gradients & Diffusion

The fundamental concept for respiration is that the concentration of O_2 and CO_2 dissolved in the fluid the ECF (plasma) creates partial pressure gradients of O_2 and CO_2 to create pressure driven flow. Pressure driven flow for gases is called diffusion. There is a relationship between concentration and partial pressure. The higher the concentration of a particular molecule in a gas (as a fraction of all the molecules), the higher the partial pressure.¹ Specifically, the partial pressure of a particular molecule in a gas is proportional to the quantity of particular molecules as a fraction of the total molecules of the gas multiplied by the total pressure of the gas. The partial pressure of O_2 being:

$$PO_2 = FO_2 \times TotalPressure(mmHg) \quad (11.1)$$

Where FO_2 is the fraction of O_2 in the gas resulting in total pressure ($mmHg$). The fraction of air in the environment (atmosphere, atm) that is O_2 is 0.21 (21%). Air in the environment is inspired so it is called the fraction of inspired O_2 (F_iO_2). At sea level the total air pressure is 760 $mmHg$ (barometric pressure). Therefore, the $P_{atm}O_2 = 0.21 \times 760mmHg$, so $P_{atm}O_2 = 159.5mmHg$.

11.2.2 Partial Pressure Driven Diffusion

Figure 11.1 is copied from Chapter 7. The gradients for O_2 and CO_2 are partial pressures (PO_2 and PCO_2). The partial pressure of each compartment is further noted with a subscript, with partial pressure of O_2 in the cell being P_cO_2 and in the tissue as P_tO_2 . The gradient for PO_2 is from the arterial end of the capillary (partial pressure of arterial blood, $P_aO_2 \approx 100mmHg$) into the tissue ($P_tO_2 \approx 40mmHg$), and into the cell ($P_cO_2 \approx 20mmHg$). Of these values the P_cO_2 is most variable and depends on the rate of utilization of O_2 in the mitochondria (aerobic metabolism). The partial pressure driven diffusion of cellular respiration relies on the continuous use of O_2 in the mitochondria; and the continuous use of O_2 in the mitochondria depends on the availability of O_2 diffusing in from the arterial blood in the

¹ Note: if there is a higher concentration but all gases in the mixture have also increased their concentration so that the fraction of any particular gas is the same, the partial pressure is higher because the total pressure is higher.

capillary. If aerobic metabolism in a cell stopped then the P_cO_2 would quickly equilibrate with the P_aO_2 at 100 mmHg. If blood flow stopped (ischemia), or if there was a lower partial pressure of O_2 in the arterial blood ($P_aO_2 \leq 60\text{mmHg}$) (hypoxemia) then diffusion of O_2 would be slower and the rate of aerobic metabolism would be limited (hypoxia). The difference between arterial blood P_aO_2 and the venous blood P_vO_2 is due to diffusion of O_2 out of capillary blood into the interstitial fluid and cells. Internal respiration creates venous blood (blood with less O_2 and more CO_2 than arterial blood).

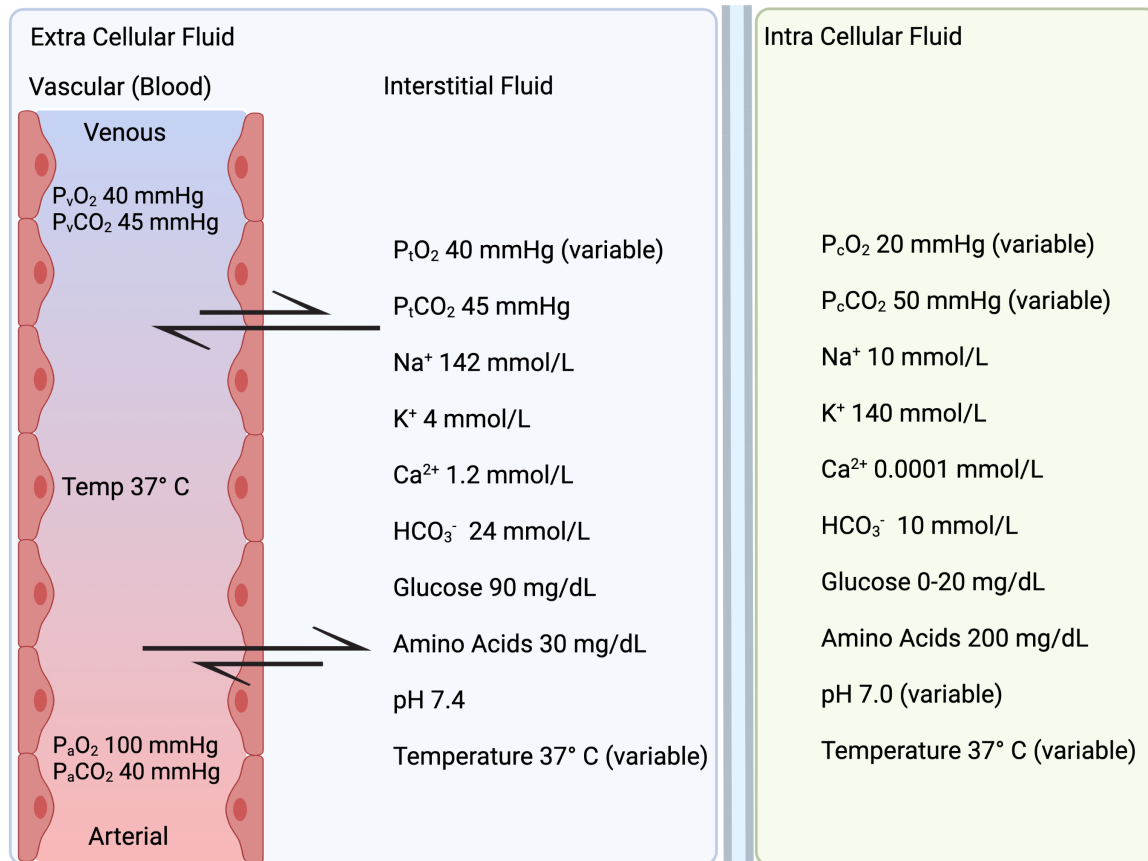


Fig. 11.1: Extra-cellular (Vascular & Interstitial) and Intra-cellular Fluid (Created with BioRender.com)

11.2.3 Factors Influencing Diffusion

The factors influencing diffusion are provided in the following equation:

$$\dot{D} = \frac{A}{T} \times d \times (P_1 - P_2) \quad (11.2)$$

- $\dot{D}$ is diffusion as a rate (amount of substances over time)
- A is the surface area that diffusion can occur
- T is the thickness of the membrane through which diffusion can occur

- d is a diffusion coefficient that is related to the solubility and molecular weight of the substance diffusing
- $P_1 - P_2$ is the partial pressure gradient (ΔP)

Equation 11.2 assumes that the membrane is permeable to the substance being diffused. Which means there are no membrane restrictions to diffusion other than the surface area or thickness in the equation. Since that is the situation for the diffusion of O_2 and CO_2 across the cell membrane, capillary membrane and alveolar-capillary membrane this assumption is appropriate for respiration.²

Only for those interested in what happened to Ohm's law

Equation 11.2 appears to jump away from the convention of Ohm's law. But it's in there. Resistance for diffusion is:

$$R_D = \frac{T \times MW}{A \times S} \quad (11.3)$$

With T representing the thickness of the membrane; MW the molecular weight; A the surface area; and S the solubility coefficient for the molecule. In Equation 11.2 the diffusion coefficient includes the solubility and the molecular weight. When you consider them separately, along with thickness and surface area, there's a parameter for resistance.

So that:

$$\dot{D} = \frac{\Delta P}{R_D} \quad (11.4)$$

So, no worries - Ohm's law is in there underlying pressure driven flow.

11.2.3.1 Diffusion of O_2 and CO_2

Since a nearly equal amount of O_2 and CO_2 is diffused across the cell, capillary and alveolar-capillary membrane during the microcirculation of blood the diffusion rates are, for all intents and purposes, equal under physiological conditions. They are always crossing the same membranes and therefore have the same A and T . As can be seen in Figure 11.1 the partial pressure gradients ($P_1 - P_2$) for O_2 and CO_2 are not equal. This is because the diffusion coefficient (d) is different between O_2 and CO_2 . CO_2 has a lower partial pressure gradient because it has a higher d due to a higher solubility. The implication of this the higher d for CO_2 is nearly equal diffusion capability despite much lower pressure gradients.³

11.2.4 Tissue & Cellular O_2 Uptake

The diffusion of O_2 into the tissue (interstitial space) reduces the partial pressure of O_2 in the blood (See Figure 11.2). The PO_2 at the venous end of the capillary is the P_vO_2 and reflects how much O_2 has been taken in by the cells. An important concept is that the P_vO_2 from different areas of the body can be quite different. The regional P_vO_2 depends on the regional metabolism. While the P_aO_2 is the same throughout the arterial circulation, the P_vO_2 is different throughout the venous circulation, but continues to mix and become similar until the final mixing chambers of the right atrium and ventricles are reached. In other words, the P_vO_2 in the right ventricle represents a weighted mean from the entire body. And the P_vO_2 from the right ventricle is what gets sent to the lungs, which means all of the alveoli receiving blood get blood with the same PO_2 . Under resting conditions approximately 5 mL O_2 /100mL of blood is taken in by tissues [HH20]. However, this is an average and the true value is variable across the body with tissues such as the heart, brain and breathing skeletal muscles being slightly higher at rest, and inactive skeletal muscle lower when at rest (per unit volume).

² Note, if the membrane was only semi-permeable to O_2 and CO_2 then to model the situation more accurately the equation would need a permeability coefficient, as we have seen for filtration across the capillary.

³ This relationship between solubility, partial pressure and diffusion is more complicated than this since not only does solubility influence the diffusion coefficient and the rate of diffusion for a given partial pressure gradient, but solubility is inversely proportional to the partial pressure of the substance that is soluble based on Henry's law. These factors are beyond the depth of this text [CB17].

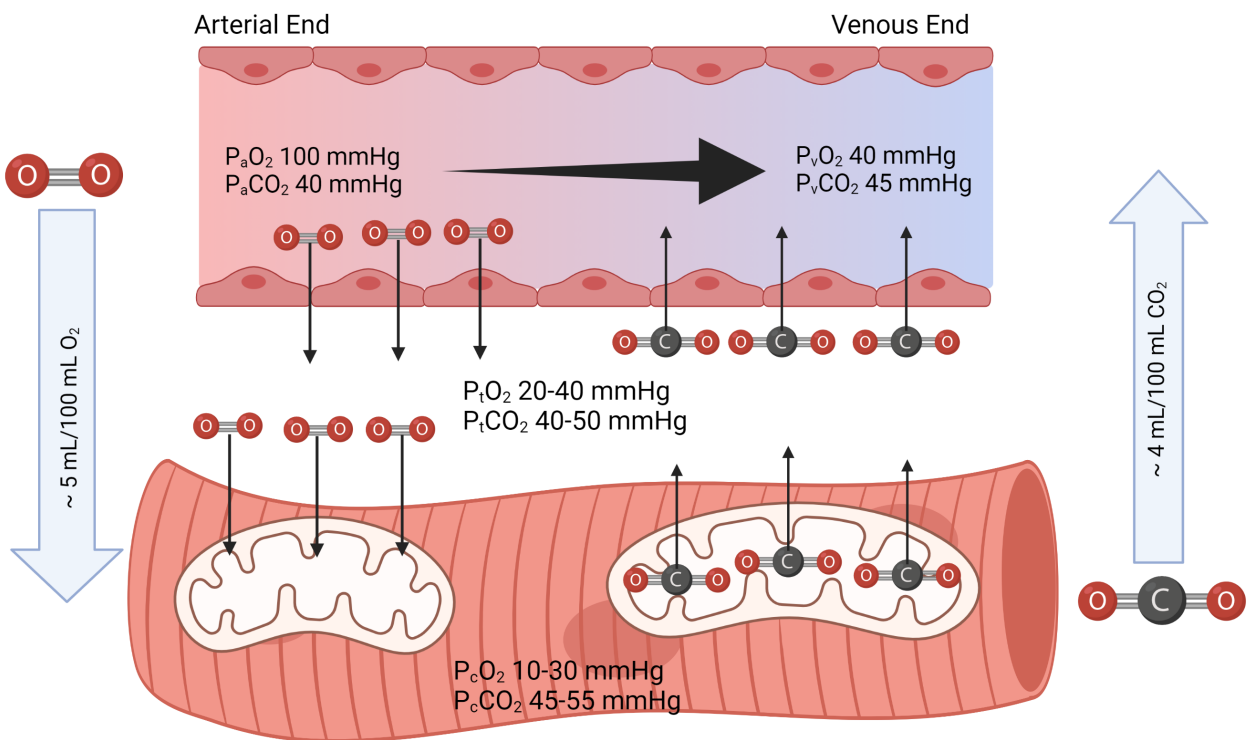


Fig. 11.2: Tissue Diffusion of O_2 & CO_2 . The ratio of $\frac{CO_2}{O_2}$ (respiratory exchange ratio, RER) in this figure is 0.8, which is based on mixed substrate utilization in the energetic pathways. Utilization of pure glucose would result in a ratio of 1.0 and the CO_2 leaving the cell would equal the O_2 entering the cell. (Created with BioRender.com)

11.2.4.1 Effect of Cellular Respiration

There is a linear relationship between tissue metabolism, cellular respiration and O_2 utilization. Since the rate of the energetic pathways is related to the ATP:ADP ratio (normally 5:1), alterations to the amount of ADP in a cell tend to result in proportional changes to the rate of O_2 utilization. For example, if cellular ADP is $\frac{1}{2}$ of its normal value then O_2 utilization will be $\frac{1}{2}$ its normal value (ATP:ADP ratio might be 11:1). If cellular ADP is $1.5\times$ higher than normal, then O_2 utilization is $1.5\times$ higher than normal (ATP:ADP ratio might be 3:1) [HH20]. The take home message is that the rate of O_2 utilization in the energetic pathways determines the rate of uptake of O_2 by the tissues and cells, and the venous O_2 .

11.2.4.2 Myoglobin

Myoglobin was introduced in Chapter 6 on Energetics and highlighted as a fundamental difference between SO, FOG and FG muscle fibers (Table 6.3). SO muscle fibers express (synthesize and contain) more myoglobin. Myoglobin is a protein in the same family as hemoglobin. It is capable of binding O_2 molecules, which removes them from solution and thus lowers the P_cO_2 which encourages more diffusion by maintaining the partial pressure gradient. The binding affinity of myoglobin for O_2 is based on the P_cO_2 with the myoglobin- O_2 dissociation curve (See Figure 11.3).

As O_2 utilization increases in a muscle fiber the P_cO_2 will drop which encourages dissociation (release) of O_2 from myoglobin and discourages O_2 entering the muscle from binding to myoglobin. With increased temperature ($40^\circ C$) the curve shifts right which encourages additional dissociation (release) of O_2 from myoglobin. Not depicted in Figure 11.3 is an additional rightward shift of the curve if pH drops below the normal resting muscle fiber value of 7.0. The rightward shifts of the myoglobin- O_2 dissociation curve in a muscle fiber encourage the release of O_2 . These rightward shifts (increased temperature and decreased pH) are associated with increased metabolism, which is precisely when the

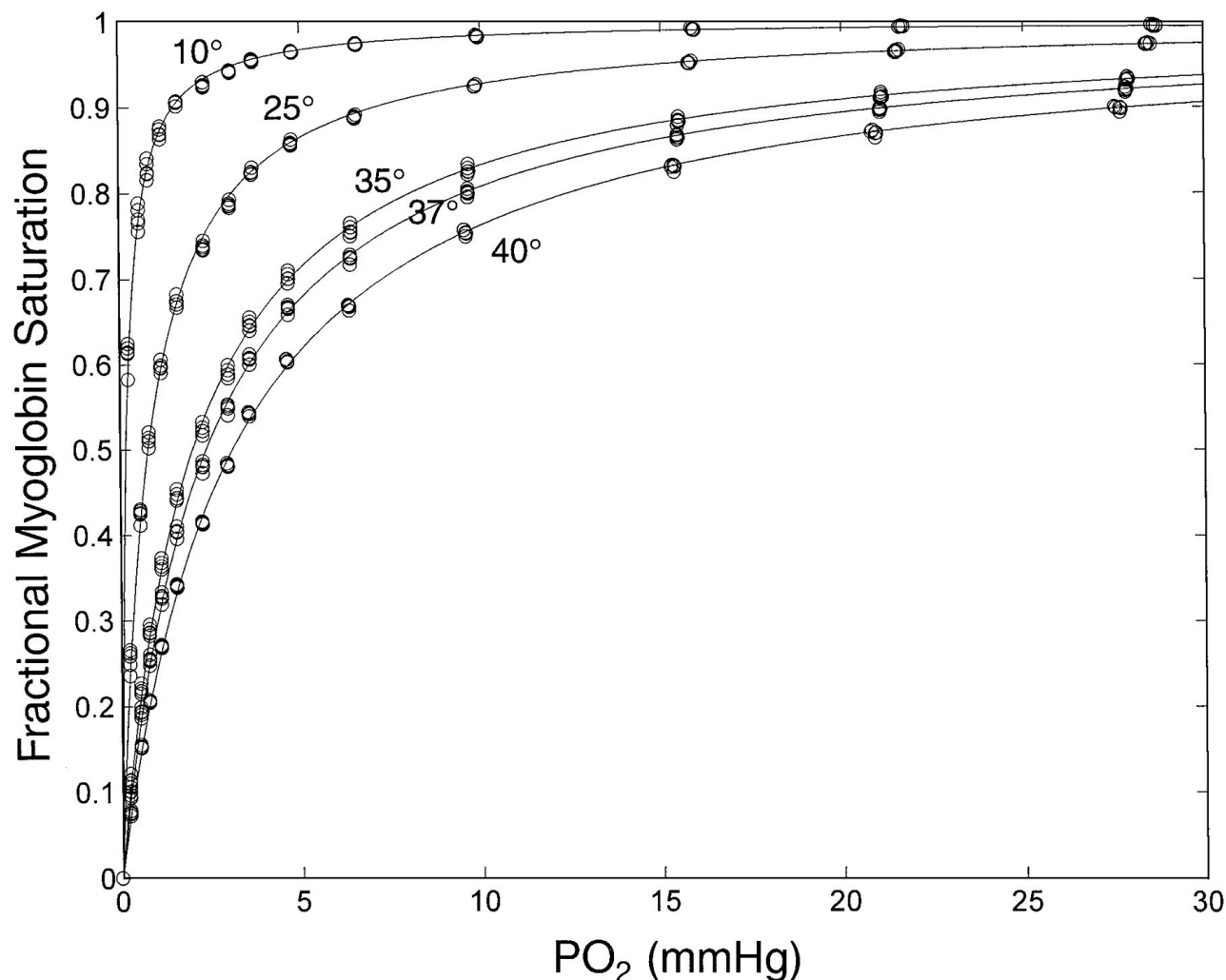


Fig. 11.3: Myoglobin- O_2 dissociation curve as a function of temperature at an intracellular pH of 7.0 (From [Sch+97])

fiber needs myoglobin to release its storage of O_2 for diffusion into the mitochondria. At rest (normal temperature and pH) the leftward shift encourages more myoglobin to bind O_2 . Even lower than normal body temperature (normal body temperature is $37^\circ C$), encourages additional storage. It's important to note that regardless of these shifts, with a low enough $P_c O_2$, myoglobin releases O_2 , and with a high enough $P_c O_2$, myoglobin binds and therefore stores O_2 .

11.2.5 Tissue CO_2 Release

Partial pressure driven diffusion of CO_2 follows the same reasoning as with the diffusion of O_2 but in the opposite direction (See Figure 11.2). The cellular partial pressure of CO_2 ($P_c CO_2$) is the highest partial pressure which drives diffusion of CO_2 out of the cell into the interstitial fluid. The $P_t CO_2$ is higher than the $P_a CO_2$ in the arterial blood at the arterial end of the capillary, which drives CO_2 into the capillary. The $P_c CO_2$ is variable based on the rate of CO_2 produced during aerobic metabolism (TCA). The diffusion of CO_2 into the blood increases the PCO_2 in blood flowing through the capillary. At the venous end the $P_v CO_2$ equilibrates to whatever is in the tissue which is higher than the $P_a CO_2$. There tends to be wider fluctuation in $P_a CO_2$ than in $P_a O_2$ due to the much lower atmospheric partial pressure of CO_2 . For reasons that will be partially elucidated in this chapter and more fully explained in Chapter 12 on Ventilation,

variation in ventilation (breathing rate and volume) can create large fluctuations in P_aCO_2 with relatively little change in P_aO_2 . Under resting conditions approximately $4\text{ mL } CO_2/100\text{ mL}$ of blood is released from the tissues.

11.2.6 Respiratory Exchange Ratio (RER)

The RER, introduced in Chapter 6 on Energetics, is based on the volume of CO_2 produced and released into the blood (and eventually expired by breathing); and the volume of O_2 taken in from the blood and utilized in the mitochondria.

$$RER = \frac{\dot{V}CO_2}{\dot{V}O_2} \quad (11.5)$$

With mixed substrate utilization for energy (fats and carbohydrates) the RER is approximately 0.8, which is based at a cellular level on the release of approximately $4\text{ mL } CO_2/100\text{ mL}$ and the uptake of approximately $5\text{ mL } O_2/100\text{ mL}$. With a purely carbohydrate metabolism the $RER = 1.0$ and with purely fats it is ≈ 0.7 .

11.3 Blood Transport

Whether discussing venous blood (from tissues and cells to the lungs), or arterial blood (from the lungs to the tissues and cells), O_2 and CO_2 are being transported in the blood. The transport of O_2 and CO_2 in the blood has overlap (shared mechanisms) albeit with different proportions per mechanism. Overall, CO_2 has an additional transport option that is responsible for a large proportion of CO_2 transport.

A small amount of O_2 (approximately 1-2%) and CO_2 (approximately 7%) is transported simply as O_2 and CO_2 dissolved in the blood. This portion contributes to the partial pressures and therefore has a large role in diffusion and as a factor in determining the amount being carried in the blood. But it is clearly not the largest quantity of either O_2 or CO_2 being transported.

11.3.1 Red Blood Cells

Red blood cells (RBCs), also referred to as erythrocytes, constitute approximately 45% of the total blood volume (Hematocrit, HcT). RBCs do not have a nucleus and therefore do not divide. They are regenerated and replaced regularly in the bone marrow through the process called erythropoiesis. Erythropoiesis is stimulated by the hormone erythropoietin (EPO) which is released from the kidneys.

EPO is secreted by the kidneys at a regular rate, possibly in response to bile filtration, to compensate for normal cellular turnover (See Figure 11.4). EPO release increases in response to cellular hypoxia and stimulates additional erythropoiesis as a compensatory mechanism in an attempt to solve the problem of cellular hypoxia. Common causes of cellular hypoxia resulting in elevated levels of EPO include altitude or low barometric pressure (hypobaric) exposure which reduce environmental and therefore arterial PO_2 ; anemia (reduced RBCs), and hypoxemia due to chronic lung disease. RBCs contain hemoglobin (Hgb).

Hgb is an iron-containing oxygen-transport protein. While HcT is the percent of blood that is composed of RBCs, Hgb is the number of grams per liter of Hgb in blood (usually per deciliter or 100 milliliter⁴). A normal Hgb level is approximately 15 g/dL with a range 12 to 20 g/dL .

Hgb has an O_2 -binding capacity of $1.34\text{ mL } O_2/\text{g}$ which increases the total blood oxygen capacity seventy-fold compared to dissolved oxygen in blood. A Hgb molecule can carry up to four oxygen molecules. When fully saturated with O_2 and with a Hgb of 15 g/dL , there is approximately $(15 \times 1.34 = 20.1\text{ mL } O_2/100\text{ mL})$ of blood. This estimate varies predictably and proportional to the amount of actual saturation and Hgb.

! Anemia

⁴ 1 dL is equal to 100 mL

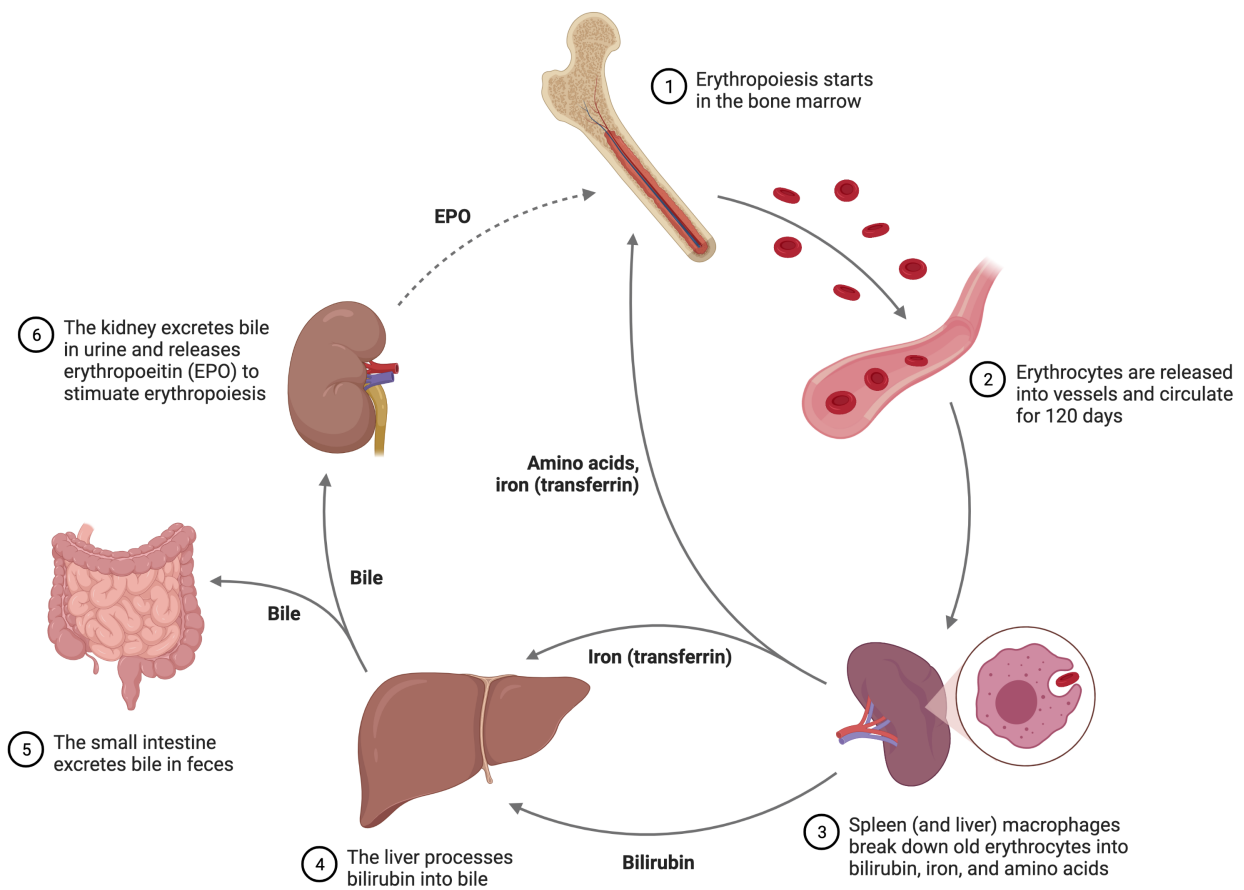


Fig. 11.4: Life Cycle of Erythrocytes (Created with RioRender)

Anemia with HcT in the range of 0.3 - 0.4 is common following surgery and a cause of fatigue and reduced endurance that will recover in several days when kidneys are healthy due to an increase in EPO release and associated erythropoiesis (or sooner if facilitated by a blood, or RBC, transfusion). Causes of anemia that are not reversible, and that continue until anemia progresses until HcT is less than 0.1, can result in death. Anemia is a, but not the only, cause of hypoxemia. With anemia, there is reduced oxygen carrying capacity in the blood. Chronic renal conditions, including chronic conditions such as heart failure that cause chronic renal insufficiency, will often result in anemia. When this occurs treatment can include the medication Epoetin alfa (Epogen, Procrit) which is a human made form of erythropoietin that stimulates RBC production in the bone marrow.

11.3.1.1 Transfusions & ABO Blood Types

RBCs can have A or B antigens on their cell membrane. Blood can occur with four possible types, A, B, AB or O (O being an indication that the RBC membrane has neither A or B antigens). Whatever antigen a RBC membrane has, it's plasma contains anti-bodies for the opposite antigen. If A antigen is present on the RBC then the plasma has Anti-B antibodies. If A and B antigen is present (AB blood) then the plasma has neither antibody. If no antigen is present (O blood) then the plasma has both antibodies. The type of blood influences whether the blood can be donated or received as summarized in Figure 11.5.

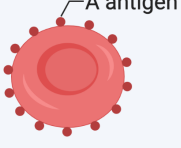
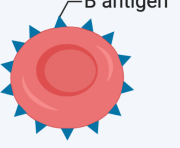
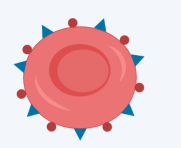
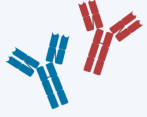
	A	B	AB	O
Antigens present on erythrocytes				
Antibodies present in plasma	 Anti-B antibodies	 Anti-A antibodies	No antibodies	 Anti-A antibodies Anti-B antibodies
Can donate to	Type A Type AB	Type B Type AB	Type AB	All
Can receive from	Type A Type O	Type B Type O	All	Type O

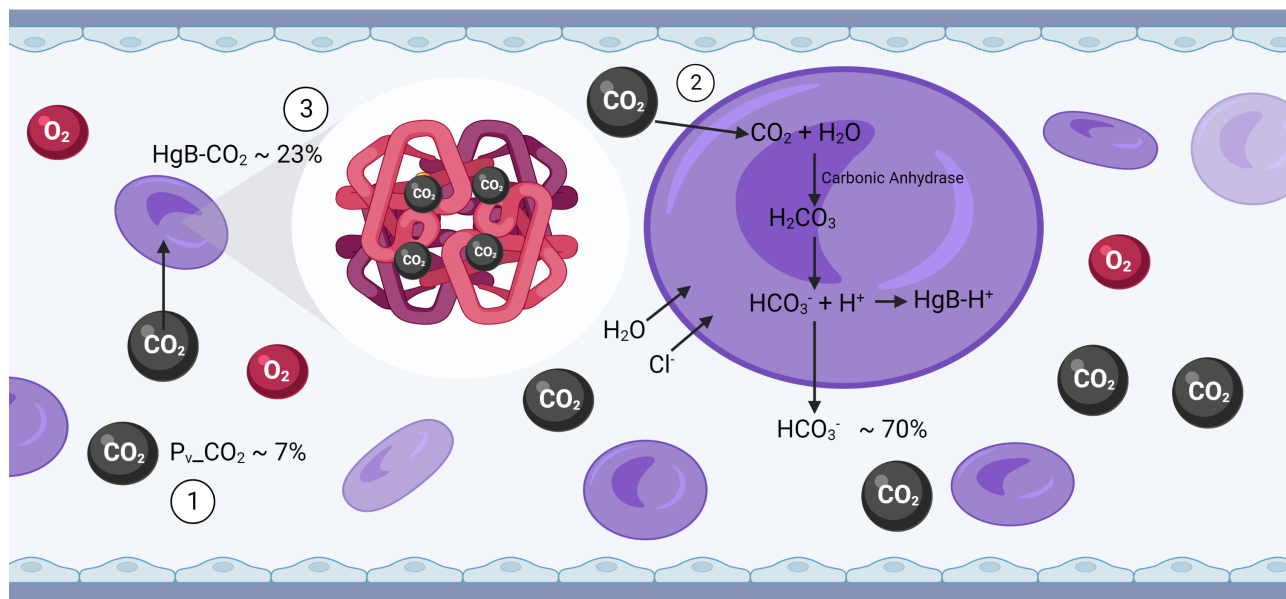
Fig. 11.5: ABO Blood Types (Created with BioRender.com)

11.3.2 Blood Transport of CO_2

There are three ways that CO_2 is transported in the blood (See Figure 11.6).

1. CO_2 is transported directly in the blood in a small quantity (7%) which generates the partial pressure of CO_2 in the blood (P_aCO_2 in arterial blood, and P_vCO_2 in the venous blood).
2. The majority of CO_2 is transported as bicarbonate (HCO_3^-) (approximately 70%) through the following chemical reaction within the red blood cells (RBCs). Within the RBC $CO_2 + H_2O \rightarrow H_2CO_3$ which is facilitated by the enzyme carbonic anhydrase. To buffer this acid, a hemoglobin (Hgb) will bind with H^+ to form $Hgb - H^+$ (which stays in the RBC) and a molecule of HCO_3^- which is released into the plasma through an exchange with Cl^- in a process known as the chloride shift.
3. The remaining CO_2 (approximately 23%) binds directly with Hgb to form $Hgb - CO_2$ which is referred to as carbaminohemoglobin (not to be confused with carboxyhemoglobin which is $Hgb - CO$ from the inhalation of carbon monoxide, the result being carbon monoxide poisoning).

These processes are facilitated by the movement of CO_2 into the RBCs which occurs due to the PCO_2 gradient between the plasma and the RBC. As the above reactions proceed the RBC PCO_2 drops and CO_2 will diffuse into the RBCs until the PCO_2 in the plasma and the RBC are equalized. As a reminder, gradient that pushes CO_2 into the plasma and then the RBCs originates in the mitochondria where CO_2 is being generated.

Fig. 11.6: Blood Transport of CO_2 (Created with BioRender.com)

11.3.3 Blood Transport of O_2

O_2 is transported directly in the blood in a small quantity (1-2%) which generates the partial pressure of O_2 in the blood ($P_a\text{O}_2$ in arterial blood, and $P_v\text{O}_2$ in the venous blood).

The majority of O_2 is transported in RBCs, attached to HgB. The partial pressure (PO_2) gradients facilitate the movement across membranes, including into the RBCs, which then allow O_2 -HgB binding. The affinity for O_2 -HgB binding is related to the PO_2 as well as to the number of binding sites of a HgB molecule already bound with O_2 . With each binding site that has an O_2 bound to it, there is an increased affinity for the other binding sites to bind. The binding of O_2 to HgB is also influenced by the temperature (increased temperature reduces O_2 -HgB affinity), and the pH (acidity reduces O_2 -HgB affinity), as well as a 2,3-bisphosphoglycerate (BPG) (increased BPG reduces affinity). Of these factors that shift the curve, temperature and pH vary quickly based on body status (exercise vs. rest), and between parts of the body (working muscle vs. nonworking muscle; and between muscle and lungs). But BPG varies across longer time periods as a RBC adaptation. BPG is found in RBCs and preferentially binds to HgB when O_2 is not bound to HgB. If there is an increase in BPG there is reduced affinity of HgB to bind with O_2 so that oxygen unloads from HgB more quickly when PO_2 drops. At first this sounds negative. However, reducing the affinity of HgB for O_2 is helpful when the goal is to release O_2 so that it can diffuse into the metabolically active cells requiring O_2 . Fluctuations in BPG tend to be RBC adaptations in response to chronic conditions such as chronic anemia, pulmonary disease and as an adaptation to high altitude.

The relationship between the PO_2 and the amount of O_2 bound to HgB expressed as the percentage of binding sites of HgB with O_2 (oxygen saturation, SO_2) is given in the oxyhemoglobin dissociation curve. The features of this curve depict, first, the relationship between PO_2 and SO_2 ; and shifts in the curve depict the influence of pH, PCO_2 , temperature, and BPG. The curve is usually considered from the perspective of arterial blood and therefore the relationship between $P_a\text{O}_2$ and $S_a\text{O}_2$ and the diffusion of O_2 during loading of HgB with O_2 . When considering $P_a\text{O}_2$ and $S_a\text{O}_2$ the oxygen content of arterial blood that is being delivered is being assessed, reduced values indicate hypoxemia. However, it is equally valid to consider the curve in the capillary or venous blood when considering the utilization of O_2 by the cells of the body; and the influence of the O_2 -HgB affinity on releasing (unloading) O_2 into the tissues. When $S_a\text{O}_2$ is estimated with a pulsed oximeter it is abbreviated as $S_p\text{O}_2$.

11.3.3.1 Oxyhemoglobin Dissociation Curve

Figure 11.7 depicts the oxyhemoglobin dissociation curve simulated with two different pH. The Y-axis is the SO_2 and the X-axis is the PO_2 . The PO_2 is the primary determinant O_2 -Hgb binding, and pH is a second determinant. The sigmoidal relationship identifies an important area in this relationship. A critical juncture where reductions of $PO_2 < 60$ include rapid reductions in SO_2 due to a reduced O_2 -Hgb binding affinity. When considering capillary and venous blood, this is a helpful attribute for unloading O_2 for diffusion into cells. However, if arterial O_2 partial pressure (P_aO_2) drops below 60 mmHg then there is barely enough O_2 -Hgb binding (hypoxemia is a $P_aO_2 < 60$ or $S_aO_2 < 90$). And further drops in P_aO_2 result in rapid and significant drops in S_aO_2 . The implications of the sigmoidal shape of the curve is further discussed in Chapter 12 on Ventilation since the most common reason for drops in P_aO_2 are changes and insufficiency in ventilation.

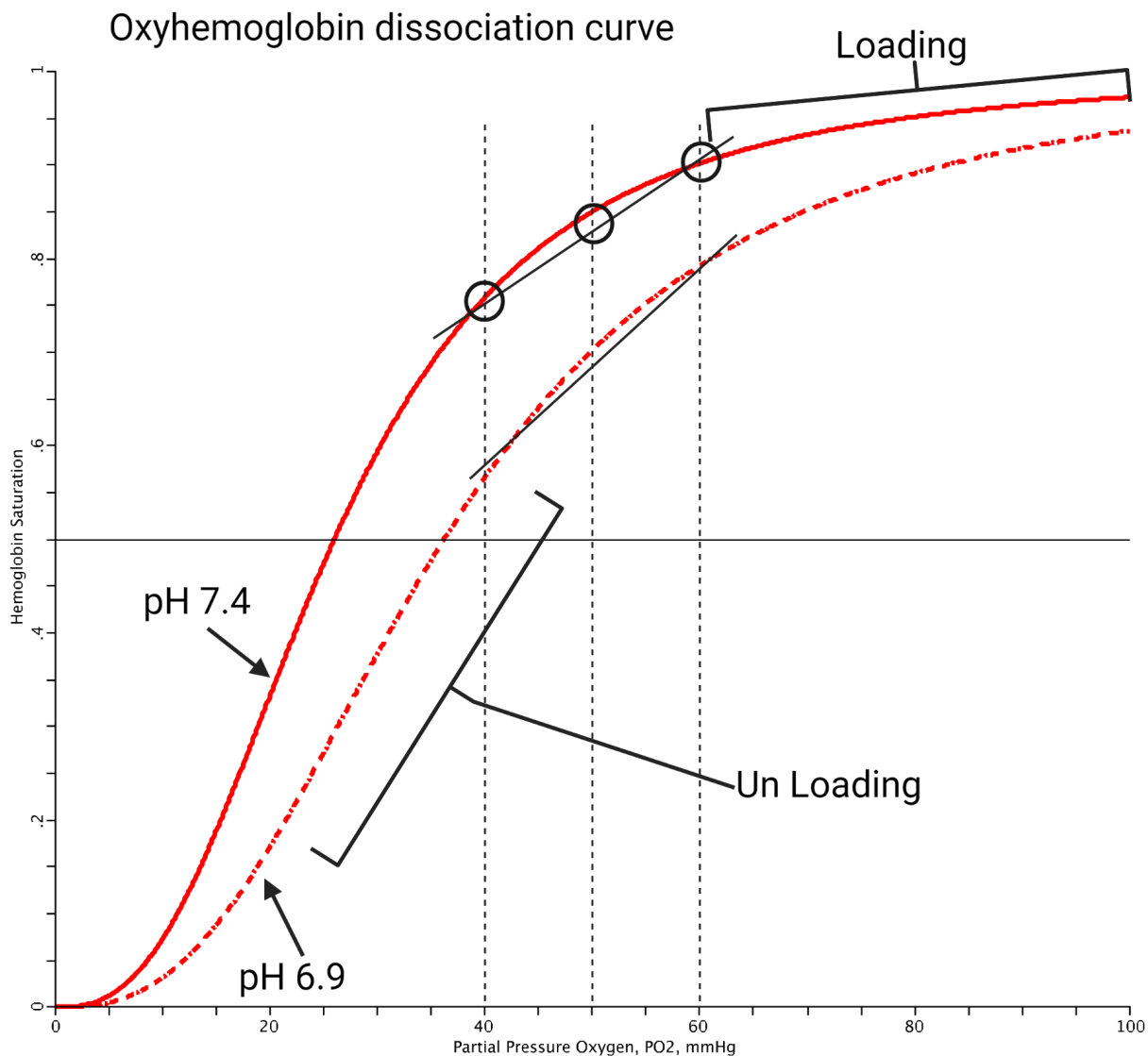


Fig. 11.7: Oxyhemoglobin Dissociation Curve (Simulation performed with JSim for physiological simulation under the Physiome project, <https://www.imagwiki.nibib.nih.gov/physiome/jsim>)

Table 11.1 depicts a linear approximation at the critical juncture in the oxyhemoglobin curve that is worthwhile to remember in the interpretation of arterial blood gases (ABGs) or pulsed oximetry. If S_pO_2 is 90%, then P_aO_2 can be estimated at 90 mmHg, and vice versa. A further drop in P_aO_2 of 10 mmHg to 50 mmHg will result in a drop in S_pO_2 to 80%.

S_aO_2	70	80	90
P_aO_2	40	50	60

Table 11.1: Linear Approximation of Oxyhemoglobin Dissociation Curve

Figure 11.7 also depicts a right shift to the oxyhemoglobin dissociation curve based on pH. Besides PO_2 , pH has the largest impact on HgB binding affinity. Fluctuations to pH occur as the blood circulates from working skeletal muscles to the lungs. In working skeletal muscle the pH tends to be lower due to the production of H^+ from the energetic pathways. The broken line in Figure 11.7 estimates the oxyhemoglobin dissociation curve in the skeletal muscle during exercise (favors unloading of O_2) when pH can drop to as low as 6.9. As can be seen in the figure, at any given PO_2 , less HgB is bound to oxygen with a lower pH. In working muscle this is helpful because it means more O_2 will be released from HgB and available for diffusion.

11.4 Pulmonary Circulation

For overall balanced circulation, where blood returns to the RA and leaves through the LV, the pulmonary circulation must equate to the systemic circulation. Therefore, all prior discussions about $\dot{Q}$, apply to the pulmonary circulation. One primary difference between the pulmonary and systemic circulation is the resistance. Resistance is much lower in the pulmonary circulation. Pulmonary circulation has much less length, and length is directly proportional to resistance (less length, less resistance). With less resistance the RV does not need to exert as much pressure (or tension) to produce circulation through the pulmonary circuit. Therefore, while the volume (L/min) through the pulmonary and systemic circulations are the same, the pressures are different by an order of magnitude (approximately 10x lower; Pulmonary Artery Pressure (PAP) is approximately 15-25/0-5 mmHg during RV systole / RV diastole).

The majority of pulmonary circulation is to the alveoli (small air sacs of the lung). Pulmonary (alveolar) circulation is a low pressure - high volume system with such a high density of capillaries that the blood flows almost as sheets in close contact surrounding the alveoli. Blood leaving the right ventricle (RV) flows into pulmonary arteries (venous blood) and into the alveolar circulation for respiration (gas exchange) with the alveoli to become arterial blood. After flowing through the capillaries of the alveoli this arterial blood returns to the left atrium (LA) through the pulmonary veins. Approximately 1-2% of cardiac output is sent to the bronchial arteries of the lungs for oxygenation of lung cells. This circulation is from the systemic circulation (from the left ventricle), but it returns its venous blood to the heart through the pulmonary veins into the LA. This small amount of venous blood mixes with the otherwise arterial blood in the LA and has a minor, inconsequential, effect on the P_aO_2 and S_aO_2 .

11.4.1 Pulmonary Circulation Pressures

Pulmonary artery pressure (PAP) is synonymous with arterial blood pressure (BP) and has both systolic and diastolic components. PAP is much lower than BP under normal circumstances, approximately 15-30 mmHg / 5-10 mmHg with pulmonary capillary pressures approximately 7 mmHg. Central pulmonary venous pressure is the pressure in the LA, approximately 0-2 mmHg, which is similar to the central systemic venous pressure in the RA.

Intensive (critical) care unit (ICU) monitoring for patients with cardiac or other blood flow or volume conditions are often monitored with a catheter that records pressure in the pulmonary artery for continuous PAP monitoring that can occasionally be pushed further into the pulmonary circulation to estimate the pulmonary capillary pressure with a measurement referred to as the pulmonary capillary wedge pressure (PCWP). The PAP and PCWP are utilized to assess whether left sided heart function is adequate for the current blood volume, or whether the pulmonary circulation

pressures are rising. If the pulmonary circulation pressures are too high then the pulmonary microcirculation favors filtration which could result in pulmonary edema or effusion.

11.4.2 Pulmonary Microcirculation & Lymphatics

The same four pressures determine the microcirculation in the lungs as in peripheral capillaries and tissues with a few important differences. The pulmonary capillary hydrostatic pressure is lower and the interstitial fluid pressure is negative due to the loss of fluid through the alveolar membrane and resultant evaporation (incidental water loss through breathing). Alveolar capillary permeability is high, allowing extra amounts of protein to leak from the capillaries; therefore, interstitial fluid osmotic pressure is high. The alveolar walls are thin but are generally impermeable to the proteins and solute (but are permeable to water). Total microcirculation outward pressure moving out of the capillaries exceeds the inward pressure by about 1 *mmHg* promoting a small continual loss of fluid from the pulmonary capillaries. The capillary fluid loss is removed from the area partly through lymphatic flow and partly through alveolar evaporation of water. Some of the fluid seeps through the lung interstitial tissue and ends up in the pleural cavity surrounding the lungs, where it is typically drained into the lymphatics that drain the pleural space.

The same safety factors exist to prevent edema as in the peripheral tissue, and the same factors can cause edema. An important difference between peripheral tissue edema and pulmonary tissue edema is that pulmonary tissue edema will become pulmonary edema (as in alveolar edema). Pulmonary (alveolar) edema is life threatening because it impairs pulmonary respiration and can cause hypoxemia. The most common cause of pulmonary edema is left sided heart failure. If there is a slow accumulation (more often caused by reduced vascular osmotic pressure, or poor lymph function) of pulmonary tissue edema it may cross the pleural membranes and result in pleural edema (commonly referred to as a pleural effusion). The most common causes of pleural effusion include: blockage of lymphatic drainage from the pleural cavity; heart failure causes excessively high peripheral and pulmonary capillary pressures; decreased vascular osmotic pressure; increased capillary permeability caused by infection or any other source of inflammation of the pleural surfaces.

11.4.3 Regulation of Pulmonary Circulation

The quantity of pulmonary circulation equals the quantity and rate of systemic circulation. The venous return that enters the right side of the heart and flows to the pulmonary circulation must equal the output of the left ventricle. This equality does not have to be beat to beat, but certainly must be equal when integrated over time (i.e., minute by minute). Otherwise there would be an accumulation of blood in the pulmonary circulation that impairs blood flow, ventilation and respiration, if more entered the lungs than was ejected from the LV. Or, a continuous decrease in LV output, if more was ejected from the LV than came through the pulmonary circulation. Therefore, the overall regulation of pulmonary circulation quantity includes the same processes that regulate systemic circulation as covered in Chapters 8 and 9.

As with the peripheral circulation the distribution of pulmonary circulation is dependent on local regulation. One fundamental difference is that local regulation of pulmonary (alveolar) circulation through tone (vaso dilation and constriction) of arterioles is based on **hypoxic vasoconstriction**. This is the opposite of muscular arterioles. In the muscular (systemic circulation) arterioles local indications of low O_2 concentration and high metabolism result in vasodilation in an effort to increase blood flow to those capillaries. In the pulmonary circulation arterioles, low O_2 concentration results in arteriole vasoconstriction, in an effort to shift blood away from those capillaries. This makes complete sense for the distribution of circulation in the lungs, where the reason for blood to be delivered to capillaries surrounding alveoli is to obtain O_2 , not to deliver O_2 . If some alveoli have poor airflow and low P_AO_2 then it is better for less blood to circulate to that alveolar-capillary network. Arterioles of alveoli-capillary networks that have high P_AO_2 vasodilate, to promote circulation to those alveoli with more O_2 .

Chapter 12 on Ventilation considers the concept of ventilation - perfusion $\dot{V}/\dot{Q}$ matching, the impact of body position on $\dot{V}/\dot{Q}$ matching zones, and the impact of $\dot{V}/\dot{Q}$ matching on the $P_AO_2 - P_aO_2$ gradient, otherwise known as the A-a gradient.

Respiratory sinus arrhythmia (RSA) refers to the observation that in healthy individuals at rest there is cyclic variation in heart rate with the respiratory (breathing) frequency. With inhalation (inspiration) HR tends to go up a few beats per minute; with exhalation (expiration) HR tends to diminish a few beats per minute. The RSA is a dominant source of what is called high frequency heart rate variability (HRV) and is associated with parasympathetic nervous system activity (vagal cardiac tone). It has been hypothesized that RSA helps to maximize blood flow through the pulmonary circulation when alveoli have more ventilation as a feedforward (not feedback) system of control. This control is feedforward because the HR increases in response to the same trigger for inhalation, not in response to there being more oxygen in the lungs. There is evidence that RSA helps to maximize ventilation - perfusion matching at rest [HY03].

11.5 External (Alveolar) Respiration

External (alveolar) respiration is gas exchange between the alveolar capillary blood and the alveoli. As with in the systemic capillaries and tissues diffusion occurs with the partial pressure gradients of O_2 and CO_2 . Diffusion is influenced by the factors in Equation 11.2. The structure of alveoli (thin membrane with a high surface area) and alveolar capillaries (dense network forming sheets in close contact with the thin alveolar membrane) are well equipped for rapid diffusion and full equilibration of O_2 and CO_2 both at rest and with higher rates of pulmonary circulation such as with high cardiac output during maximal exercise.

11.5.1 Alveolar O_2 Diffusion

A ventilated alveoli at sea level has a partial pressure of oxygen (P_{AO_2}) of approximately 104 mmHg. The atmosphere at sea level has a partial pressure of O_2 of approximately 159.5 mmHg. The PO_2 from the atmosphere to the arterial blood drops as it passes into the alveoli because air is humidified (some of the pressure is now due to water molecules); because CO_2 is being diffused into the alveoli (some of the pressure is due to CO_2); and because O_2 is being diffused out of the alveoli (there is a regular reduction in the concentration of O_2). Maintaining a $P_{AO_2} \approx 104\text{mmHg}$ requires continuous air exchange with the atmosphere (alveolar ventilation, V_A , the primary topic for the next Chapter 12 on Ventilation).

The full equilibration of blood at a single alveoli is depicted in Figure 11.8. Note that the $P_aO_2 = P_{AO_2}$. This requires alveoli with normal membrane thickness and surface area in contact with capillaries, and normal capillary membranes. During equilibration of O_2 binding to HgB based on the P_aO_2 according to the oxyhemoglobin curve (Figure 11.7) ensures the blood O_2 carrying capacity is sufficient to support cellular respiration.

11.5.1.1 Atmospheric Limits

The atmospheric $P_{atm}O_2$ places a limit on how high the alveolar partial pressure (P_{AO_2}) can be and therefore the highest value of P_{AO_2} and P_aO_2 . This upper limit means now matter how much someone breathes, there is an upper limit on diffusion of oxygen into the blood. That upper limit is reduced with lower atmospheric $P_{atm}O_2$, which is the case with higher altitude. A hyperbaric (higher pressure) chamber increases $P_{atm}O_2$, and subsequently P_{AO_2} , and P_aO_2 .

! Supplemental O_2

The use of supplemental O_2 changes the P_{AO_2} and P_aO_2 by increasing the F_iO_2 . Since $PO_2 = F_iO_2 \times \text{Pressure}$, then if F_iO_2 is 0.4 at sea level the $P_{atm}O_2 = 0.4 \times 760 = 304\text{mmHg}$, and the $P_{AO_2} \approx 200\text{mmHg}$. Supplemental O_2 is utilized when there are impairments in alveolar respiration or alveolar ventilation. The goal is to treat hypoxemia by improving the partial pressure gradient of O_2 in ventilated alveoli to make up for problems in alveoli that otherwise limit diffusion (poor circulation, increased diffusion membrane thickness, decreased diffusion surface area, or poor ventilation).

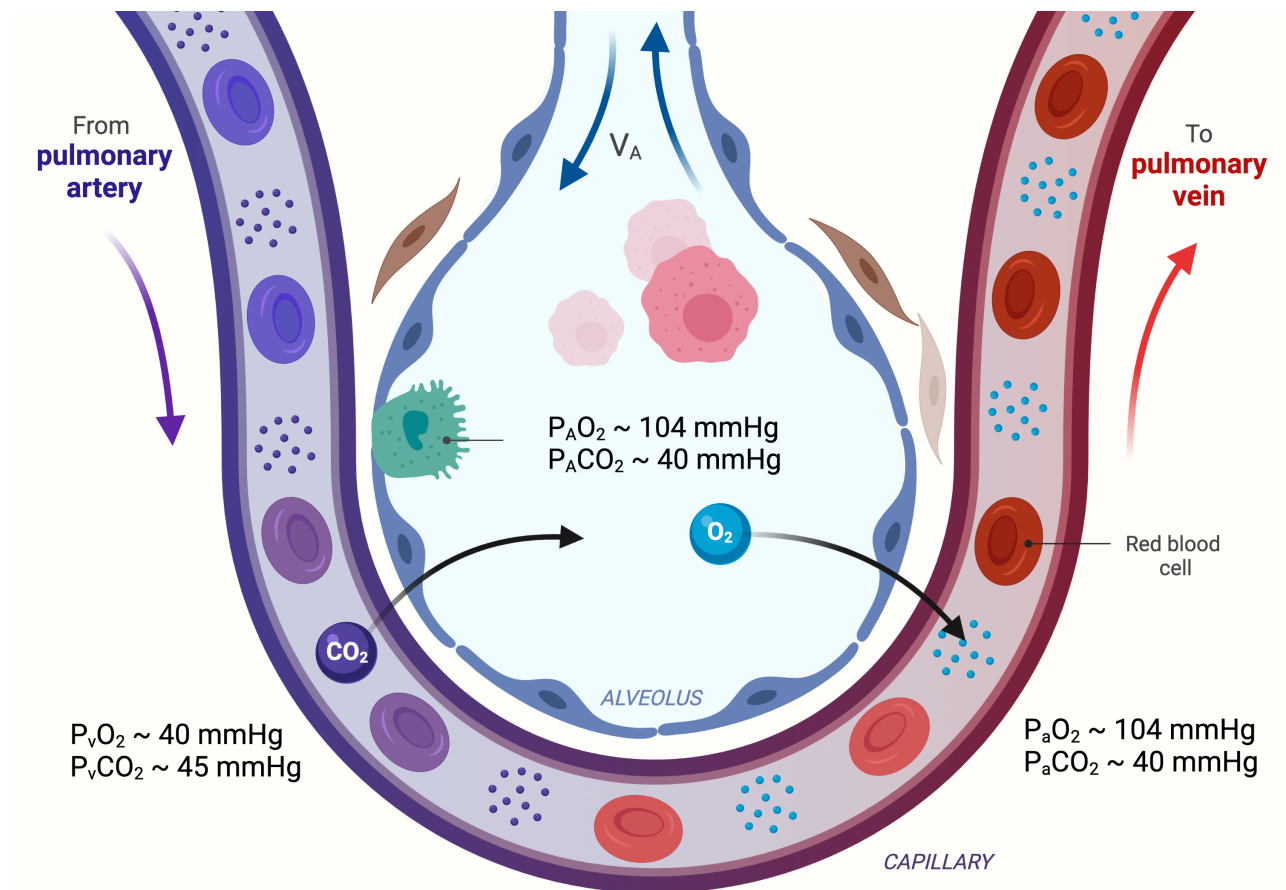


Fig. 11.8: Equilibration of Alveolar Diffusion (Created with BioRender.com)

11.5.2 Alveolar CO_2 Diffusion

A ventilated alveoli at sea level will have a partial pressure of carbon dioxide (P_{ACO_2}) of approximately 40 mmHg. The PCO_2 drops dramatically from the arterial blood to the atmosphere ($P_{atmCO_2} \approx 0.3 \text{ mmHg}$). The full equilibration of blood so that the $P_aCO_2 = P_{ACO_2}$ requires alveoli with normal membrane thickness and surface area in contact with capillaries, and normal capillary membranes (See Figure 11.8). During equilibration of CO_2 the processes discussed in the section on CO_2 transport in the blood in reverse to release CO_2 from the blood (See Figure 11.6).

11.5.2.1 Atmospheric Limits

The atmospheric PCO_2 places a very low limit on the alveolar partial pressure (P_{ACO_2}) and therefore there is a very low limit for P_aCO_2 . The lack of a limit means high alveolar ventilation (V_A) can greatly reduce the P_aCO_2 . The situations of hyperventilation (lower than normal P_{aCO_2}), and hypoventilation (higher than normal P_{aCO_2}) are discussed in Chapter 12.

! P_aCO_2 and Brain Blood Flow

Hyperventilation, which is often associated with pain and anxiety, includes expiring more CO_2 than is being produced metabolically. It reduces PCO_2 levels and can result in an increase in blood pH (alkaline). A consequence of reduced blood pH is vasoconstriction which decreases cerebral blood flow. This relationship provides an explanation for

lightheadedness during hyperventilation (reduced P_aCO_2), and signs and symptoms attributed to hyperventilation (or over breathing) syndrome [CC08].

11.6 Practice Connections

11.6.1 Arterial Blood Gases

Arterial Blood Gases (ABGs) are measured from a sample of arterial blood and provide information on the arterial (plasma) partial pressure of O_2 and CO_2 , the arterial blood pH, and the arterial blood bicarbonate (HCO_3^-). There are several standards for reporting these values in tabular form. A common convention and normal values are depicted in Table 11.2

P_aO_2	P_aCO_2	pH	HCO_3^-
95-100 mmHg	35-45 mmHg	7.35-7.45	22-26 mmol/L

Table 11.2: Normal Ranges of the Arterial Blood Gases

ABGs are extremely valuable measurements for the examination of respiration impairments. Some impairments reduce the P_aO_2 (hypoxemia) but do not impact the P_aCO_2 and therefore do not create acid base imbalance. Whereas other impairments reduce P_aO_2 and increase P_aCO_2 (hypercarbia, also known as hypercapnea) and therefore create acid base imbalances (respiratory acidosis).

11.6.2 Acid - Base Balance & Disorders

Acid base balance is when pH is at its normal blood level of 7.35-7.45. Acid base balance is maintained due to buffering capacity in both ICF and ECF. The ECF buffering capacity of blood is related to renal function (long term) and respiration (short term). Disorders of acid base balance are classified as acidosis when pH is below 7.35, and alkalosis when pH is above 7.45. There are two general causes of imbalances, respiratory (due to altered P_aCO_2), or metabolic (primarily renal) due to altered HCO_3^- . The four possible situations are:

- Respiratory Acidosis: Decreased pH due to elevated P_aCO_2
- Respiratory Alkalosis: Increased pH due to decreased P_aCO_2
- Metabolic Acidosis: Decreased pH due to decreased HCO_3^-
- Metabolic Alkalosis: Increased pH due to increased HCO_3^-

Respiratory and renal mechanisms can compensate for acid base disturbances created by the other system. Respiratory compensation refers to the use of respiratory adjustments to P_aCO_2 to compensate (normalize, but not completely, pH) when there is a metabolic disturbance. Renal compensation refers to the use of renal clearance adjustments to HCO_3^- and H^+ to compensate (normalize, but not completely, pH) when there is a respiratory disturbance. Respiratory compensation is rapid since it simply involves changing breathing rate and depth (ventilation), it can occur in seconds to minutes. Renal compensation involves renal clearance and possibly production of more HCO_3^- , so it takes hours to days, depending on what is required, GFR, and urine output (which depends on fluid intake).

11.6.2.1 Interpretation of ABGs to Evaluate Acid Base Balance

Figure 11.9 provides a decision tree for the evaluation of acid base balance. The process starts with an assessment of the pH. Since compensation can return pH close to the normal of 7.4 (and therefore within the normal range of 7.35-7.45)

the process considers the range of pH but decisions of acidosis and alkalosis are initially based on whether pH is below or above 7.4. The decision tree starts with the claim that there is an imbalance even if pH is within the range of 7.35-7.45, so the interpretative process does not simply stop if pH is within the normal range due to compensation.

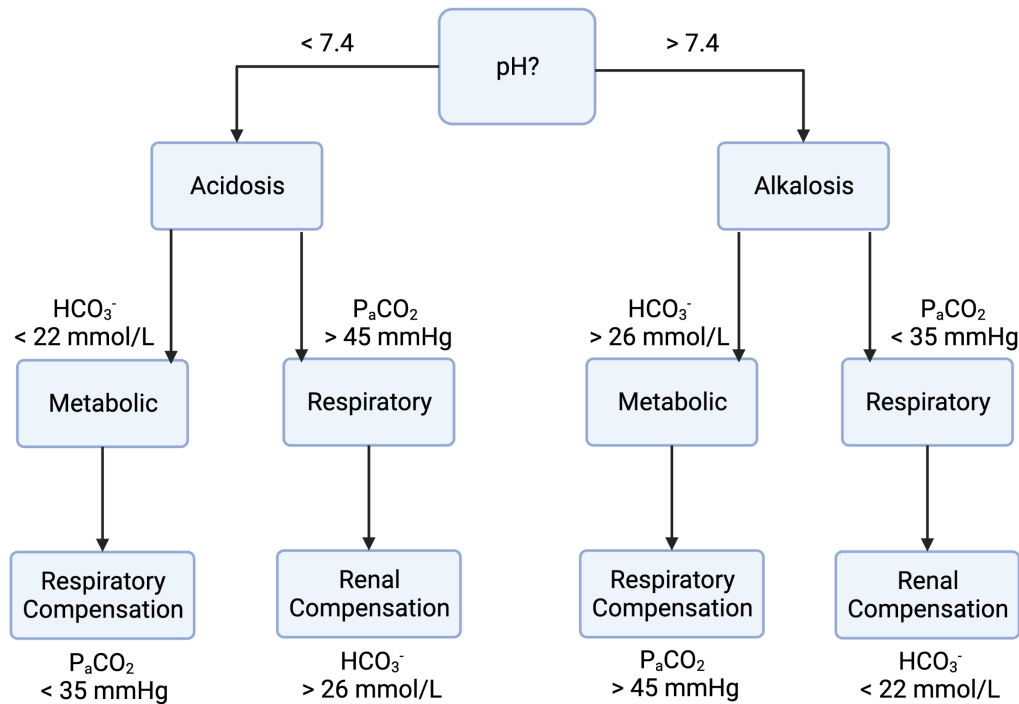


Fig. 11.9: Interpretation of ABGs to Evaluate Acid Base Balance

Individuals with COPD (emphysema and/or chronic bronchitis are the primary causes) tend to have respiration impairments that result in respiratory acidosis due to elevated P_aCO_2 . Interpretation of ABGs can evaluate whether the current status of a patient is acute or chronic. If a patient has respiratory acidosis that is not compensated then the patient is having an acute exacerbation of their chronic respiration impairment. However, if the respiratory acidosis is compensated then this degree of respiratory acidosis has had time for renal compensation to occur. It may still be relatively new for the patient, but since renal compensation can take days, it is not acute in the sense that it is unlikely to be new at the time of measurement. The clinician needs to consider "acute" vs. "chronic" with regard to respiration impairments since problems can be acute but still be a day old. The best way to know whether someone's respiration impairment is worsened is to compare to previous ABG measurements, ideally those taken when not in distress.

11.6.3 Pulmonary Function Testing - DLCO

DLCO is a pulmonary function test that evaluates the diffusion capacity of the lung for carbon monoxide (CO). It is used to evaluate the extent that O_2 diffuses across the alveoli-capillary membrane into the capillary blood impacting the P_aO_2 . The test involves measuring the partial pressure difference between inspired and expired CO . It relies on the strong affinity and large absorption capacity of red blood cells for carbon monoxide and demonstrates gas uptake by the

capillary blood. DLCO is also affected by the HgB, carboxyhemoglobin, age and sex. DLCO is measured in $ml/min/kPa$ (kPa as an SI unit of pressure, the kilopascal). However, DLCO is usually reported as a percent of predicted value based on age and sex. Values between 80-120% predicted are generally accepted as normal. Values below 80% indicate possible limitations to O_2 diffusion in the lung. With a low DLCO it is less likely that capillary blood will equilibrate with P_AO_2 while in the alveolar capillary, and is therefore more likely that P_aO_2 will be reduced. DLCO may be reduced in individuals with conditions that impact the thickness or surface area of alveoli (Equation 11.2) such as emphysema, pulmonary fibrosis, or pulmonary hypertension. DLCO helps with examination of whether someone has a respiration impairment and the risk that they will have hypoxemia at rest, or become hypoxemic with exertion.

11.6.4 Pulsed Oximetry

Pulsed oximetry indirectly monitors the O_2 saturation of arterial blood (S_pO_2) as opposed to a measurement directly from arterial blood (S_aO_2). A pulsed oximetry device is commonly referred to as an oximeter, or more colloquially a "pulse ox". Oximeters use an electronic processor and two small light-emitting diodes (LEDs) facing a photodiode through a translucent part of the skin, such as a fingertip or earlobe. Each LED emits a different wavelength of light. Absorption of light at these wavelengths differs significantly between O_2 -HgB and HgB. The amount of light not absorbed from each LED is measured, and separate normalized signals are produced for each wavelength. These signals fluctuate in time because the amount of arterial blood that is present increases (literally pulses) with each heartbeat. The signals are used to calculate an estimate of S_aO_2 known as S_pO_2 . Most units are accurate within 2.5% down to approximately 70% saturation. The signals also reflect the volume of pulsatile capillary flow (photoplethysmogram (PPG)) that pulses with left ventricular systole and can therefore also be used to determine pulse rate (as an estimate of heart rate), and indices related to blood flow volume and timing.

Poor blood flow to the extremity may cause erroneously low readings which can be as simple as a cold limb, or vasoconstriction (so not all low readings are low). There are no known causes of erroneously high readings (so normal / high readings tend to be accurate). This means that as an indicator of hypoxemia, oximetry has more false positives and limited false negatives. Such a diagnostic accuracy measurement is said to have low specificity and high sensitivity and are better for ruling out a situation than they are for ruling in a situation. For example, if S_pO_2 is normal then we are confident it is normal so we feel confident we have ruled out hypoxemia. If S_pO_2 is less than 89% we have less confidence about hypoxemia because there is a risk of a false positive, so we have less confidence to rule in hypoxemia. Other problems can include incorrect sensor application (such as not placing the LEDs on the thinner skin surface area); highly calloused skin; movement; finger nail polish; or higher skin pigmentation. Oximeters that include a PPG waveform are usually more expensive but allow an assessment of sensor accuracy by verification that the sensor is producing a steady pulse waveform. In the absence of a PPG waveform another approach to confirm accuracy is to verify the pulse rate estimate from the oximeter is the same as the radial pulse.

S_pO_2 helps to examine whether someone has hypoxemia at the time of measurement. Since it is easily measured, it provides continuous information about hypoxemia. For example, in someone with COPD that has a reduced DLCO but is not hypoxemic at rest, using an oximeter to monitor the S_pO_2 allows the clinician to know the point in time that the patient became hypoxemic, either due to exertion with exercise, ADLs, or position changes.

11.7 Summary

The mass balance of O_2 and CO_2 requires the diffusion of these gases at several points along what is a long chain of intake (for O_2) and output (for CO_2). Cellular and internal respiration, blood transport, and external respiration are all necessary and important steps in this system wide gas exchange. Understanding these steps and the factors that influence the steps are important for understanding the normal response to increased muscle need for O_2 , and situations that include impairments at any point along the pathway for O_2 . Clinical approaches exist to evaluate the current status of respiration (ABGs and oximetry). ABGs provide more information but require arterial blood sampling. Oximetry provides less information, but is easily obtained for continuous monitoring. DLCO provides an evaluation of the risk of hypoxemia.

11.7.1 Next Step

This chapter presents an incomplete story of overall pulmonary support for the mass balance of O_2 and CO_2 . It leaves off with alveolar respiration, and only considers alveolar respiration at one alveolus. In the next chapter we consider ventilation (breathing) mechanics and the process of drawing air into and pushing air out of the alveoli. Without continuous alveolar ventilation (V_A) there is no alveolar respiration; and without ventilation (V_e), there is no V_A . We also consider the importance of the distribution of V_e throughout the lungs that results in regional differences in V_A , and the distribution of pulmonary perfusion ($\dot{Q}$). Alveolar respiration cannot occur unless there is V_A matched with capillary blood flow (perfusion).

References

- [CB17] Kevin M. Christmas and James B. Bassingthwaite. *Equations for O_2 and CO_2 solubilities in saline and plasma: combining temperature and density dependences*. en. Journal of Applied Physiology 122.5 (May 2017), pp. 1313–1320. ISSN: 8750-7587, 1522-1601. DOI: [10.1152/japplphysiol.01124.2016](https://doi.org/10.1152/japplphysiol.01124.2016). URL: <https://www.physiology.org/doi/10.1152/japplphysiol.01124.2016> (visited on 10/31/2021).
- [CC08] Rosalba Courtney and Marc Cohen. *Investigating the Claims of Konstantin Buteyko, M.D., Ph.D.: The Relationship of Breath Holding Time to End Tidal CO_2 and Other Proposed Measures of Dysfunctional Breathing*. en. The Journal of Alternative and Complementary Medicine 14.2 (Mar. 2008), pp. 115–123. ISSN: 1075-5535, 1557-7708. DOI: [10.1089/acm.2007.7204](https://doi.org/10.1089/acm.2007.7204). URL: <http://www.liebertpub.com/doi/10.1089/acm.2007.7204> (visited on 03/29/2021).
- [HH20] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences, 2020.
- [HY03] J Hayano and F Yasuma. *Hypothesis: respiratory sinus arrhythmia is an intrinsic resting function of cardiopulmonary system*. en. Cardiovascular Research 58.1 (Apr. 2003), pp. 1–9. ISSN: 00086363. DOI: [10.1016/S0008-6363\(02\)00851-9](https://doi.org/10.1016/S0008-6363(02)00851-9). URL: [https://academic.oup.com/cardiovascres/article-lookup/doi/10.1016/S0008-6363\(02\)00851-9](https://academic.oup.com/cardiovascres/article-lookup/doi/10.1016/S0008-6363(02)00851-9) (visited on 07/19/2022).
- [Sch+97] Kenneth A. Schenkman et al. *Myoglobin oxygen dissociation by multiwavelength spectroscopy*. en. Journal of Applied Physiology 82.1 (Jan. 1997), pp. 86–92. ISSN: 8750-7587, 1522-1601. DOI: [10.1152/jappl.1997.82.1.86](https://doi.org/10.1152/jappl.1997.82.1.86). URL: <https://www.physiology.org/doi/10.1152/jappl.1997.82.1.86> (visited on 07/21/2022).

Sample Questions

1. What does DLCO stand for and what is it directly and inversely proportional to?
2. If cells are using only carbohydrate for energy metabolism (glycogen to CO_2 pathway only), and they are using 5 ml O_2 / 100 mL of blood. What is the rate of CO_2 production?
3. Based on the oxyhemoglobin and oxymyoglobin dissociation curves, what has the strongest effect on oxygen loading (increased binding, increased affinity) and unloading (decreased binding, decreased affinity), and therefore S_aO_2 and $MgB - O_2$?
4. What is a physiological reason for increasing fatigue related to heart failure that involves renal function and blood oxygen levels?
5. How do pulmonary arterioles respond to a drop in alveolar P_AO_2 ?
6. Interpret: $P_aO_2 = 60$ $P_aCO_2 = 60$ pH = 7.23 $HCO_3^- = 22$
7. Interpret: $P_aO_2 = 95$ $P_aCO_2 = 30$ pH = 7.4 $HCO_3^- = 35$

Chapter 12

Ventilation

Updated on May 22, 2024

Contents

12.1	Ventilation Overview	194
12.2	Airways	195
12.3	Volumes & Capacities	200
12.4	Mechanics: The Ventilatory Pump	201
12.5	Ventilation / Perfusion Matching	210
12.6	Regulation	213
12.7	Practice Connections	216
12.8	Summary	218

Oxygen intake and carbon dioxide output are necessary for skeletal muscle energetics. Sustaining a flow of O_2 into the mitochondria requires the intake of O_2 by drawing air into the lungs (ventilation) and the alveoli (alveolar ventilation). Similarly, the flow of CO_2 out of the mitochondria requires expiring CO_2 out of the alveoli and lungs. This chapter covers alveolar ventilation (V_A), ventilation (V_e), ventilation/perfusion (V/Q) matching, the regulation of respiration through ventilation, and an introduction to several practice considerations.

Objectives include:

1. Explain the relationship between ventilation, alveolar ventilation and ventilation/perfusion matching (assumes you know what all of these things are and factors that influence them which are covered throughout the entire chapter not just the overview).
2. Explain the airways of ventilation.
3. Explain the impact of breathing pattern on anatomical dead space.
4. Explain bronchopulmonary hygiene and airway clearance.
5. Explain the benefit of breathing through the nasal passageways.
6. Explain bronchodilation and bronchoconstriction.
7. Explain ventilation volumes and capacities including the connection to the thorax and relate to chest wall excursion as ROM
8. Explain the mechanics of the ventilatory pump
9. Explain ventilation regulation
10. Explain practice considerations of ventilation
11. Integrate the concepts of this chapter with those of the Respiration and Circulation Chapters.

12.1 Ventilation Overview

Chapter 11 on Respiration established the need for alveolar ventilation and alveolar circulation. There are three recurring and important preliminary concepts: alveolar ventilation (V_A), ventilation (V_e) and ventilation / perfusion matching (V/Q).

12.1.1 Important Preliminary Concepts

- Alveolar ventilation (V_A) is the movement of air into, and out of, the alveoli. Without continuous V_A there is no alveolar respiration.
- V_e is the movement of air into, and out of, the lungs. V_e is necessary for V_A . Without ventilation (V_e), there is no V_A .
- Pulmonary perfusion is based on circulation (Q) and must be primarily distributed to alveoli receiving alveolar ventilation (V_A). V/Q matching refers to the matching of V_A and pulmonary circulation (Q). Both V_A and Q are necessary for alveolar respiration. Without V_A matching with pulmonary Q there is no alveolar respiration.

V_e is the volume of air inspired (breathe in) or expired (breathe out) in a time period. It is most commonly measured in mL/min or L/min . As a volume per minute, it is the volume of air per breath times the number of breaths per minute (bpm):

$$V_e(mL/min) = V_t(mL) \times RR(bpm) \quad (12.1)$$

In Equation 12.1, the abbreviation V_t stands for tidal volume. Tidal volume is the volume of air inspired (or expired) with one breath. The underlying assumption for both V_e and V_t is that the volume of air inspired equals the volume of air expired. While this might not be the case breath by breath, it is the case over a minute (or longer) time periods.¹

? Question

If there is an imbalance between V_t in and V_t out, it tends to be limited to transient periods that involve dynamic changes in V_t . For example, when V_t increases in the transient periods from rest to exercise, or during exercise when there are variations in O_2 needs.

What would happen if the volume inspired did not equal the volume expired over a long period of time? Try this, breathe more in than out, how long can you do that? Now breathe more out than in, how long can you do that?

Not all V_e ends in the alveoli (V_A), either per breath (V_t), or per minute (V_e). And, not all V_A is going to be to alveoli with perfusion (receive capillary circulation). Figure 12.1 depicts the general concept that V_e is going to be greater than V_A ; and that V_A is going to be greater than the V/Q . Pulmonary regulation is largely based on minimizing the discrepancy between V_e , V_A and V/Q .

The importance of this concept is that the alveolar respiration relies on the match between alveolar ventilation (V_A) and perfusion (Q). The consequences of all pulmonary (lung), and many cardiovascular conditions, ultimately rest on how they impact the ability to ventilate (V_e), the ability to turn as much V_e into alveolar ventilation (V_A) as possible, or the ability to have as much V_A perfusion as possible. Take home message: Many cardiopulmonary conditions can be considered based on the impact they have on V/Q matching.

¹ I've avoided analogies between ventilation and cardiac output for the sake of space and time. However, for those inclined to look for them, you'll find them.

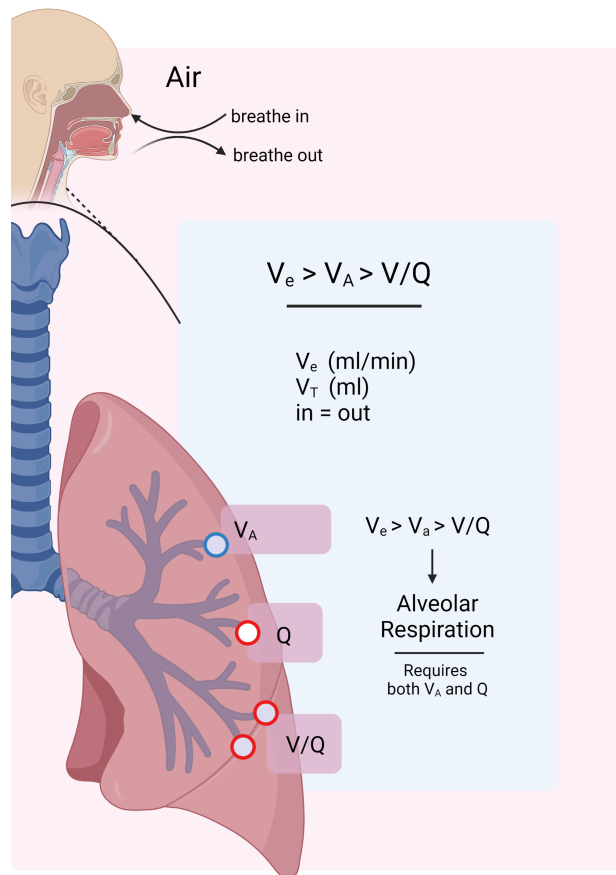


Fig. 12.1: Ventilation & Alveolar Ventilation
(Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

12.1.2 Dead Space Ventilation (V_D)

Dead space ventilation is ventilation that does not reach the alveoli (anatomical dead space), or that reaches the alveoli but does not receive perfusion (alveolar dead space). The combination of anatomical and alveolar dead space is called physiological dead space (V_D). To understand anatomical dead space (V_{Dana}) it is important to understand the structure and function of the airways.

12.2 Airways

The pathway of ventilation involves passage of air through a set of airways. Only the alveoli are capable of gas exchange (diffusion, respiration). The difference between V_e and V_A starts with the airways.

Figure 12.2 includes the airways connecting the environment to the alveoli.

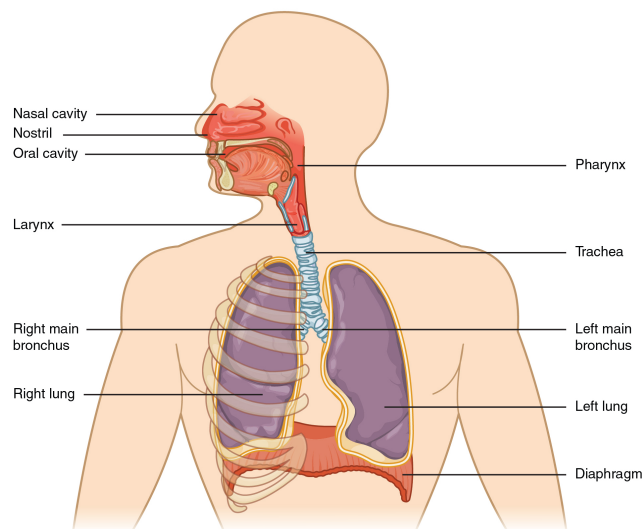


Fig. 12.2: Major Airways and Anatomy of Ventilation (OpenStax Anatomy, CCBYSA4.0)

12.2.1 Anatomical Dead Space (V_{Dana})

Any air that enters the airways is part of V_e , but it is not part of V_A unless it reaches an alveoli. For example, at the end of an inspiration the air remaining in the upper airways (nasal passages, the back of the throat, trachea, bronchi, bronchioles) is part of V_e , but not part of V_A (See Figure 12.3).

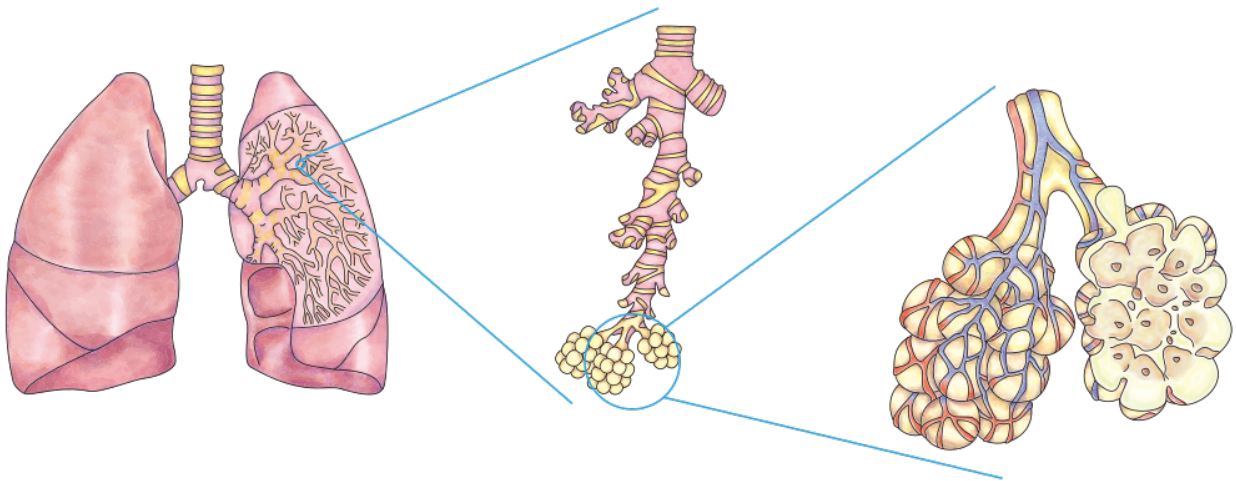


Fig. 12.3: Detail of Airways from Trachea to Alveoli (WikiMedia Commons, CCBYSA4.0)

Air in the airways that does not reach the alveoli (airway conduction zone) is anatomical dead space (V_{Dana}). V_{Dana} is a considered fixed volume (mL) per breath. For example, if $V_{Dana} = 100mL$, and $400 mL$ is inspired with tidal volume (V_T), then $300 (mL)$ reaches alveoli (V_A).

The air that does reach the alveoli is alveolar ventilation (V_A), and is the ventilation minus the anatomical dead space ventilation:

$$V_A(mL/min) = V_e(mL/min) - V_{Dana}(mL/min) \quad (12.2)$$

? Question

Does $V_A(mL/min)$ change if $V_e(mL/min)$ stays the same but is achieved with different volume (mL) per breath at different breathing frequencies (respiratory rates)?

Answer this by figuring out the $V_A(mL/min)$ and $V_e(mL/min)$ if $V_{Dana} = 100mL$ per breath, when breathing 24 breaths per minute at $250 mL/breath$; as compared to breathing 12 breaths per minutes at $500 mL/breath$; as compared to breathing 6 breaths per minute at $1000 mL/breath$. With each breathing pattern, $V_{Dana} = 100mL$ per breath since this is a fixed volume.

The answer requires using Equation 12.1 ($V_e(mL/min) = V_t(mL) \times RR(bpm)$) and substituting V_e with $V_t(mL) \times RR(bpm)$ and adjust V_t based on V_{Dana} since anatomical dead space occurs with each breath:

$$V_A(mL/min) = (V_t(mL) - V_{Dana}(mL)) \times RR(bpm) \quad (12.3)$$

12.2.2 Upper Airways

The upper airways include the nasal passageways, the nasopharynx, and the pharynx. Down to the pharynx the oral passages are also an option for breathing, but mouth (oral) breathing is sub optimal. The nasopharynx passes into the oropharynx which passes into the hypopharynx (also called the laryngopharynx). At this point the epiglottis is the gatekeeper between the larynx (for air) and the esophagus (for food and drink). The larynx is the beginning of the lower airways; it includes the vocal cords and then continues into the trachea.

All airway passages down as far as the terminal bronchioles are lined with epithelial cells that include, in various locations, mucus production cells and cilia. Cilia are small cellular extensions that promote the movement of mucus across the membrane surface. Cilia move continually in a direction that moves mucus toward the pharynx. Mucus and cilia in the airways are combined as a defense mechanism referred to as mucociliary activity. Together they promote the trapping of particles in mucus (including pathogens) and moving the mucus for removal. Cilia movement in the lower airways (trachea and bronchi) is upward, whereas those in the upper airways is downward. Continual cilia movement moves the mucus toward the pharynx. The mucus and its trapped particles are then swallowed or coughed out of the body.

12.2.2.1 Bronchopulmonary Hygiene & Airway Clearance

Mucociliary activity, along with regular deep breathing, is an important component of bronchopulmonary hygiene (BPH); while throat (pharynx) clearing, sneezing, coughing are airway clearance. Techniques to facilitate airway clearance are airway clearance techniques (ACT). BPH and ACT include several manual techniques within the physical therapy scope of practice. Just as increasing tissue extensibility through manual techniques may facilitate the range of motion necessary to move during therapeutic exercise to then maintain tissue extensibility; improving ventilation and respiration through manual techniques may facilitate the respiration necessary to move during exercise to then maintain respiration.

Please don't tell me that "respiratory therapy" does this, of course they do. But that does not negate the fact that this is part of the PT scope of practice. If PTs did not do the things in their scope of practice that other professionals do, physical therapists would not do anything. Saying "respiratory therapy" does this - is like saying "massage therapists" do this about soft tissue mobilization; or that "chiropractors" do this about joint mobs/manipulations; or that "athletic trainers" or "OTs" do this about rehabilitation; or that "exercise physiologists" or "personal trainers" or "strength and conditioning specialists" do this about exercise prescriptions. Of course that is all true. Doing one thing is NOT what makes any professional the professional that they are, it is doing those things in context and the variety of things being done.

12.2.2.2 Nasal Passageways

Nasal passageways are lined with epithelial cells rich in olfactory sensors (sense of smell) and goblet cells which are responsible for creating mucus. The passages include a set of turbinates that create turbulent flow of air through the passageways which is important for immunity, humidification, adding nitric oxide and the sense of smell. The turbulent flow of air encourages interaction between the air entering the the nasal passages and the cellular epithelial mucus membrane. Large particles (dust, pathogens, allergens) are very likely to become stuck in the mucus of the nasal passage membranes, serving an important immune and protective function. Interaction with the nasal passage membrane also humidifies the air (adds water) and adds nitric oxide (NO). NO enhances local defense mechanisms via direct inhibition of pathogen growth and stimulation of mucociliary activity. NO from the nose and sinuses is added to the air with every breath taken through the nose. NO reaches the alveoli in a more diluted form and enhances alveolar respiration by local vasodilation of pulmonary capillaries [Tor+02; Lun08]. For these, and many other reasons, nasal breathing is better for ventilation and respiration and promotes more efficient breathing. Nasal breathing tends to encourage and promote further nasal breathing. Mouth breathing tends to make nasal breathing more difficult. If someone that uses mouth breathing attempts to nasal breathe they may need to go through a transition period with intentional nasal breathing to help re-establish air movement through the nasal passageways. Mouth breathing has also been linked to sub-optimal changes to upper airways that ultimately make nose breathing more difficult which then promotes more mouth breathing. Mouth breathing is associated with snoring, sleep apnea and hyperventilation (over breathing) dysfunction [GCB14].

12.2.2.3 Lower Airways

The trachea is the airway that brings air to the many divisions that terminate in alveoli. The trachea is supported by cartilage rings that keep it open during wide fluctuations in thoracic pressures (coughing, sneezing, valsalva manuever²). The trachea divides at the carina into two main bronchi (right and left) for the right and left lungs (respectively) (See Figure 12.4).

The main, and larger bronchi, also contain cartilage rings for support along with smooth muscle. Each bronchi splits into two branches for approximately 20-25 generations until ending with alveoli [HH20]. The final branch of the bronchi is a terminal bronchiole and is sometimes referred to as the respiratory bronchiole (See Figure 12.5). The lower airways are also lined with epithelial cells that produce mucus and have cilia, and thus also have mucociliary activity that contributes to bronchopulmonary hygiene (BPH) and airway clearance.

Similar to the artery to arteriole, each branch of the the bronchi to bronchiole gets progressively smaller, while the sum of space gets larger; and the influence of smooth muscle on the radius gets more important. For example, the main bronchi is larger than any single branch, but the branches, summed together, offer far more space than the main bronchi. As the branches get smaller with each generation of splitting the walls get thinner and the influence of smooth muscle on the walls gets more influential. Most bronchodilation and bronchoconstriction occurs in the smaller bronchioles than in the larger bronchi.

Under normal conditions resistance to air flow is highest in the larger bronchi because there are relatively few bronchi compared to bronchioles. But with various pulmonary conditions the smaller bronchioles play a greater role in determining air flow resistance for two reasons. First, they are easily occluded because of their small size. Second, they are easily constricted because they have a greater proportion of smooth muscle fibers in their walls in relation to their radius and they do not have supportive cartilage rings.

The factors impacting air flow, like blood flow, through a cylinder such as an airway (or blood vessel) can be considered with Pouiselle's Law:

$$Flow = \frac{\pi \Delta P r^4}{\eta L} \quad (12.4)$$

And similar to blood flow, the resistance to airflow is largely determined, and manipulated, by changes to the radius through bronchodilation and bronchoconstriction.

² Valsalva manuever includes forced expiration against a closed epiglottis and the development of substantial intrathoracic, including intrapulmonary, positive pressure. It often occurs during straining for a bowel movement, or during lifting and is thought to provide stability support and counteract compression of the spine when combined with increased abdominal pressure. Risks include hernia, lung collapse (pneumothorax), and alterations to blood flow leading first to elevated and then dropped blood pressure.

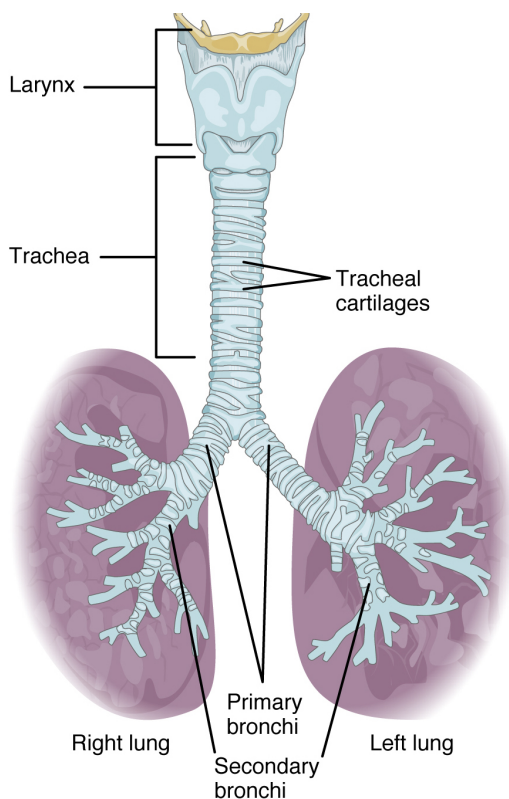


Fig. 12.4: Trachea and Main Bronchi Branches (OpenStax, CCBYSA4.0)

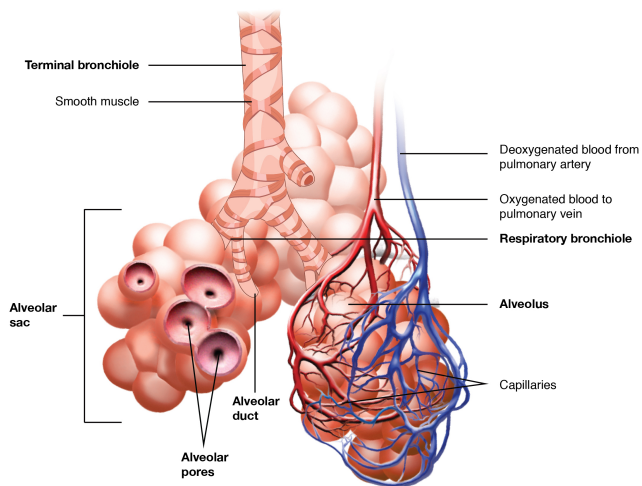


Fig. 12.5: Terminal Bronchiole & Alveoli (OpenStax, CCBYSA4.0)

Bronchodilation & bronchoconstriction are regulated by sympathetic nervous system innervation and endocrine function (circulating epinephrine and norepinephrine). Control of the bronchioles by sympathetic nerve fibers is relatively weak because few nerve fibers penetrate deeply and reach the bronchioles. However, all bronchi and bronchioles are exposed to circulating norepinephrine and epinephrine released from the adrenal medulla. These hormones, and in particular epinephrine because of its greater stimulation of β -adrenergic receptors, dilate bronchi and bronchioles. A few parasympathetic nerve fibers, from the vagus nerve, reach the bronchi and bronchioles. These nerves secrete acetylcholine, which constrict the bronchi and bronchioles. Conditions that create bronchoconstriction, such as asthma, are therefore commonly treated with β -adrenergic agonist inhalers (inhaled medication), sometimes in combination with

cholinergic antagonist inhalers. A side effect of β -adrenergic agonist inhalers is anxiety and rapid heart rate. This side effect would be much more pronounced if the β -adrenergic agonist was taken orally and not with an inhaler because of its impact on the heart.

12.3 Volumes & Capacities

Ventilatory volumes and capacities are depicted in Figure 12.6. Capacities are combinations of volumes.

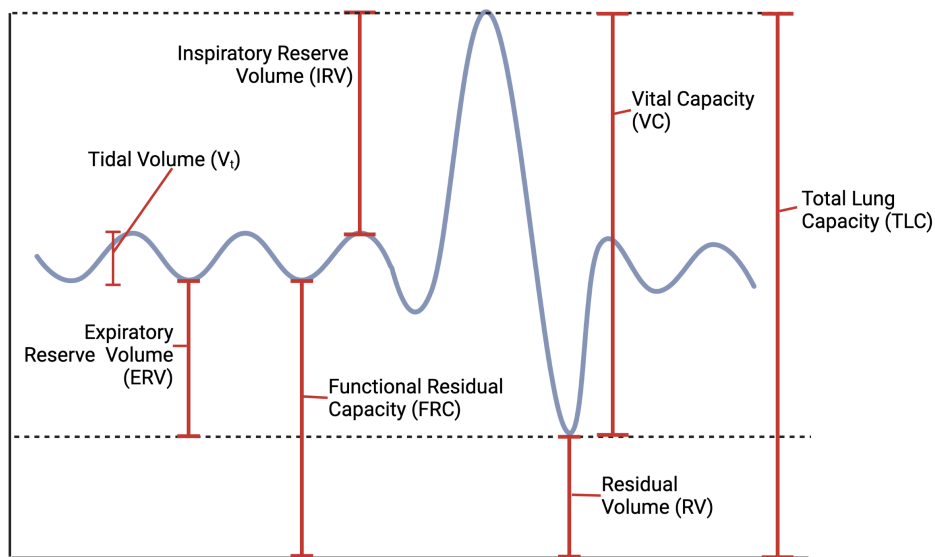


Fig. 12.6: Ventilatory Volumes & Capacities

- Tidal Volume (V_t): Volume of breathing. V_t varies in order to change ventilation V_e , for the purpose of changing V_A to meet the needs of alveolar respiration as it maintains P_{aCO_2} and P_aO_2 ; or adjusts P_{aCO_2} as part of acid base balance regulation.
- Inspiratory Reserve Volume (IRV) is the volume that can be inhaled after a normal V_t inspiration
- Expiratory Reserve Volume (ERV) is the volume that can be exhaled after the end of a normal V_t expiration
- Residual Volume (RV) is the volume that cannot be exhaled
- Vital Capacity (VC) is $V_t + ERV + IRV$, it is the largest volume that can be breathed in after a complete exhalation, or out after a complete inspiration
- Functional Residual Capacity (FRC) is $ERV + RV$, it is the volume of air remaining in the lungs after a normal expiration
- Total Lung Capacity (TLC) is $IRV + V_t + ERV + RV$, it total capacity for air in the lungs

12.3.1 The Musculoskeletal Side of Ventilation

Volumes & capacities are measures of the amount of air that can be ventilated (or that cannot be ventilated in the case of RV which cannot be expired). They are helpful for understanding the functional capabilities of the lungs and describe changes to the lung in various states of disease. For example, with the obstructive pulmonary diseases the RV tends to be higher, which reduces VC (and all other values, except FRC). The volumes & capacities also represent a functional measure of the thoracic cavity musculoskeletal range of motion (chest wall mobility) which includes the diaphragm,

muscles of the ribs (intercostal muscles), the mobility of the ribs, and the posture of the spine. For example, in situations with diaphragm weakness, reductions in *VC* are directly related to an inability of the diaphragm to fully descend against the combined resistance of pulmonary elasticity, thoracic muscles and joints, and abdominal contents. In patients with many different neurological conditions there is involvement of ventilatory muscles or spinal posture which greatly influence the ventilatory volume & capacities that further limit activities and result in pulmonary conditions such as pneumonia.

12.3.2 Volumes & Capacities as range of motion (ROM) - Chest Wall Excursion (CWE)

Occasionally the volumes and capacities are utilized as an indication of the chest wall range of motion (ROM), which is referred to as chest wall excursion (CWE). During testing of the strength of inspiratory muscles patients are asked to expire all their air prior to the test because it is best to test those muscles when they are stretched. Those muscles are stretched when they are in the CWE associated with and at the ROM of, *RV*. When doing a forced expiration, or when training someone to improve their cough, they are asked to take in a full breath, they breathe in until they are at the CWE associated with *TLC*. CWE can also be measured as a circumference, but not without limitations. One limitation is that the CWE with a tape measure is measuring circumference in one plane (even if it is an oblique plane, it is only one plane). Whereas the change in thoracic cavity size occurs in all three planes (all three dimensions). If CWE is measured with a tape measure, it is usually recommended that it is measured in more than one place (for example, above, at and below the xiphoid process). CWE is also assessed qualitatively simply by placing the hands on various locations of the chest wall and assessing the symmetry, amount and quality of movement.

12.3.3 Homeostatic Regulation with Volitional Control

Ventilation is certainly part of an intricate set of homeostatic regulation mechanisms. Yet it is also under volitional control. The volitional control allows speech (vocalization), and it also allows the opportunity to use breathing to manipulate (hack) homeostatic systems that cannot otherwise be regulated directly (e.g. the use of deep breathing to increase parasympathetic nervous activity and lower heart rate). Physical therapists contribute uniquely to problems related to breathing and ventilation because of how much these processes rely on the neuromuscular mechanics of the breathing muscles. Like posture, gait and many movements, the breathing muscles and patterns of ventilation are mostly automatic yet are able to be voluntarily controlled, and therefore manipulated, adapted and trained. Emotional states are related to the autonomic nervous system (bi-directional), and the autonomic nervous system (ANS) is related to breathing (bi-directional). The bi-directional connection between ANS and breathing can be used to influence ANS by voluntarily altering breathing (See Figure 12.7). The downside of this is that breathing dysfunction, even when not due to pulmonary disease but rather caused by altered breathing coordination (mechanics) or patterns, can result in altered emotions. Questionnaires to assess whether breathing coordination (mechanics)(including too much mouth breathing) or patterns may be impacting health include a host of signs and symptoms related to the ANS [GCB14].

12.4 Mechanics: The Ventilatory Pump

The ventilatory pump creates pressure for ventilation with the musculoskeletal system of the thorax (chest wall), including the diaphragm as the floor of the thoracic cavity (and roof of the abdominal cavity, See Figure 12.8). The diaphragm is positioned to contribute to the variation of pressures between the thoracic and abdominal cavities which makes it uniquely involved in breathing as well as the coordination of breathing with other aspects of dynamic core stability.

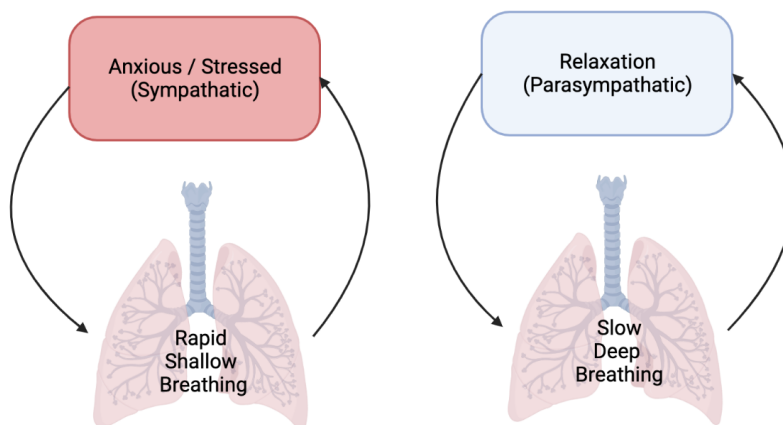


Fig. 12.7: Bi-Directional Homeostatic Feedback Between Emotion and Breathing

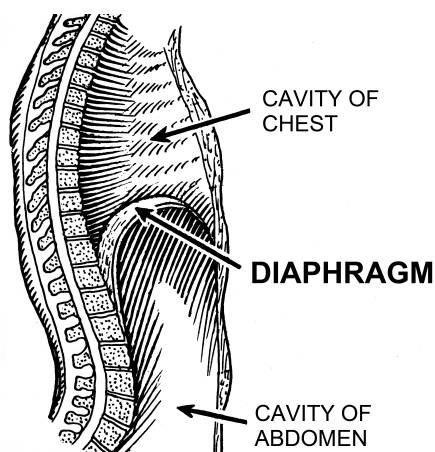


Fig. 12.8: Lateral View of the diaphragm dome and the separation between the thoracic and abdominal cavity (CCBYSA4.0, WikiMedia Commons)

12.4.1 Inspiration: A Negative Pressure Pump

To breathe air into the lungs the ventilatory pump creates negative pressure. For breathing, the terms negative and positive pressure are always being compared with the atmospheric pressure. Negative pressure means less than atmospheric pressure; and positive pressure means greater than atmospheric pressure. When the ventilatory pump generates negative pressure it pulls air into the lungs. It does this by expanding the volume of the thoracic cavity, which lowers the pressure below the atmospheric pressure to establish the necessary pressure gradient for air flow. Based on the familiar equation: $Flow = (P_1 - P_2)/R$; the flow is directly proportional to the pressure gradient created, and inversely proportional to the resistance (and the resistance is largely dependent on the radius of the airways).

12.4.2 Expiration: A Positive Pressure Pump

For breathing air out of the lungs the ventilatory pump is a positive pressure pump. Positive pressure is generated by reducing the volume of the thoracic cavity, which increases the pressure above the atmospheric pressure to establish the necessary pressure gradient for air flow. Air flow is still directly proportional to the pressure gradient, and inversely proportional to resistance. Resistance is still largely dependent on the radius of the airways.

One difference with expiration is that the positive pressure, combined with the fact that during expiration the lungs and airways are getting smaller, it is much more likely that resistance to expiratory flow will impact exhalation (than resistance to inspiratory flow will impact inhalation). The practical implications are that with any condition that impacts airway radius, such as asthma, emphysema and chronic bronchitis, the functional impact is on expiratory flow rates more than inspiratory flow rates. When a limitation is made on expiratory flow rates the term utilized to describe the pattern is obstructive.

Forced expiratory activities - such as airway clearance with sneezing or coughing - requires a rapid increase in positive thoracic pressure. The difference between a sneeze (besides being typically through the nose) and a cough is closure of the epiglottis to allow for a larger build up of positive pressure in the thoracic cavity and a sudden release of that pressure when the epiglottis opens.

12.4.3 Pressure Equilibration

Airflow through airways is rapid enough that equilibration can occur at any thoracic cavity volume. Once equilibration occurs, airflow stops. This means that at any thoracic cavity volume, other than TLC or RV , changes in volume can cause either inspiration or expiration depending on how the changes in volume change the pressure. At this point it's helpful to remind readers of the Ideal Gas Law.

The common form of the Ideal Gas Law is:

$$PV = nRT \quad (12.5)$$

Where, P is pressure; V is volume, n is the number of moles in the gas (the amount of gas), R is a gas constant, and T is the temperature in Kelvin. For our purposes we can simplify by removing R and T (although there are slight changes during ventilation from the temperature of atmospheric air and the warmed air at body temperature, these are relatively small on the Kelvin scale, and since we are not actually calculating just considering relationships, we'll be fine).

A simplified version of the Ideal Gas Law, that solves for P is:

$$P = \frac{n}{V} \quad (12.6)$$

An increase in the thoracic cavity V will reduce the P . The reduced P results in air flow into the lungs through airways until the amount of air in the lungs n raises P until it equals the atmospheric pressure and air stops flowing into the lungs. At that point, if thoracic cavity V can be increased more then P will drop and more air will flow into the lungs until n increases to the point that P in the lungs equilibrates with the atmospheric pressure. The same process can be considered for exhalation. A reduction in thoracic cavity V increases the P , air flows out of the lungs until n is reduced until P equals the atmospheric pressure.

Breathing in, Breathing out is a cyclic variation in the difference between the thoracic (pulmonary) P and the atmospheric pressure, brought on by the cyclic variation in the thoracic (pulmonary) V , which is equilibrated by the movement of air into (increase n) and out of (decrease n) the lungs.

12.4.4 FRC : The Resting Position

When considering the mechanics of the ventilatory pump during breathing it is useful to understand the conditions surrounding the functional residual capacity (FRC). FRC is the resting position (volume) of the thoracic cavity, and therefore the amount of air in the lungs at this position equals the FRC . (That's a nice, intentionally, circular explanation.) FRC is the V of the thoracic cavity when there is no active tension being developed by the breathing muscles. The V of the thoracic cavity can still change based on body position, but it is not changing due to active tension of breathing muscles (either inspiratory or expiratory muscles). If you "stop breathing" - and there is no airflow or active tension in your inspiratory or expiratory muscles - you are at FRC . If you do this in standing and then sit down there may be slightly less V at FRC . If you are seated upright and slouch or lean forward then you reduce the V and FRC . The FRC is the V in the lungs when there is no active attempt to breath, but it is positionally (supine, prone, sitting, standing) and posturally (spine neutral, flexed, extended) dependent.

***FRC* is the balance point** between an inward elastic recoil force of the lung tissue, and an outward elastic recoil force of the musculoskeletal thoracic cavity and is influenced by other forces such as gravity (and the impact that gravity has on the thoracic cavity) (See Figure 12.9). If the lungs were removed from the body (or detached due to a loss in negative pleural pressure), they would be smaller (less V) than at the *FRC*. Without the influence of the lungs, the thoracic cavity would be larger (more V) than at the *FRC*.

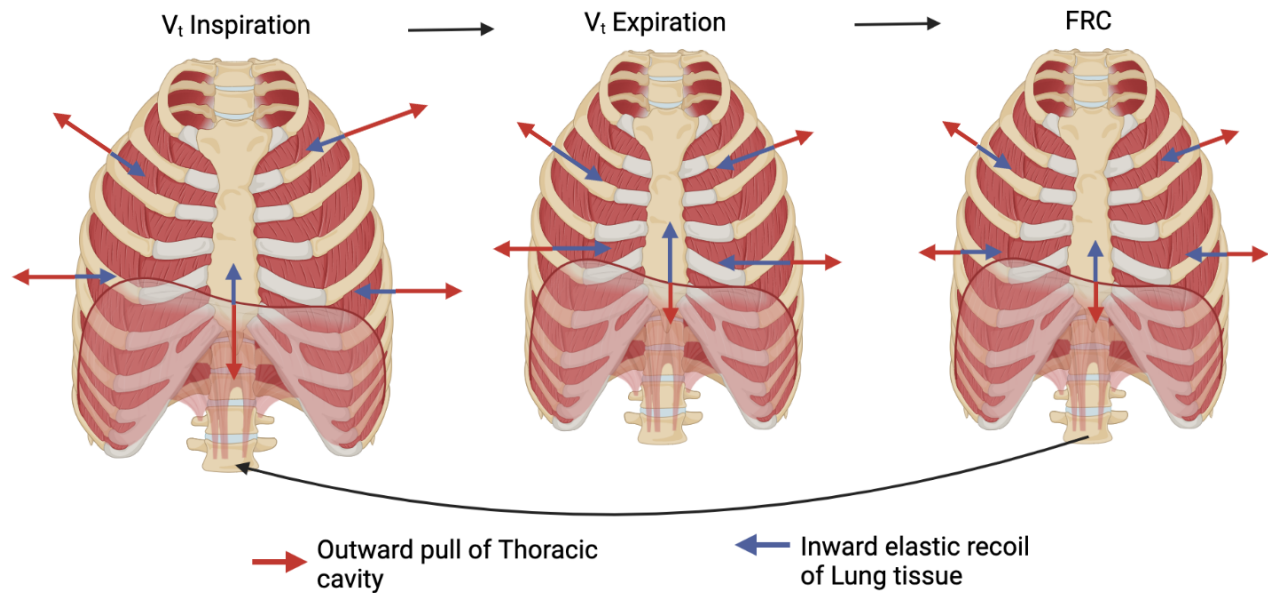


Fig. 12.9: Inward Elastic Recoil of Lung Tissue, Outward Pull of Thoracic Cavity. *FRC* occurs at the end of passive expiration at the balance point of these forces. (Created with BioRender.com)

The balance of forces point of *FRC* has a large influence on breathing patterns because it has a large influence on optimal breathing volumes. During resting tidal volume (V_t) breathing each expiration ends at *FRC* and therefore is a passive process. Another way to say this is that a resting V_t expiration ends at *FRC* so that it can be a passive process. Inspiration requires muscles of inspiration to expand the thoracic cavity V , working against the passive recoil force of the lung tissue. But then the passive recoil force of the lung returns the V to *FRC* which promotes expiration. Therefore, with resting V_t breathing only the muscles of inspiration must be utilized.

During resting V_t breathing, initial increases in thoracic cavity V are easier to perform because the thoracic cavity is naturally larger than at *FRC*. However, with increases in thoracic cavity V both the lungs and the thoracic cavity create a passive elastic force that the muscles of inspiration must oppose for continued inspiration. This is one of the reasons that during exercise, breathing patterns typically increase V_t by using more of the *ERV* as compared to the *IRV*.

Forceful expiration requires a larger and more rapid change in P , which requires a larger and more rapid change in V , which requires active tension in the muscles of expiration. Increasing the rate of breathing *RR* alone, or along with an increase in V_t , requires active tension of the expiratory muscles so that it occurs more completely (volumes less than *FRC*), and more quickly.

12.4.5 Muscles of Inspiration: Increase Thoracic Cavity Volume

12.4.5.1 Diaphragm

The diaphragm is a membranous muscle and is the most important inspiratory muscle, with motor innervation solely from the phrenic nerves (C3-C5). The diaphragm is extremely active. Muscle fibres within the diaphragm can reduce

their length by up to 40% between residual volume and total lung capacity and spend 35% of each day in an active tension state [LT20]. The origins of the crural part of the diaphragm are the lumbar vertebrae and the arcuate ligaments, while the costal parts arise from the lower ribs and sternum. Both parts are inserted into the central tendon (See Figure 12.10).

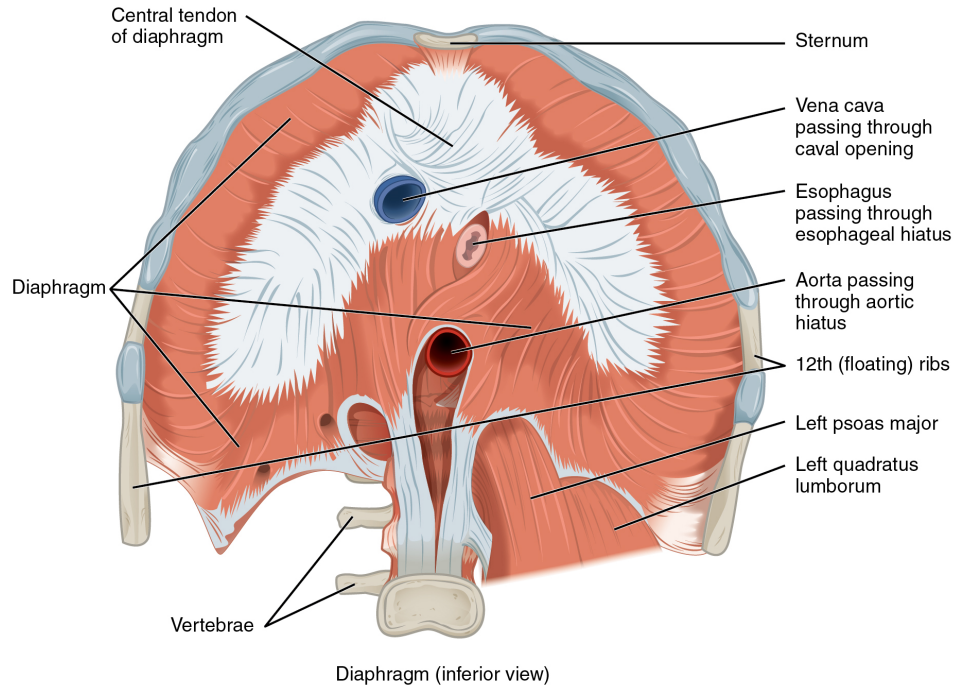


Fig. 12.10: Inferior View of Diaphragm Attachments (OpenStax, CCBYSA4.0)

Under normal circumstances, a zone of apposition exists around the outside of the diaphragm where it is in direct contact with the inside of the rib cage, with no lung in between. The parietal pleura continues to allow movement of the diaphragm. When at the upright *FRC* approximately 55% of the diaphragm surface area is in the zone of apposition (See Figure 12.11).

There are three actions that the diaphragm creates to increase thoracic cavity V and these three actions are combined during breathing. Diaphragm mechanics, and these three actions, can be considered with the analogy to a piston in a cylinder. With the diaphragm being the piston, and the thoracic cavity being the cylinder. A graphic of these three diaphragm actions is depicted in Figure 12.11.

1. **Piston:** Downward movement of the diaphragm simply by shortening the zone of apposition around the whole cylinder and leaving the dome shape unchanged. This action very efficiently converts diaphragm muscle shortening into changes in V .
2. **Non Piston:** Diaphragm tension reduces the curvature and therefore the dome shape. This does not change the zone of apposition but results in an, albeit smaller, increases in V .
3. **Expand the Cylinder:** While the diaphragm is shortening it is also pulling upward on the ribs. This upward pull on the lower rib cage results in expansion of the lower thorax and therefore thoracic cavity V . This contributes to the "bucket handle" movement of the lower rib cage during breathing.

There are many factors that influence how the above three actions that diaphragm mechanics contribute to inspiration. What is clear is that the larger the inspiratory volume, the more diaphragm activity is required and the likelihood that all three mechanical approaches are involved. The most important factors influencing these mechanics are activity (which influences the amount of air that must be ventilated (V_e)); position; and posture. For example, in supine the "Expand the Cylinder" approach tends to dominate because abdominal mass reduces downward movement of the diaphragm which increases upward pull on the lower ribs.

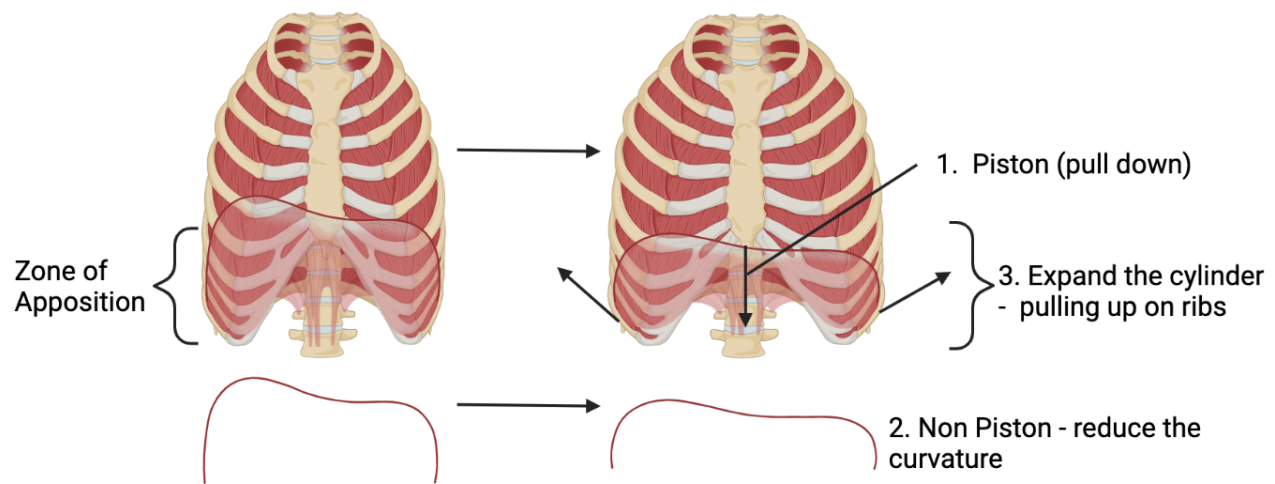


Fig. 12.11: Three Actions of the Diaphragm to Expand the Thoracic Cavity (Created in BioRender.com)

12.4.5.2 Intercostals & Scalenes

The exact coordination and involvement of the external and internal intercostals during breathing is complicated and most descriptions (including this one) is a simplification [LT20; DKW05]. There are several contradictory sources about which set of intercostals are involved with inspiration vs. expiration [LT20; HH20]. To avoid those contradictions, and since there is little practical value to those arguments, we simply consider the intercostals and do not attempt to assign a particular breathing role to the external or internal sets. We rather consider both the internal and external as contributing to either lifting the ribs (for inspiration) or lowering the ribs (for expiration) from their axes at the costotransverse and costovertebral joints in the thoracic spine. We consider this based on motor coordination patterns [GCB14].

The inspiration motor coordination pattern for the intercostals starts with the activation of the scalenes. Scalenes are active even during quiet breathing. They stabilize the first rib. The first rib is then the anchor for each rib below it, in succession, to pull upward on the rib below it. They support, or pull up, rib 1. From this base of support the intercostals between rib 1 and 2 stabilize, or pull up, rib 2. From the base of support at rib 2, the intercostals between rib 2 and 3 stabilize, or pull up, rib 3. The pattern continues to rib 12 - the stabilized rib n , allows the intercostals between rib n and $n+1$ to stabilize, or pull up, rib $n+1$.

Based on the costotransverse and costovertebral joint axes, and in combination with attachment to the sternum, the upward movement of the upper ribs (1 - 6 or so), is described as "pump handle" which primarily includes a change in the anteroposterior (AP) diameter of the thoracic cavity. The upward movement of the lower ribs (7ish through 12) is described as "bucket handle" which primarily includes a change in the medial-lateral (ML) diameter of the thoracic cavity.

Even during quiet breathing there is active tension generated in the scalenes and intercostals, even if there is no movement of the ribs. These isometric contributions prevent inward movement of the ribs due to the drop in thoracic pressure by the diaphragm.

! Thoracic Paradox

In premature infants with underdeveloped intercostal muscles, or people with C5 to C8 spinal cord injury that have innervation of the diaphragm but do not have innervation of the intercostal muscles, thoracic paradox breathing mechanics can impair ventilation. Thoracic paradox includes a diaphragm piston like activation that reduces thoracic pressure and pulls the ribs inward which reduces the change in thoracic V and limits ventilation. It is called thoracic paradox because the thorax is not doing what is expected during inspiration.

12.4.5.3 Accessory Inspiratory Muscles

Accessory muscles are not active during normal breathing, but as ventilation increases there is a greater need to increase the thoracic cavity size, and therefore greater chest wall excursion (CWE).

Increasing the dimensions of the thoracic cavity requires more tension and range of contraction in the inspiratory muscles and can include recruitment of accessory muscles. Other reasons for accessory muscle activation may be when there are limitations in the primary muscles or when there are pulmonary conditions making ventilation less efficient. Accessory muscles include the sternocleidomastoids, extensors of the vertebral column, pectoralis minor (and to a lesser extent major), trapezius and the serratus anterior and posterior muscles. For many of these muscles to serve an accessory role to inspiration the upper extremity must be stabilized so that tension in the muscle moves the thorax, not the limbs. In people with severe pulmonary disease that rely on accessory muscle activation for breathing, activities that use the upper extremities can be considerably limited.

Spinal Extension increases thoracic cavity size. Think of the rib cage as an accordion. With spinal extension the accordion opens. Spinal extension (and the spinal extensors) is therefore an inspiratory accessory movement (and therefore spinal extensors are accessory muscles).

12.4.6 Muscles of Expiration: Decrease Thoracic Cavity Volume

12.4.6.1 Intercostals

The motor coordination pattern for using the intercostals to decrease thoracic cavity volume starts with the abdominal muscles (rectus abdominus, external and internal obliques). The abdominal muscles stabilize the lower rib cage so that intercostal contraction pulls down on the rib above it to reduce the size of the thoracic cavity (cylinder). This is the opposite coordination pattern of inspiration when the scalenes stabilized the first rib.

12.4.6.2 Abdominal Muscles

The abdominal muscles have three roles in expiration (particularly forced expiration).

1. Stabilize the lower ribs so that intercostal activation pulls ribs down.
2. Pull the lower ribs down (they pull the lower ribs down much more than the scalenes can pull the first rib up - so the role of the scalenes for inspiration, while important, is really limited to stabilizing that first rib). The pulling down on the lower ribs significantly contributes to reducing the size of the thoracic cavity (the cylinder) and therefore the abdominals are a primary muscle of forced expiration.
3. Reduce abdominal cavity volume → increases abdominal cavity pressure → pushes up the relaxed diaphragm → reduces thoracic cavity volume and resets the piston (stretches the diaphragm for the next inspiration)

12.4.6.3 Accessory Expiratory Movement

As you probably figured out, if spinal extension is an accessory inspiratory movement then spinal flexion is an accessory expiratory movement. Of course, the abdominal muscles (collectively) are the spinal flexors and they are primary expiratory muscles, so they are not also accessory expiratory muscles. It is simply the case that if spinal flexion accompanies expiration, either intentionally to get as much air out as possible, or unintentionally due to the force of the abdominal muscles during expiration, then it is an expiratory accessory movement. Note: forced expiration that must include abdominal muscle activation but does not include spinal flexion, must involve spinal extensors to generate an equal, but opposite torque so that the spinal flexor muscles perform the three tasks they perform for expiration, but do not flex the spine.

12.4.7 Pulmonary & Pleural Pressures

The discussion about the mechanics of ventilation have rightly assumed coherence between the lungs and the thoracic cavity. When the thoracic cavity increases volume, the lungs increase volume which increases the volume of the alveoli. When the thoracic cavity decreases its volume, the lungs decrease volume which decreases the volume of the alveoli. Under all normal circumstances these events are true.

But.... anatomically there is no direct connection between the lungs and the inner wall of the thoracic cavity. There is a space, it's a small space and it's filled with fluid, but it's a space. It is the pleural space and it is filled with pleural fluid. The lungs are lined with a lung pleural membrane and the thoracic cavity is lined with a thoracic pleural membrane. Between them is a thin space with pleural fluid. This arrangement allows the lungs to move within the thoracic cavity with relatively little friction (similar to the pericardial space and the pericardial fluid around the heart).

To keep the lungs inflated and moving along with the thoracic cavity the pleural space is kept at a slightly negative pressure (negative in comparison to the atmosphere, and the atmosphere is taken as 0 cm of water (cmH_2O)³). This is called the intrapleural pressure. The normal intrapleural pressure at the beginning of inspiration is about -4mmHg , which is the amount of suction required to hold the lungs at their resting volume. During inspiration the thoracic cavity expands and the intrapleural pressure drops to approximately -7.5mmHg . These values are being shared simply to convey that they are slightly negative.

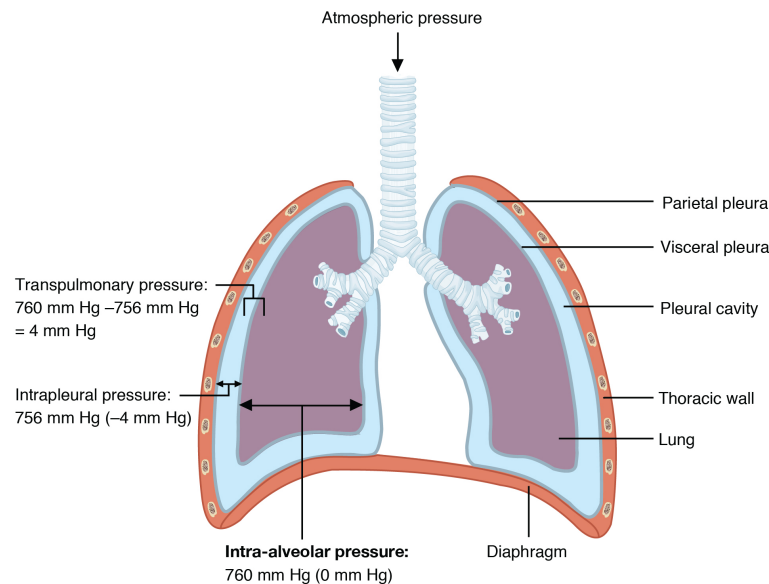


Fig. 12.12: Intra Pulmonary Pressures (OpenStax, CCBYSA4.0)

During regular (resting V_T) expiration the lungs recoil and pull the thoracic cavity thus maintaining the negative pressure of the pleural space. Due to the movement of the thoracic cavity along with the lungs the intrapleural pressure returns to -4mmHg .

During forced expiration the intrapleural pressures become slightly positive as the thoracic cavity pushes on the lungs to reduce their volume. The positive intrapleural pressure can result in a greater degree volume loss in alveoli, and reduction in the size of bronchioles and bronchi than due to the reduction in the volume of the thoracic cavity. In other words, during forced expiration the thoracic cavity directly reduces the lung size and this reduction is further influenced by a positive intrapleural pressure that also decreases lung airways and alveoli volume. This effect is more pronounced with a greater force for forced expiration because there is a higher positive intrapleural pressure.

³ In ventilation physiology it is common to simply refer to the atmosphere as 0 which is a relative measurement based on the fact that when not breathing the atmosphere and alveoli pressures equilibrate and that for ventilation it is only differences in these pressures that matter. However, when calculating the partial pressures of gases, we consider the absolute atmospheric pressure which, at sea level, is 760 mmHg ($1033\text{ cmH}_2\text{O}$)

! Pneumothorax

Pneumothorax is the collapse of a lung which can be precipitated by a loss of negative pleural pressure due to trauma. In certain trauma scenarios air may escape the lung and enter the pleural cavity resulting in positive pleural pressure which collapses the lung. This situation is called a tension pneumothorax and requires the positive pleural pressure to be removed with a tube placed into the chest (chest tube). Chest tubes are commonly used after thoracic surgery (any surgery that opens the thoracic cavity) to prevent the accumulation of edema and also to maintain negative pressure via suction to keep the lungs inflated.

! Pleural Effusion

Pleural Effusion is the accumulation of fluid in the pleural space which then compresses the lung.

12.4.8 Lung Compliance

Lung Compliance is the change in lung volume for each unit of change in transpulmonary pressure. Transpulmonary pressure is the difference between the alveolar and pleural pressures. For example (and not to memorize), the normal total compliance of both lungs together in a healthy adult is about $200 \text{ mL/cmH}_2\text{O}$ (for each $1 \text{ cmH}_2\text{O}$ change in pressure there would be a 200 mL change in volume). Lung compliance depends on the elastic recoil forces of the lung. The elastic recoil forces are determined by the elastin and collagen fibers and the surface tension of alveoli. Increased elastic recoil forces (including increased surface tension) decreases compliance. For example, if there is more elastic recoil the compliance may be $100 \text{ mL/cmH}_2\text{O}$ which means only 100 mL of volume for a $1 \text{ cmH}_2\text{O}$ change in pressure).

12.4.8.1 Surface tension of the alveoli

Surface tension of the alveoli is the tendency for the H_2O molecules on the surface to attract each other and contract (reduce the volume) of the alveoli. This force is strong enough that it can result in the collapse of the alveoli. Collapse of alveoli is called **atelectasis**, and one of the primary reasons patients are encouraged to do deep breathing and start moving around early after surgery (which keeps PTs in acute care hospitals busy) is to prevent, or reverse, atelectasis.

12.4.8.2 Surfactant

Surfactant is secreted by alveolar epithelial cells and reduces the surface tension in the alveoli substantially (to about $\frac{1}{12}$ the surface tension without surfactant). This substantial reduction in surface tension helps to maintain normal lung compliance, and to keep alveoli open. Since there is still some surface tension alveoli have a tendency to get smaller over time (hours). As alveoli get smaller there is less surface area which leads to increased surface tension because surfactant is pushed out of the alveoli. This reduces compliance and increases the risk of collapse (atelectasis). A protective mechanism against this is an occasional deep breath (cycles in inspiratory volume) open the alveoli which allows surfactant to move back into the alveoli. Following surgery with general anaesthesia there is a reduced drive to breath which lowers breathing volume and reduces the occasional deep breath from occurring (yawns, sighs, or just from moving around). This places individuals at a higher risk of atelectasis, which causes problems with oxygenation and also can be a precursor to pneumonia. Another effect is that clearing the anaesthesia from the body is facilitated by respiration and ventilation, so impaired respiration slows down clearance of the anaesthetic that is impairing the drive to breathe more deeply in the first place. Therefore, it is common for post operative rooms and surgical units to regularly encourage deep breathing as patients recover from surgery with exercises and even biofeedback such as an incentive spirometer (provides visual feedback about how much volume has been inhaled).

! Neonatal Respiratory Distress Syndrome (NRDS)

Neonatal Respiratory Distress Syndrome (NRDS) occurs in premature infants that have not yet started to create and release surfactant. They have very low lung compliance and require mechanical ventilation for breathing. But even the mechanical ventilation is challenging because as the ventilator pushes a sufficient volume of air into the lungs for ventilation the pressure rises substantially, because of the low compliance, which can then cause pressure trauma (baro-trauma) and worsen the overall pulmonary situation with lung inflammation. The availability of and delivery technologies for artificial surfactant has significantly improved the outcomes in such situations.

12.5 Ventilation / Perfusion Matching

Ventilation serves respiration. This is not the only function of ventilation, but it is certainly a primary function. For ventilation to serve respiration it must result in alveolar ventilation, V_A . And most of that fresh V_A must match pulmonary capillary perfusion (Q). V/Q matching requires minimizing the situations that result in mismatching. There are two general situations that result in mismatching, dead space ventilation (V_D) and right-left pulmonary shunt (RL shunt). Minimizing these situations involves the pulmonary circulation regulation discussed in Chapter 11 on Respiration in the section on the Regulation of Pulmonary Circulation. This is a good time to review that section of the Respiration Chapter if needed.

12.5.1 Dead Space Ventilation (V_D)

As introduced in the Ventilation Overview section, there are three classifications for dead space: anatomical, alveolar and physiological. To summarize:

- Anatomical dead space ventilation (V_{Dana}) is air ventilated but that does not make it into alveoli, $V_{Dana} = V_e - V_A$
- Alveolar dead space ventilation (V_{Dalv}) occurs in alveoli that receive ventilation but do not receive alveolar circulation
- Physiological dead space ventilation (V_D) is the sum of anatomical and physiological dead space, $V_D = V_{Dana} + V_{Dalv}$

Anatomical dead space ventilation (V_{Dana}) was discussed in the section on Airways, since it is the volume of air that is ventilated but remains in the conductive zone of the airways (does not make it to the alveoli). V_{Dana} is minimized by regulating, and optimizing, the breathing pattern volume and rate. Larger volumes certainly reduce V_{Dana} , but if volumes are too large then the work of breathing becomes more difficult for three reasons:

1. Inspiratory muscles shorten and it's more difficult to generate the required active tension based on their length - tension curve.
2. The thoracic cavity increases and starts to resist further increases due to passive tension so the muscles must work harder (at a time that it is harder for them to work based on #1)
3. Lung compliance starts to decrease because as lungs expand the lung elastic recoil forces (like passive tension) increase (and recall that lung elastic forces are inversely proportional to compliance)

For all of those reasons - optimizing breathing pattern requires a balance between higher volumes and lower rates, which minimize V_{Dana} but increase the work of breathing; and lower volumes and higher rates, which maximize V_{Dana} , but lower the work of breathing.

Alveolar dead space ventilation (V_{Dalv}) is minimized by the regulation of pulmonary circulation. Pulmonary arterioles dilate in areas with increased O_2 and constrict in areas with decreased O_2 . They also dilate in alveoli receiving fresh ventilation taken in through the nose that has had nitric oxide (NO) since NO is a vasodilator.

12.5.1.1 Perfusion Zones

exist throughout the lung because gravity, and therefore body position, influences distribution of V_{Dalv} and perfusion. Furthermore, the distribution of alveolar residual volume influences perfusion as described below. The lungs can be broken into three perfusion zones that are based on gravity (See Figure 12.13). The zones are:

1. Upper Zone: receives less new V_A because during expiration the alveoli do not decrease in size as much, so there is less overall change in their volume and less overall change in air ventilated; receives less Q because the pulmonary circulation is a lower pressure system so its arterial flow is influenced more by gravity (harder to pump blood up) as well as the residual volume of the alveoli. The Upper Zone has more of the FRC and RV . With a higher residual volume of air in the alveoli the pulmonary capillaries are compressed (in addition to the constriction of the arterioles due to lower O_2 delivery).
2. Middle Zone: receives more V_A than the Upper, but less than the Lower Zones; and more Q than the upper, but less than the Lower. There is a lower alveoli residual volume than in the Upper zone, and more than in the Lower zone.
3. Lower Zone: receives more new V_A because during expiration the alveoli decrease in size substantially so there is more overall change in their volume and more overall ventilation; receives more Q because the pulmonary circulation is influenced by gravity. The Lower Zone has less of the FRC and RV . With less residual volume the pulmonary capillaries are not compressed, and in combination with the dilated arterioles (due to the higher O_2 values from the increased ventilation) Q is greater than the other zones.

The combination of pulmonary circulation regulation and pulmonary zones whereby gravity influences V and Q results in normal resting V/Q matching of approximately 80%. This is an empirical figure based on the use of a ventilation/perfusion lung scan, also called a V/Q scan. The imaging uses medical isotopes to evaluate the circulation of air and blood within the lungs to determine the ventilation/perfusion matching.

! Pulmonary Embolism

Pulmonary Embolism is a blood clot that travels to and lodges in a pulmonary artery or arteriole that cuts off blood flow to a section of alveoli capillary circulation. This is a life threatening situation that immediately and substantially causes alveolar dead space and thus **physiological dead space**, reducing the V/Q matching, reducing respiration and therefore P_aO_2 .

Physiological Dead Space (V_D) is the sum of anatomical and alveolar dead space and can be estimated with the Bohr Equation:

$$V_D = V_t \times \frac{(P_aCO_2 - etCO_2)}{P_aCO_2} \quad (12.7)$$

- $etCO_2$ is the end tidal partial pressure of CO_2 measured at the nose at the end of expiration

> For example

If V_t is 500 mL, P_aCO_2 is 40 mmHg, and $etCO_2$ is 35 mmHg then V_D is estimated as 62.5 mL. The units are in mL because the V_t is in mL and it is multiplied by a ratio which is unitless since the P_aCO_2 and the $etCO_2$ are both in mmHg.

? Questions

Figure out the V_D if V_t is 300 mL, P_aCO_2 is 50 mmHg, and $etCO_2$ is 30 mmHg. How does this compare with the above example?

Figure out the V_D if V_t is 300 mL, P_aCO_2 is 40 mmHg, and $etCO_2$ is 15 mmHg. How does this compare with the above examples?

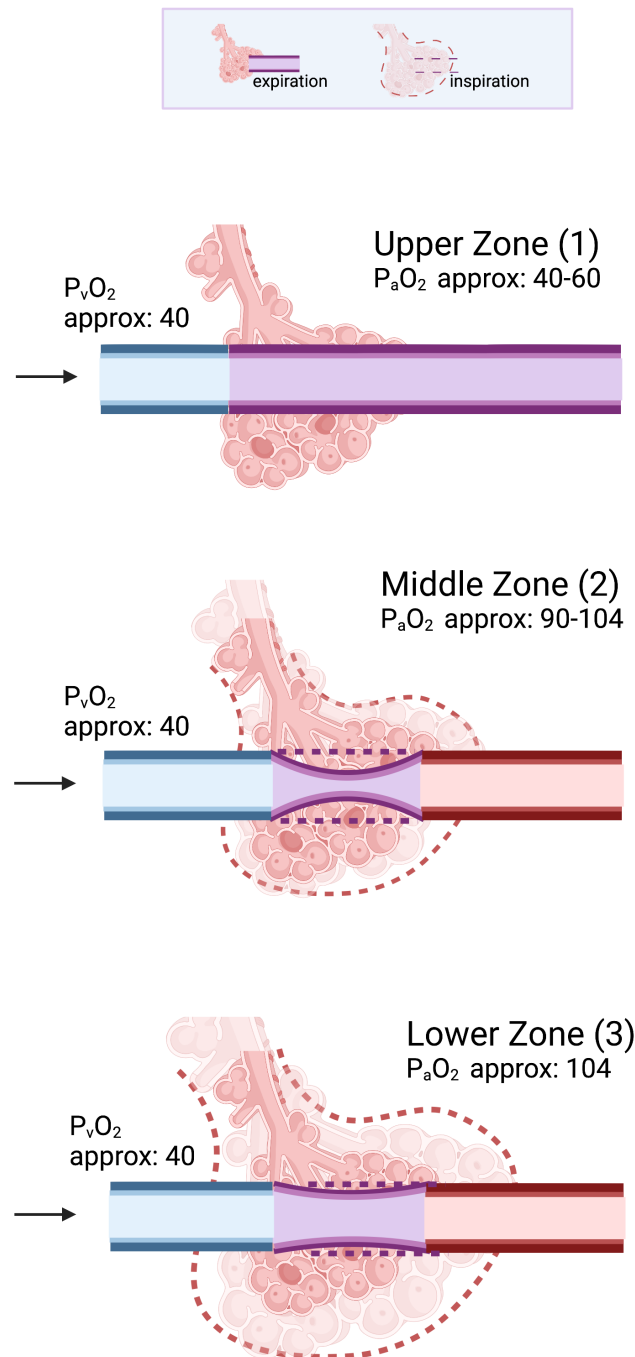


Fig. 12.13: Perfusion Zones
 (Created with Biorender.com by Ashley Ney, SPT, PSU DPT Class of 2025)

12.5.2 Right-Left Shunt

Another situation that can result a V/Q mismatch is a right-left (RL) shunt. When alveolar dead space causes a mismatch it is because ventilated alveoli do not receive any (or enough) perfusion. A RL shunt is the opposite situation. Alveoli

that do not have ventilation, or are under ventilated, receive blood flow. Because this blood flow does not exchange gases with the alveoli (because the alveoli does not have ventilation), the right sided blood (venous blood) returns unchanged and joins with the left sided blood (arterial blood).

The simplest cause of a RL shunt is not breathing (apnea). With apnea the alveoli still have air. However, since the alveoli equilibrate with blood passing through they quickly, without V_A , have the same amount of O_2 and CO_2 as venous blood. This results in elevated P_aCO_2 and reduced P_aO_2 that worsen with prolonged apnea. While holding your breath is apnea, we should reserve the term apnea for pathological causes of not breathing. Voluntary breath hold to the point of lowering P_aO_2 is very difficult so is not as good of an example of a RL shunt.

Pathological causes of RL shunts include anything that results in hypoventilation (reduced V_A), including conditions that fill the alveoli with something other than ventilated air. For example, pulmonary edema, mucus from pneumonia, inflammation with any pulmonary inflammatory condition such as COVID (which can also cause pneumonia). Chronic pulmonary conditions such as emphysema and bronchitis increase the RV which increases the number of alveoli getting blood flow but not new ventilation, a RL shunt.

> V_D & RL Shunt Thought Experiment

If all ventilation went to the right lung; and all blood flow went to the left lung the overall V/Q matching would be 0%. The right lung is completely dead space ventilation V_D ; and the left lung is a complete RL shunt.

12.5.3 A-a Gradient

The A-a gradient is the drop off that occurs between the partial pressure of O_2 from fully ventilated alveoli (P_AO_2) to the partial pressure of O_2 in mixed arterial blood (P_aO_2). The P_AO_2 , at sea level with a normal fraction of inspired oxygen ($F_iO_2 = 0.21$) is taken as 104 mmHg. This blood mixes with blood from other regions of the lung as it travels back to the left atrium. In addition to the venous blood from the bronchial pulmonary circulation, there is blood from normally occurring RL shunts given the normal resting V/Q matching of 80% (20% mismatching). Therefore, we should not be surprised that the normal P_aO_2 is less than 104 mmHg of P_AO_2 . This difference (approximately 104 - 95 = 9 mmHg) is the A-a gradient. The A-a gradient is proportional to the severity of a RL shunt (See Figure 12.14). In critical care situations the A-a gradient can be elevated even when a patient has a normal blood oxygen level because the P_AO_2 is elevated. For example, if a patient at sea level is on supplemental oxygen with an $F_iO_2 = 0.6$ (60% oxygen), then the estimated P_AO_2 is 300 mmHg, and the A-a gradient even with a normal P_aO_2 would be 205, indicating substantial RL shunting.

12.6 Regulation

Ventilation, including breathing volume, pattern and coordination, usually occurs without thought. However, the volume, pattern and coordination can be consciously controlled. Ventilation has automatic but voluntary control - like posture and gait and any learned movement. Ventilation (breathing) is the first movement people learn. Like any movements (or postures), there are many factors that can alter breathing from its optimal performance.

- **Ventilation volume** (V_e) is of volume ventilated in a time period (usually a minute).
- **Ventilation pattern** is the particular combination of tidal volume and respiratory rate to achieve V_e . Since $V_e = RR \times V_t$, where RR is the rate in breaths per minute and V_t is the tidal volume in mL , there are numerous combinations possible to achieve the same V_e . The combination of RR and V_t is the ventilation (breathing) pattern (also the pattern of breathing).
- **Ventilation coordination**, also referred to as **breathing mechanics**, refers to the various combinations of coordinated muscle activation, or the various mechanical changes to the thoracic cavity that different coordination of muscle activation produces. A major determinant of ventilation coordination (breathing mechanics) is how much diaphragm activation is involved with ventilation. More diaphragm activation results in "belly breathing", whereas less diaphragm results in "chest breathing", by necessity and activation of the intercostals and possibly the accessory muscles. Each

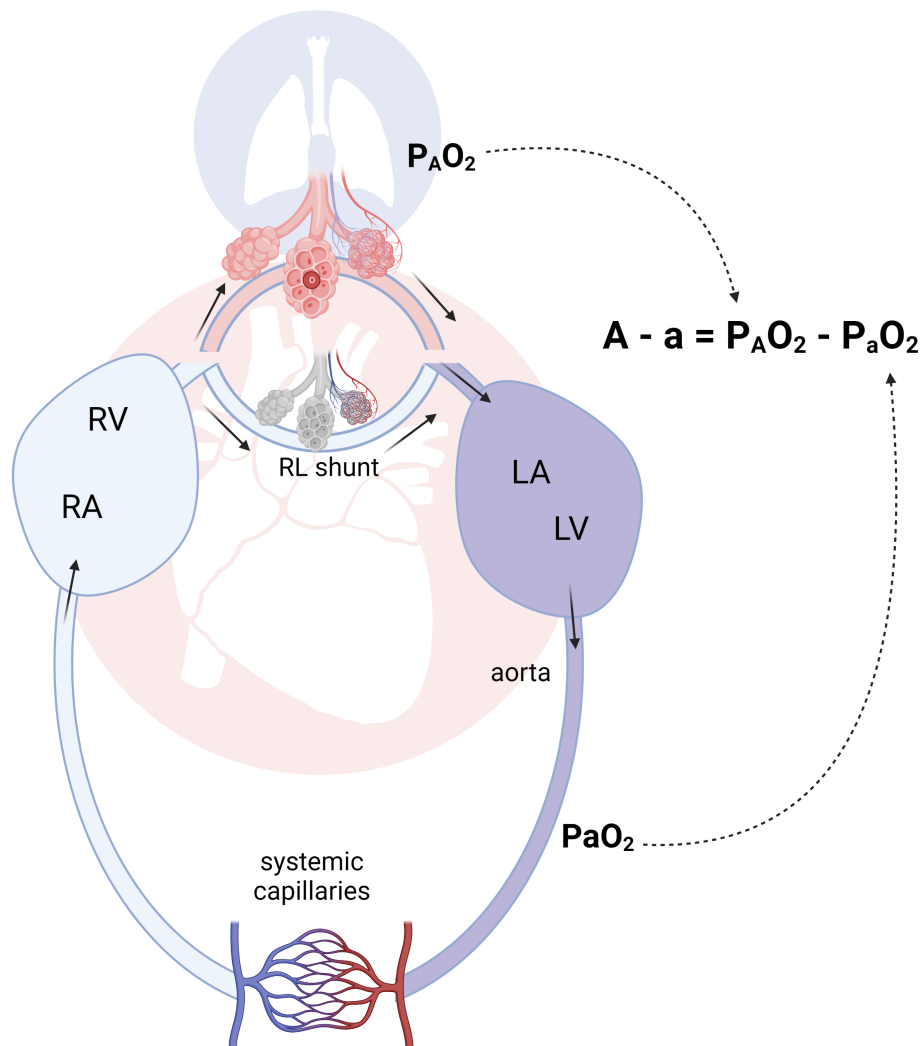


Fig. 12.14: RL Shunt & A-a Gradient (Created with BiRender.com by Ashley Ney, DPT student, Class of 2025)

of these ventilation coordination (breathing mechanics) can produce the same RR and V_t , though chest breathing has a limit on the range of V_t .

In Figure 12.15 the relationship between V_t (mL) (on the x-axis) and V_e (mL/min) (on the y-axis) with three different RR s is depicted. Each V_t , RR combination is a different breathing pattern. Since breathing mechanics that utilize the diaphragm allow a larger V_t , those mechanics are capable of the full range of the V_t (x-axis). When not using the diaphragm there is a limit to the V_t (x-axis). In Figure 12.15 the limit to V_t when not using the diaphragm is taken as 300 mL (dotted orange vertical line). If the metabolic needs of the body require V_e of 4000 mL/min (yellow horizontal line) then diaphragm breathing mechanics can achieve this with a lower RR and greater V_t . However, chest breathing mechanics limits the possible breathing patterns to those with a V_t of less than 300 mL, which requires a greater RR .

? Challenge Question

In Figure 12.15 the need for V_e is taken as 4000 mL (highlighted line). In this example, that volume is based on an assumed metabolic need. Since the metabolic need is actually met by O_2 that gets into alveoli (V_A), and since V_A is influenced by anatomical dead space (V_{Dana}), and since anatomical dead space is related to the breathing pattern (combination of RR and V_t), would the actual V_e need be the same between the $RR = 6$ vs. the $RR = 15$ pattern? How

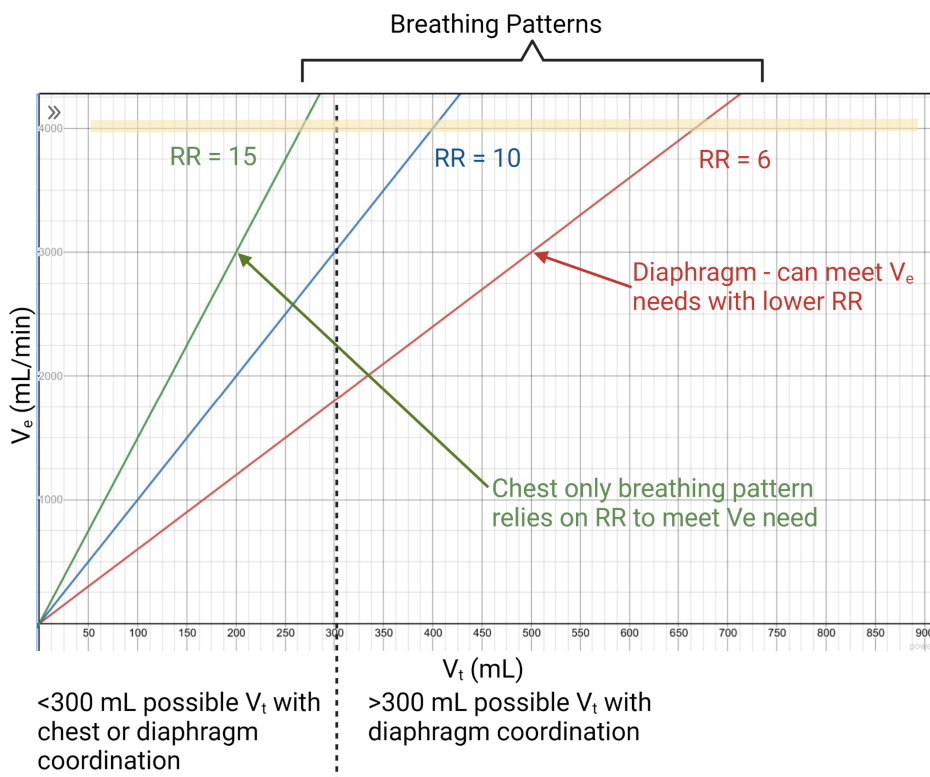


Fig. 12.15: Breathing Pattern Combinations at RR = 6, 10 or 15

would anatomical dead space factor in and change the model assumptions in Figure 12.15? Hint: Consider your answer to the question in the Airway Section on Anatomical Dead Space.

12.6.1 Ventilation Control Centers

Neurons that innervate the muscles of the respiratory system are responsible for controlling and regulating pulmonary ventilation. The major brain centers involved in pulmonary ventilation are the medulla oblongata (medullary respiratory center) and the pons (pneumotaxic center). The medullary respiratory center is responsible for the rhythm of breathing and automatic ventilation. The pneumotaxic center controls the rate and depth of ventilation. The cerebral cortex sends impulses directly to the motor neurons of the ventilatory muscles which allows for voluntary control of ventilation.

12.6.2 Factors That Affect Ventilation

The medullary respiratory center and the pneumotaxic center adjust V_t and RR in response to afferent input. The overall response patterns are dose-response, negative-feedback. The greater the stimulus, the greater the response. Increasing stimuli results in higher RR and greater volume with V_t . These both typically include forced expiration. Multiple factors are involved in stimulating the centers to adjust. Major factors are the concentration of carbon dioxide in the arterial blood (P_aCO_2), and the pH, as sensed by chemoreceptors. A central chemoreceptor is one of the specialized receptors that are located in the brain and brainstem, and a peripheral chemoreceptor is one of the specialized receptors located in the carotid arteries and aortic arch. Increased P_aCO_2 and decreased pH (acidity), stimulate the respiration centers

to increase V_e through some combination of increasing RR and V_t , with some combination of muscles (coordination, mechanics). Reduced P_aCO_2 and increased pH (alkaline), reduces stimulation and therefore reduces V_e . There are chemoreceptors for O_2 , and with hypoxia, P_aO_2 of less than 60 mmHg, these receptors are activated. The lower the P_aO_2 the stronger the stimulus. However, under normal circumstances the P_aO_2 is so far above the threshold for O_2 chemoreceptor stimulation they are rarely involved in the regulation of ventilation. If ventilation was reduced enough for hypoxia, then the increase in P_aCO_2 would be substantial enough to have strongly stimulated the CO_2 chemoreceptors to increase ventilation. To voluntarily allow hypoxia through breath hold, a period of hyperventilation that substantially reduces P_aCO_2 is necessary. This allows prolonged breath holding and enough time for P_aO_2 to drop. That hyperventilation, hypocarbia (low P_aCO_2) situation, is the underlying principle for what is popularly referred to as "Wim Hof Breathing." It can be used to create therapeutic exposures to intermittent hypoxia.

12.6.3 Emotional & Feed Forward Regulation

The hypothalamus and other brain regions associated with the limbic system also influence breathing by interacting with the respiratory centers. These regions are responsible for changes in ventilation related in response to fight-or-flight, emotions, pain, and temperature. For example, an increase in body temperature causes an increase in respiratory rate. Feeling excited or the fight-or-flight response results in an increase in V_e . These responses, particular the fight-or-flight response act as feed forward regulation. Ventilation increases prior to an increase in P_aCO_2 with exercise because before P_aCO_2 has a chance to increase, ventilation has already increased.

12.6.4 Protective Reflexes

Protective reflexes such as coughing, sneezing, bronchoconstriction, mucus secretion and mucociliary activity respond to irritants such as smoke, dust or larger particles (larger than O_2 and CO_2). These coordinated reflexes work together to capture and expel irritants and potentially harmful pathogens.

! CO_2 Retainers - O_2 Drive for Breathing

The combination of pulmonary diseases classified as chronic obstructive pulmonary disease (COPD) are progressive chronic conditions. In severe to end-stage COPD, among other things, there is chronically elevated P_aCO_2 that creates respiratory acidosis and metabolic compensation. The chronically elevated CO_2 and reduced pH can result in down regulation of the chemoreceptors to these stimuli. In these situations, P_aO_2 chemoreceptors are utilized to regulate V_e . A consequence of this situation is that P_aO_2 must be near hypoxia to increase ventilation. If someone in this situation is provided with too much supplemental O_2 , which increases their F_iO_2 , then their ventilation will decrease and their P_aCO_2 will increase further, to potentially dangerously toxic levels.

12.7 Practice Connections

12.7.1 Pulmonary Function Tests - Forced Expiratory Maneuver

The Forced Expiratory Maneuver (FEM) of Pulmonary Function Tests (PFTs) includes a full inspiration (until reaching total lung capacity (TLC)), followed by a forced expiration to expire as much air as fast as possible with the goal of expiring air until reaching the residual volume (RV). Subjects are usually seated for testing, but they are allowed to extend their spine during the full inspiration and flex their spine during the full, forced expiration. There are several measurements that are made from this single test. But, the two most important measurements for the physical therapist to recognize from an FEM of a PFT are the forced vital capacity (FVC) and the forced expiratory volume in 1 second (FEV_1). They are both measured in liters (L). Forced vital capacity (FVC) is an estimate of the vital capacity VC and

with healthy lungs $FVC = VC$. Forced expiratory volume in 1 second (FEV_1) is influenced by both expiratory volume and expiratory flow. The ratio of these two measures, the FEV_1/FVC considers whether a reduction in FEV_1 is due to a reduction in FVC (volume) or expiratory flow. When expiratory flow is normal: $FEV_1/FVC \approx 0.75$. The lower the FEV_1/FVC , the more substantial the limitation in expiratory flow.

12.7.2 Ventilatory Muscle Testing

There are numerous ways to test (and quantify) ventilatory muscle strength and endurance. The most common approach is the maximal inspiratory pressure (MIP) and the maximal expiratory pressure (MEP). MIP and MEP are isometric tests performed when the respective muscles are in their lengthened position. MIP testing is performed at residual volume (RV). The subject attempts to breath in as strongly as possible on a tube connected to a device that can measure negative pressure. There can be a small air leak to prevent the use of facial muscles from contributing mouth pressure to the negative pressure generated. MEP testing is performed at total lung capacity (TLC). The subject attempts to breath out as strongly as possible on a tube connected to a device that can measure pressure. There can be a small air leak to prevent the use of facial muscles from contributing mouth pressure to the positive pressure. Even though MIP and MEP are isometric tests of pressure they are highly correlated with other approaches to measuring both ventilatory muscle strength and endurance. As with any muscle testing, care must be taken to ensure proper muscle coordination and activation. For example, when being asked to generate a strong inspiratory effort subjects commonly associate what they are being asked to do with a big breath and they associate big breaths with upper chest and accessory muscle use. If the tester wants an estimate of diaphragm strength, care must be taken to prepare the subject, and provide feedback to encourage diaphragm activation as opposed to upper test and accessory muscles. For example, establishing that subjects understand the request to breathe in with their diaphragm even before attempting the test. And then repeating whatever instructions led to a diaphragm activation while doing the test (i.e., "Breathe in forcefully, starting from deep in your lungs, starting the movement with your abdomen which should move outward while you breathe in.")

12.7.3 Ventilatory Insufficiency Patterns

There are two major ventilatory insufficiency patterns.

- Obstructive pattern: Difficulty getting air out of the lungs during expiration due to airway resistance and bronchiole collapse. This pattern includes increased RV , increased FRC and decreased VC . PFT's include reduced FEV_1 and FVC and a reduced FEV_1/FVC ratio, which means FEV_1 decreases more than FVC because expiratory flow is significantly reduced.
- Restrictive pattern: Difficulty getting air into the lungs for any reason (intra-pulmonary reasons are within the lungs such as pneumonia or fibrotic changes; extra-pulmonary reasons are outside of the lungs such as pleural effusion, chest wall changes or diaphragm weakness). FEV_1 and FVC are decreased proportionally, FEV_1/FVC ratio is normal, which means (FEV_1) is reduced because of an overall lower volume, not because of a reduction in expiratory flow.

Chronic pulmonary conditions such as emphysema and bronchitis are termed chronic obstructive pulmonary disease (COPD) because they have an obstructive pattern. The increase in RV eventually changes the anatomy of the thorax, which changes the mechanics of the inspiratory muscles. Therefore, even though there are expiratory flow limitations (reduced FEV_1/FVC), there is also a reduction in maximal inspiratory pressure (MIP) which contributes to the shortness of breath and activity intolerance. The increased RV also results in substantially under ventilated alveoli (reduced V_A), which leads to substantial right-left shunt, which leads to substantial V/Q mismatching, and therefore substantial problems with respiration that include both decreased P_aO_2 and increased P_aCO_2 .

Asthma is a chronic condition that has acute obstructive episodes. Sometimes it is classified as a chronic obstruction, but this should only be the case when there actually is chronic obstruction (chronically reduced FEV_1/FVC ratio).

12.8 Summary

Ventilation is the movement of air into and out of the lungs and involves activation of skeletal muscles that change the shape of the thoracic cavity. This is always an active tension process for inspiration, and is usually a passive tension process for expiration. Forced (or additional) expiration requires active tension. Alveolar ventilation is the movement of air into and out of the alveoli. Alveolar ventilation that includes pulmonary perfusion (circulation) is involved in respiration. Ventilation that does not reach alveoli is called anatomical dead space ventilation. Alveolar ventilation that does not receive sufficient pulmonary circulation is called alveolar dead space ventilation. Physiological dead space (or just dead space) is the sum of anatomical and alveolar dead space. Pulmonary circulation that goes to alveoli not receiving sufficient ventilation is a right-left shunt. Dead space and right-left shunt reduce V/Q matching (the opposite of V/Q matching is V/Q mismatching).

References

- [DKW05] André De Troyer, Peter A. Kirkwood, and Theodore A. Wilson. *Respiratory Action of the Intercostal Muscles*. en. *Physiological Reviews* 85.2 (Apr. 2005), pp. 717–756. issn: 0031-9333, 1522-1210. doi: [10.1152/physrev.00007.2004](https://doi.org/10.1152/physrev.00007.2004). url: <https://www.physiology.org/doi/10.1152/physrev.00007.2004> (visited on 07/27/2022).
- [GCB14] Chris Gilbert, Leon Chaitow, and Dinah Bradley. *Recognizing and Treating Breathing Disorders*. Elsevier Health Sciences, 2014.
- [HH20] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences, 2020.
- [LT20] Andrew B Lumb and Caroline R Thomas. *Nunn's applied respiratory physiology*. Elsevier Health Sciences, 2020.
- [Lun08] Jon O. Lundberg. *Nitric Oxide and the Paranasal Sinuses*. en. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* 291.11 (Nov. 2008), pp. 1479–1484. issn: 19328486, 19328494. doi: [10.1002/ar.20782](https://doi.org/10.1002/ar.20782). url: <https://onlinelibrary.wiley.com/doi/10.1002/ar.20782> (visited on 07/26/2022).
- [Tor+02] D.C.F. Tornberg et al. *Nasal and oral contribution to inhaled and exhaled nitric oxide: a study in tracheotomized patients*. en. *European Respiratory Journal* 19.5 (May 2002), pp. 859–864. issn: 0903-1936, 1399-3003. doi: [10.1183/09031936.02.00273502](https://doi.org/10.1183/09031936.02.00273502). url: <http://erj.ersjournals.com/cgi/doi/10.1183/09031936.02.00273502> (visited on 07/26/2022).

Sample Questions

1. If $VD_{ana} = 100$ mL, $V_T = 600$ mL, respiratory rate is 12 breaths per minute, what is V_A per minute?
2. What is the amount of air in the lungs following a passive exhalation (the end of a resting tidal volume exhalation) when the inward elastic recoil of the lungs equals the outward recoil of the thorax (chest cavity)?
3. If the diaphragm encounters more resistance to descending when it contracts (such as when someone is in supine), then dominant diaphragm action is?
4. What happens with atelectasis?
5. In normal conditions, what are major triggers for stimulating deeper and faster breathing (more ventilation)?

Next Steps

There are several next steps to this book. This is not even a first edition. This is a pre-edition, a third draft of what will be a completed book. One thing is clear though - it will always be a self-published, no cost (eBook) or low cost (printing cost only) textbook. Though as the pages increase the cost for pages will increase. But no fear. If you've purchased this draft, keep it and if a new draft or when the first edition is printed, I'll give you a new printed copy *gratis*.

New Content

Chapter 13 - Digestion, Absorption and Metabolism is currently a missing but necessary chapter for **Part III** on Muscle Support. Without food and water entering the body, being digested, absorbed and metabolized the extra-cellular fluid cannot sustain muscle fibers. It has not been written yet because - at least at Plymouth State University DPT - it is not part of the first Clinical Physiology course. It is taught in the Nutrition and Exercise Prescription course. It is forthcoming and in progress. It sits at the border of Part III on muscle support and Part IV on the maintenance of long term muscle cell health, growth and repair.

Part IV is currently in process and focuses on muscle molecular biology. Specifically, how the muscle continues to fulfill its role of creating tension without being too large of a burden on the physiological support systems. Topics include maintaining muscle cells and tissue for the demands they are placed under through the processes of atrophy, hypertrophy, and repair.

Part V is also in process and will include a variety of integrated exercises. These topics have been selected to highlight the need for physiological coordination such as Fick's Equation to highlight integration of the material in Parts II and III; exercise prescription and antigravity, cold and hypoxia exposure to highlight integration of the material in Parts II, III and IV; the autonomic nervous system to highlight integration of Parts III and IV; and rehabilitation protocols to highlight integration of Part I with the all other parts as the full spectrum of hierarchy is manifest in clinical practice.

Clinical practice is a fulfillment of clinical physiology. Practice abductively reasons based on physiology to understand movement problems; and practice attempts to deductively modify physiology through interventions.

Update Content

It is hoped that all of the typos, errors and mishaps in the existing content will be fixed prior to the final publication of the first edition of the text. There are still some rather long and muddy sections that require at least one more pass through in the tried and tested locale of classroom teaching and student feedback!

Added Features

Additional features for the first edition will be the inclusion of a glossary of key terms at the end of each chapter, and an index at the end of the book. In all honesty, I really need a full printed draft to highlight those terms worthy of being in the glossary and the index. Finally, an end of book Works Cited page is necessary for those - like me - that enjoy reading

through the works cited in masse and not just by chapter. This feature will not remove the end of chapter references because both are options worthy of inclusion for different preferences.

Next Steps

Next Steps