Software Requirements Specification for Diagnose: Medical Prediction Tool for Acquired immunodeficiency syndrome (AIDS)

Andrea Clemeno

December 9, 2020

Contents

Revision History

Date	Version	Notes
October 10 2020	1.0	Initial Draft of SRS
December 3	1.1	Revised Draft of SRS
December 8	1.2	2nd Revised Draft of SRS

1 Reference Material

This section records information for easy reference.

1.1 Table of Units

Throughout this document SI (Système International d'Unités) is employed as the unit system. For each unit, the symbol is given followed by a description of the unit and the name.

symbol	**unit**	**name**
d	time	day
mL	volume	millilitre
mol	amount of substance	mole

Table 1: Table of Units

1.2 Table of Symbols

The table that follows summarizes the symbols used in this document along with their units. The symbols are listed in alphabetical order.

symbol	**unit**	**description**
k	d^{-1}	Elimination Constant
N	$\frac{\text{mol}}{\text{mL}}$	Viral Load
N_o	$\frac{\text{mol}}{\text{mL}}$	Initial Viral Load
N_p	$\frac{\text{mol}}{\text{mL}}$	Predicted Viral Load at time t_p
N_t	$\frac{\text{mol}}{\text{mL}}$	Viral Load at time t_t
n	mol	Number of virions
r	$\frac{\text{mol}}{\text{mL} * \text{d}}$	Rate of change of the viral load
t	d	Time
t_p	s	Time duration of chosen prediction period
t_t	s	Time duration between initial and secondary viral load test
V	mL	Volume

Table 2: Table of Symbols

1.3 Abbreviations and Acronyms

The table that follows summarizes the symbols used in this document that allude to different sections of the Software Requirements Specification. The symbols are listed in alphabetical order.

symbol	**description**
A	Assumption
DD	Data Definition
GD	General Definition
GS	Goal Statement
IM	Instance Model
LC	Likely Change
PS	Physical System Description
R	Requirement
SRS	Software Requirements Specification
Diagnose	Medical Diagnosis Prediction Tool for Acquired immunodeficiency syndrome (AIDS)
T	Theoretical Model
ULC	Unlikely Change

Table 3: Table of Abbreviations and Acronyms

2 Introduction

Human immunodeficiency virus 1 (HIV -1) is the most common type of HIV virus that attacks cells of the immune system needed to fight off diseases. After being infected with the HIV-1 virus, the state of a patient's health worsens and the condition can be classified as Acquired immunodeficiency syndrome (AIDs), a disease which affects 38 million people worldwide (who, 2020). Although there is no cure for AIDs, antiretroviral therapy helps control a patient's condition. Having a program that assesses the efficient of a patient's immune system and predict their viral load progression will be useful in creating a treatment plan. The program documented here is called Diagnose.

The following section provides an overview of the Software Requirements Specification (SRS) for Diagnose. This section explains the purpose of this document, the scope of the requirements, the characteristics of the intended reader, and the organization of the document.

2.1 Purpose of Document

The main purpose of this document is to outline the software requirements of the medical prediction system. Different aspects of the system including goals, assumptions, theoretical models, and definitions will be explained to ensure full understanding of the system. The following SRS document will remain abstract exploring what is being solved rather than how it will be solved.

This initial document is the start of a series of documents that will outline the development of the software tool, Diagnose. The SRS will describe the system context and constraints, the specific problem definition and solution characteristics, requirements and likely and unlikely changes for the development of the tool.

2.2 Scope of Requirements

The scope of the requirements includes the analysis of the HIV-1 viral load concentration over time after the viremia peak of the initial infection without the influence of external factors that can affect the efficiency of the immune system.

2.3 Characteristics of Intended Reader

Reviewers of this documentation should have a basic understanding of virus dynamics and high school differential calculus. The users of Diagnose must have a higher level of expertise, as explained in Section: User Characteristics (Section 3.2).

2.4 Organization of Document

The organization of the document will present the system's goals, theories, definitions, and assumptions. The template for an SRS for scientific computing will be followed. To approach Diagnose from a higher level to a more elaborate level, readers can begin with reading the system's goal statements. Subsequently, the theoretical models will elaborate on the goal statements. Lastly, readers can finish with a more understanding of the system by reading instance models of the system.

Oppositely, readers can read from the instance models to the goal statements for a more specific to general understanding.

3 General System Description

This section provides general information about the system. It identifies the interfaces between the system and its environment, describes the user characteristics, and lists the system constraints.

3.1 System Context

The flow chart in (Figure 1) shows the system context. The circles represent a user that interacts with the software. The rectangle represents the software system (Diagnose). The arrows display the input data from the user and the output data that is useful for the user. The interaction presented will be facilitated using an application programming interface (API).

- User Responsibilities:
  - Provide viral load data for patient's initial and secondary viral load tests, time duration between the tests and the time duration between the initial test and the chosen prediction date.

Figure 1: System Context

  - Ensure required software assumptions (Section 4.2.1) are appropriate for any particular problem the software addresses.

- Diagnose Responsibilities:

  - Assess the input data and determine if the required physical and software constraints (Section 4.2.6) are met.
  - Calculate the elimination rate of the virus and the predicted viral load after the chosen prediction period.

3.2 User Characteristics

The users of Diagnose should be HIV healthcare providers including doctors of medicine (MD) or doctors of osteopathic medicine (DO), nurse practitioners (NP), or a physician assistant (PA). (Content Source: CDC's HIV Treatment WorksDate last updated: May 21, 2018)

3.3 System Constraints

The system, Diagnose, is limited to only instances where viral load is eliminating. If the viral load is increasing, the immune system has not recognized the intruders and a estimation of immune system efficiency using the elimination rate cannot be determined. More specifically, the initial viral load must be greater than the viral load on the consecutive test.

4 Specific System Description

This section presents the problem description, which gives a high-level view of the problem to be solved. This is followed by the solution characteristics specification, which presents the assumptions, theories, definitions and finally the instance models.

4.1 Problem Description

AIDs is one of the most serious health challenges in our global community. Many individuals infected with HIV visit healthcare facilities with hopes to find solutions to manage their conditions. A system for the characterization of a patient's condition can prove useful in predicting the progression of viruses like HIV, and the diagnosis of diseases like AIDs. Using this system, important decisions reached by healthcare professionals will be more data-driven and quick.

Diagnose is intended to model the viral load over time for the HIV-1 Virus responsible for AIDs. This system could be used to assess risk before substantial immune destruction has occurred.

4.1.1 Terminology and Definitions

This subsection provides a list of terms that are used in the subsequent sections and their meaning, with the purpose of reducing ambiguity and making it easier to correctly understand the requirements:

- Virus: Submicroscopic parasites that infect cells.
- Viral load: The concentration of HIV virus at a point in time.
- Replication: The process where infected cells rapidly reproduce genetic material of the virus rather than it's normal process of reproducing itself. (Burrell et al., 2017)
- Productivity: Cells can be productive and non-productive. Productive means that infection occurs. The HIV-1 Virus needs to interact with a cell to be productive and replicate. (Burrell et al., 2017)
- Infected cells: Cells that interact with the virus replicate into cells altered by the virus.

- Helper T cells: Cells of the immune system that neutralize infected cells. (William C. Shiel Jr., 2018).
- Immune Response: The defensive reaction of the human body against the harmful substances like HIV.
- Elimination: Physical quantity undergoing a decline in amount.
- Viremia peak: The peak of viral load after the virus has entered the blood stream. There is a trend of elimination of viral load due to the body's immune response. (Little, 1999)
- AIDs: Acquired Immunodeficiency Syndrome develops from an increase in HIV viral load to the extent where T cell count decreases to under 200 mol/mL.(Content Source: HIV.govDate last updated: June 05, 2020)
- Diagnosis: The determination of a patient's condition reached by a healthcare professional.
- Progression: The development towards a more advanced stage.

4.1.2 Physical System Description

The physical system of Diagnose, as seen in (Figure 2), occurs within the human body and includes the following elements:

PS1: The HIV Virion.

PS2: An infected cell with the virus.

PS3: The helper T cells.

PS4: The human body.

PS5: The skin barrier of the human body.

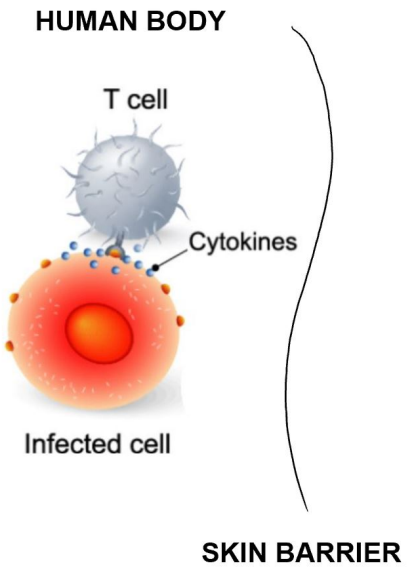

Figure 2: The Physical System
(Contributors, 2019)

4.1.3 Goal Statements

Given two HIV-1 viral load test datum, time duration between the tests and the chosen prediction period, the goal statements are:

GS1: determine-elimination-rate: Determine the rate of elimination of the HIV virus due to immune response.

GS2: predict-VL: Predict Viral Load after chosen prediction period.

4.2 Solution Characteristics Specification

The instance models that govern Diagnose are presented in Subsection 4.2.5. The information to understand the meaning of the instance models and their derivation is also presented in Assumptions, Theoretical Models, General Definitions and Data Definitions, so that the instance models can be verified.

4.2.1 Assumptions

This section describes the theoretical assumptions that the system will be based on. The theoretical models will be easier to understand by explaining these assumptions.

A1: initial-inf: Initial infection of a HIV patient.

A2: const-growth: The virions will invade uninfected cells at a constant rate (Ref By: LC2).

A3: const-volume: The dimensions of the location associated with the infection remains constant.(Ref By:A4, UC3)

A4: const-conditions: Temperature in the infection location remains constant. Higher temperatures boost virus replication and viral load (Roesch et al., 2012). (Ref By: UC3)

A5: all-productive: All viruses within the volume have productively infected cells. (Ref By: UC3)

A6: always-elimination: After the viremia peak, no significant upward trends occur.

A7: neglect-drugs: The effect of antibiotic drugs or therapy on the elimination rate will not be considered. (Ref By: IM1, IM2)

A8: neglect-sick: The effect of other infections on the elimination rate will not be considered. (Ref By: IM1, IM2)

A9: proportional: The elimination of the virus is assumed to be proportional to the amount of viruses present. (Ref By: IM1, IM2)

4.2.2 Theoretical Models

This section focuses on the general equations that Diagnose is based on.

Number	T1:expElim
Label	**Rate of Change of Viral Load**
Equation	$r = \frac{dN_t}{dt} = -kN_o$
Description	r is the rate of change of the viral load ($\frac{mol}{mLd}$) t is the time (d) N_t is the viral load at time t ($\frac{mol}{mL}$) k is the elimination constant (d^-1) N_o is the initial viral load ($\frac{mol}{mL}$)
Source	(Libretexts, 2020)
Ref. By	GD1

4.2.3 General Definitions

This section collects the laws and equations that will be used to build the instance models.

Number	GD1: vLoadt
Label	**VLoadt as a function of time for constant decay rate**
Equation	$N_t = N_o e^{-kt}$
Description	N_o is the initial amount that will undergo an exponential decrease. N_o is equivalent to $N(t = 0)$. N_t is the viral load at time t $(\frac{mol}{mL})$ k is the clearing constant (d^{-1}) t is the time (d)
Source	(Hobbie and Roth, 1970)
Ref. By	IM1 , IM2

Detailed derivation of viral load at time t:

Using the First-Order rate Law in TM: expElim, we have:

$$\frac{dN}{dt} = -kN_{\mathrm{o}}$$

Where N_{t} denotes the viral load at time t, N_{o} denotes the initial viral load and k denotes the elimination constant. When rearranging for integration, we have:

$$\int_{N_{\mathrm{o}}}^{N_{\mathrm{t}}} 1\, dN_{\mathrm{t}} = -\int_{0}^{t} k\, dt$$

Performing the integration, we have the required equation:

$$N_{\mathrm{t}} = N_{\mathrm{o}} e^{-kt}$$

4.2.4 Data Definitions

This section collects and defines all the data needed to build the instance models.

Number	DD1: Viral load
Label	**Viral Load**
Symbol	N
Units	$\frac{mol}{mL}$
Equation	$N = \frac{n}{V}$
Description	N is the viral load at a given time ($\frac{mol}{mL}$). n is the number of virions (mol). V is the volume (mL).
Sources	(Fischer et al., 2004)
Ref. By	IM1 , IM2

4.2.5 Instance Models

This section transforms the problem defined in Section 4.1 into one which is expressed in mathematical terms. It uses concrete symbols defined in Section 4.2.4 to replace the abstract symbols in the models identified in Sections 4.2.2 and 4.2.3.

The goals GS1: determine-elimination-rate and GS2: predict-VL-30 are met by IM1and IM2.

Number	IM1: calofElimConst
Label	**Calculation of elimination rate**
Input	N_t, N_o, t_t The input is constrained so that $N_o > N_t > 0$, $t_t > 0$
Output	k, such that $k = \frac{ln(N_o) - ln(N_t)}{t_t}$
Description	k is the elimination constant (d^{-1}) N_o is the initial viral load $(\frac{mol}{mL})$ N_t is the viral load at time t $(\frac{mol}{mL})$ t_t is the time duration between the initial viral load test and secondary viral load test (d)
Sources	(Libretexts, 2020)
Ref. By	IM2

Detailed derivation of elimination constant:

Using the relationship for the viral load seen in DD: viralLoad with respect to time in GD: vLoadt, we have:

$$N_{\mathrm{t}} = N_{\mathrm{o}} e^{-kt_{\mathrm{t}}}$$

Where N_{t} denotes the viral load at time t_{t}, N_{o} denotes the initial viral load and k denotes the elimination constant. When isolating for the elimination constant, we have:

$$\frac{N_{\mathrm{t}}}{N_{\mathrm{o}}} = e^{-kt_{\mathrm{t}}}$$

To isolate further, the natural logarithm is applied:

$$\ln\left(\frac{N_\mathrm{t}}{N_\mathrm{o}}\right) = -kt_\mathrm{t}$$

After using the logarithmic quotient property, we have the required equation:

$$k = \frac{\ln\left(N_\mathrm{o}\right) - \ln\left(N_\mathrm{t}\right)}{t_\mathrm{t}}$$

Number	IM2: calofPredictedVL
Label	**Calculation of predicted viral load**
Input	N_o, k, t_p from IM1 The input is constrained so that $N_o > N_t > 0$, $t_p > t_t > 0$ and $k > 0$
Output	N_p, such that $N_p = N_o e^{-kt_p}$
Description	N_p is the predicted viral load after 30 days ($\frac{mol}{mL}$) N_o is the initial viral load ($\frac{mol}{mL}$) k is the clearing constant (d^{-1}) t_p is the chosen prediction period (d)
Sources	(Hobbie and Roth, 1970)
Ref. By	-

Detailed derivation of predicted viral load after 30 days:

Using the relationship for the viral load seen in DD: viralLoad with respect to time in GD: vLoadt, we have:

$$N_\mathrm{t} = N_\mathrm{o} e^{-kt_\mathrm{t}}$$

Where N_{t} denotes the viral load at time t, N_{o} denotes the initial viral load and k denotes the elimination constant. When predicting the viral load after 30 days, the elimination constant found in IM: calofElimConst, is used instead of N_{t} therefore we have:

$$N_{\mathrm{p}} = N_{\mathrm{o}} e^{-k t_{\mathrm{p}}}$$

4.2.6 Input Data Constraints

Table 4 shows the data constraints on the input variables. The column for physical constraints gives the physical limitations on the range of values that can be taken by the variable. The column for software constraints restricts the range of inputs to reasonable values. The software constraints will be helpful in the design stage for picking suitable algorithms. The constraints are conservative, to give the user of the model the flexibility to experiment with unusual situations. The column of typical values is intended to provide a feel for a common scenario. The uncertainty column provides an estimate of the confidence with which the physical quantities can be measured. This information would be part of the input if one were performing an uncertainty quantification exercise.

Table 4: Input Variables

Var	Physical Constraints	Software Constraints	Typical Value	Uncertainty
N_o	$N_o > 0$	$N_o > N_t$	$10 * 10^6 \frac{mol}{mL}$	10%
N_t	$N_t > 0$	$N_o > N_t > 0$	$5 * 10^6 \frac{mol}{mL}$	10%
t_p	$t_p > 0$	$t_p > t_t$	30 d	N/A
t_t	$t_p > t_t > 0$	$t_p > t_t$	1 d	N/A

4.2.7 Properties of a Correct Solution

A correct solution must exhibit an elimination constant greater than 0 and a predicted viral load less than the initial and secondary viral load, N_o and N_t, respectively. Table 5 shows the data constraints for the output variables.

Table 5: Output Variables

Var	**Physical Constraints**
k	$k > 0$
N_p	$N_o > N_p > 0$

5 Requirements

This section provides the functional requirements, the business tasks that the software is expected to complete, and the non-functional requirements, the qualities that the software is expected to exhibit.

5.1 Functional Requirements

R1: Input-Values: Input the values from subsection 4.2.6.

R2: Verify-Input-Values: The software will ensure that the inputs are not out of bounds and in accordance with the data constraints. If any inputs are out of bounds, an error message is displayed in subsection 4.2.6.

R3: Calculate-Values: Software calculates the elimination rate of HIV and the viral load at the chosen prediction period from IM1 and IM2.

R4: Verify-Output: The output values will be cross referenced with the result constraints from subsection 4.2.7.

R5: Output-Values: Output related requirements including the elimination rate and predicted viral load.

5.2 Non-functional Requirements

This section provides the non-functional requirements, the qualities that the software is expected to exhibit.

NR1: Verifiable: The Drasil generated python code is tested using static and dynamic code analysis, linting, unit and system testing and continuous integration methods outlined in the Verification and Validation plan of Diagnose (Clemeno, 2020).

NR2: Understandability: The code will be generated using the Drasil Framework into modules to address each requirement in subsection 5.1.

NR3: Reusable: The code will be created using Drasil. The framework has data libraries of scientific theories, concepts, quantities and units commonly used in scientific problems. The created libraries of Diagnose can be reused as required. (Owojaiye, 2020)

NR4: Maintainablility: The traceability between requirements, assumptions, theoretical models, general definitions, data definitions, instance models, likely changes, unlikely changes, and modules is completely recorded in traceability matrices in the SRS document to measure maintainability. The traceability will be measured by ensuring relationship from inputs to instance model is recorded and identifiable.

NR5: Portablility: The generated code is able to run in different environments.

6 Likely Changes

This section lists the likely changes to be made to software

LC1: More-Inputs: The software may be expanded to include more inputs from the user to increase the accuracy of the output.

LC2: More-outputs: The program may be extended to include more outputs like assessing the progression to AIDs or suggestion for therapy.

7 Unlikely Changes

This section lists the unlikely changes to be made to the software.

UC1: determine-elimination-rate: The goal of determining the elimination rate of the virus will not likely change.

UC2: external-input: There will always be a source of input data external to the software.

UC3: constant-conditions: The input constraints will not likely change.

8 Traceability Matrices and Graphs

The purpose of the traceability matrices is to provide easy references on what has to be additionally modified if a certain component is changed. Every time a component is changed, the items in the column of that component that are marked with an "X" may have to be modified as well. Table 6 shows the

dependencies of theoretical models, general definitions, data definitions, and instance models with each other. Table 7 shows the dependencies of instance models, requirements, and data constraints on each other. Table 8 shows the dependencies of theoretical models, general definitions, data definitions, instance models, and likely changes on the assumptions.

	T1	GD1	DD1	IM1	IM2
T1					
GD1					
DD1		X			
IM1	X		X		X
IM2	X		X	X	

Table 6: Traceability Matrix Showing the Connections Between Items of Different Sections

	4.2.6	R1	R2	R3	R4	R5
IM1	X			X	X	X
IM2	X			X	X	X

Table 7: Traceability Matrix Showing the Connections Between Requirements and Instance Models

The purpose of the traceability graphs is also to provide easy references on what has to be additionally modified if a certain component is changed. The arrows in the graphs represent dependencies. The component at the tail of an arrow is depended on by the component at the head of that arrow. Therefore, if a component is changed, the components that it points to should also be changed.

	A1	A2	A3	A4	A5	A6	A7	A8	A9
T1	X					X			
GD1		X	X	X					X
DD1		X	X	X	X				X
IM1		X	X	X	X				X
IM2		X	X	X	X	X			X
LC1	X								
LC2					X	X			X

Table 8: Traceability Matrix Showing the Connections Between Assumptions and Other Items

9 References

References

Hiv/aids, 2020. URL https://www.who.int/news-room/fact-sheets/detail/hiv-aids.

Christopher J. Burrell, Colin R. Howard, and Frederick A. Murphy. Chapter 4 - virus replication. In Christopher J. Burrell, Colin R. Howard, and Frederick A. Murphy, editors, *Fenner and White's Medical Virology (Fifth Edition)*, pages 39 – 55. Academic Press, London, fifth edition edition, 2017. ISBN 978-0-12-375156-0. doi: https://doi.org/10.1016/B978-0-12-375156-0.00004-7. URL http://www.sciencedirect.com/science/article/pii/B9780123751560000047.

Andrea Clemeno. Verification and validation plan, Nov 2020. URL https://github.com/andreamclemeno/CAS741-Diagnose/blob/master/docs/VnVPlan/BlankProjectTemplate_docs_VnVPlan_VnVPlan.pdf.

2018 Content Source: CDC's HIV Treatment WorksDate last updated: May 21. Types of providers, Nov 2018. URL https://www.hiv.gov/hiv-basics/starting-hiv-care/find-a-provider/types-of-providers.

2020 Content Source: HIV.govDate last updated: June 05. What are hiv and aids?, Jun 2020. URL https://www.hiv.gov/hiv-basics/overview/about-hiv-and-aids/what-are-hiv-and-aids.

Anonymous Contributors. $group_{5p}resentation_{3-i}mmune_system---wiki$, 2019.$URL$. [Online; accessed 11-October-2020].

Ulrike R. Fischer, Willy Weisz, Claudia Wieltschnig, Alexander K. T. Kirschner, and Branko Velimirov. Benthic and pelagic viral decay experiments: a model-based analysis and its applicability. *Applied and Environmental Microbiology*, 70(11):6706–6713, 2004. ISSN 0099-2240. 10.1128/AEM.70.11.6706-6713.2004. URL https://aem.asm.org/content/70/11/6706.

Russell K. Hobbie and Bradley J. Roth. Exponential growth and decay, Jan 1970. URL https://link.springer.com/chapter/10.1007/978-0-387-49885-0_2.

Libretexts. 14.4: The change of concentration with time (integrated rate laws), Aug 2020. URL https://chem.libretexts.org/Bookshelves/General_Chemistry/Map:_Chemistry_-_The_Central_Science_(Brown_et_al.)/14:_Chemical_Kinetics/14.4:_The_Change_of_Concentration_with_Time_(Integrated_Rate_Laws).

S J et al. Little. Viral dynamics of acute hiv-1 infection., 1999. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2195636/#.

Oluwaseun Owojaiye. Creating your project in drasil, 2020. URL https://github.com/JacquesCarette/Drasil/wiki/Creating-Your-Project-in-Drasil.

Ferdinand Roesch, Oussama Meziane, Anna Kula, Sébastien Nisole, Françoise Porrot, Ian Anderson, Fabrizio Mammano, Ariberto Fassati, Alessandro Marcello, Monsef Benkirane, and Olivier Schwartz. Hyperthermia stimulates hiv-1 replication. *PLOS Pathogens*, 8(7):1–14, 07 2012. 10.1371/journal.ppat.1002792. URL https://doi.org/10.1371/journal.ppat.1002792.

MD William C. Shiel Jr. Definition of t cell, Dec 2018. URL https://www.medicinenet.com/script/main/art.asp?articlekey=11300.