TRUISM (TRANSFUSION INVENTORY PROCESS MINING FRAMEWORK): A PROCESS MINING-BASED FRAMEWORK FOR UNDERSTANDING BLOOD TRANSFUSION PRODUCT INVENTORIES

by

Calvino Ka-Wing Cheng

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
March 2017

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Transfusion product inventories are a vital part of modern health care delivery as they bridge supply and demand to deliver a quality blood product in a timely manner. Blood products are also costly, scarce, and there are also ethical considerations. The management of these blood product resources is becoming more complex, especially in the context of an increasingly diverse product menu, and in an environment where cost and efficiency pressures are becoming more prevalent. To meaningfully manage and understand real blood product inventories, the transfusion community has relied upon a value-based perspective to understand their functioning, using key performance indicators such as inventory size and wastage rates. These methodologies are limited in their ability to fully describe inventory performance. Although other tools, such as modelling and simulation have been used, they can only represent a real inventory in an idealized state and are used theoretically but not for day-to-day description of an inventory. Although papers in the past decade have alluded to blood product inventories as processes, and although inventories possess the basic attributes of a process, the conventional transfusion inventory management paradigm has not evaluated them with process-based methodologies. Therefore, to fully evaluate the inventory, the complementary use of process-based tools and value-based tools should occur. This research describes the creation of the TRansfUsion Inventory proceSs Mining (TRUISM) Framework which is a generalizable approach applicable to transfusion service inventory data for the purpose of generating valid process mining results. The research describes the subsequent knowledge translation of the process mining methodology for two case studies on red cell unit wastage and automated blood fridge usage. The results of these case studies demonstrate that further valuable process-based insight and guidelines for process improvement can be generated using a process-centric view to complement the conventional value-centric one. This may allow global transfusion inventory managers to gain further insight into their practices and potentially improve inventory performance.

LIST OF ABBREVIATIONS USED

AABB	American Association of Blood Banks
BPM	Business Process Management (systems or software)
BSMS	Blood Stocks Management Scheme
BTS	Blood Transfusion Services
CBS	Canadian Blood Services
CCL	Cerner Command Language
CDN	Canadian [reference to currency]
CMV	Cytomegalovirus
CP	Concurrent Probing methodology (usability.org)
CRM	Customer Relationship Management (systems or software)
CT	Crossmatch:Transfuse [ratio]
CTA	Concurrent Think Aloud methodology (usability.org)
CSA	Canadian Standards Association
CSV	Comma separated values
CVAR	coefficient of variation of daily transfusion
COBCON	Cost of Blood Consensus Conference
CZBTS	Central Zone Blood Transfusion Services
DGH or DG	Dartmouth General Hospital
DOH	Days of inventory on hand
ERP	Enterprise Resource Planning (systems or software)
FIFO	First in, first out
FP	frozen plasma
FFP	fresh frozen plasma
FSM	Final state machine
GB	Gigabytes
HI	Halifax Infirmary
HC or HCH	Hants Community Hospital
HLA	Human leukocyte antigen
ISI	Issuable stock index

KPI	Key performance indicator
LIFO	Last in, first out
LIS	Laboratory information system
MB/s	Megabytes/second
MDU	Medical Day Unit (outpatient oncology treatment)
MFLOPS	mega floating-point operations per second
MSI	Medical Surgical Intensive Care Unit
MXML	Mining eXtensible Markup Language
NSHA	Nova Scotia Health Authority
PAIS	process aware information systems
PC	Personal computer
PInG	Pathology Informatics Group
PMMF	Process Mining Methodology Framework
QA	Quality assurance
RAM	Random Access Memory
RBC	Red blood cell
RCA	Root cause analysis
RP	Retrospective probing methodology (usability.org)
RSL	Residual shelf life
RTA	Retrospective Think Aloud methodology (usability.org)
SME	Subject matter expert
SOP	Standard operating procedure
TRUISM	TRansfUsion Inventory proceSs Mining Framework
VANESA	Vital Appropriate Knowledge (silent K) Empowers Stock Analysis
VGH or VG	Victoria General Hospital
VM	Virtual machine
WAPI	Wastage as a percentage of issuance
WHO	World Health Organization
WMS	Warehouse management system
XES	eXtensible Event Stream

ACKNOWLEDGMENTS

I would like to acknowledge…

My supervisor, Dr. Raza Abidi, for coaching me to think and write scientifically, and for enduring my writing style as I wrote and rewrote the thesis. Thank you for making me think bigger than the research.

My committee members: Dr. Robert Liwski for his consistent cheering, encouragement and outside-the-box application questioning, Dr. Alex Easton for his ability to ask why it matters, Dr. Samina Abidi for her help on the evaluation framework and providing a balanced view, and Mr. Bryan Crocker for his encyclopedic functional knowledge of the laboratory business process and laboratory information system. Dr. Jason Acker, my external examiner, for his participation in the process and thoughtful insight.

Members of the Pathology Informatics Group at the Nova Scotia Health Authority, with special thanks to Mr. Andrew Kumar-Misir, who wrote the data extract using his query building skills and, in collaboration with Ms. Stephanie Watson, helped in the data validation. Andrew and Stephanie, thank you for being there to help me troubleshoot things and for your assistance in getting the data.

My Division Chief and Director of the Blood Transfusion Service, Dr. Irene Sadek, for her encouragement to get the PhD done, and for her willingness to let me try things out on a real-life blood transfusion service under her watch. Thank you for investing in my professional and academic development.

My Department Head, Dr. Godfrey Heathcote, for his encouragement to finish the PhD. Thank you for your great advice over the years, and for investing in me.

My colleagues of the Division of Hematopathology, who were there to support me through the nearly ten years of this endeavour, especially more recently as I wrote the thesis. I want to thank Drs. Jason Quinn and David Conrad for their assistance in helping to craft the process mining methods paper which put this methodology out in the transfusion community.

Thank you to Mr. Daniel Liwski, Dr. Thomas Covello, who helped to crunch data on side-topics related to the thesis, but not directly incorporated into the thesis, so that I could functionally validate that I was thinking on the right track.

The Blood Transfusion Service at Nova Scotia Health Authority, Central Zone, for their encouragement and the fielding of my constant questions about their processes. Thank you for your willingness to be involved in the process mining focus groups. I want to thank Ms. Joan Macleod for allowing me the freedom to experiment with the real-life blood transfusion service under her watch as well.

Ms. Eileen Kaiser, Ms. Heidi Devlin, Ms. Karen Cleveland for helping me book things, like rooms, meetings with my Committee, and setting the audiovisual equipment up. Thank you for making my life easier.

Mr. David Barteaux and Ms. Robin Parker for helping me to physically acquire articles and help in the search strategy for the literature review.

My parents, my in-laws, and my extended family for praying for me and encouraging me throughout this journey. I also want to thank Sharon's dad, who was on a similar PhD journey, for stoking my competitive side, as we raced to the end to complete our PhD theses. I would also like to thank Rev. Jacob Birch for keeping me accountable to my deadlines and for praying for me.

Finally, but most importantly, I want to thank my family for suffering with me as I spent countless days off, early mornings, late nights, holidays, and much airtime talking about the research. I want to thank my wife, Sharon, for her unfailing support while I went on this crazy adventure, and for encouraging me not to give up prematurely. Thank you for listening to my crazy ideas and rants about the research, and for the countless meals and snacks that made the process go so much more enjoyably. Thank you for supporting me in prayer. Thank you for taking care of household matters so that I could concentrate, and thank you for your wisdom and insight. Thank you for your superb thesis editing skillz. I thank God for you always and I am so blessed to have you as my life partner. I also want to thank my kids, Allister, Maxwell, Cassidy, and Declan, all of whom went on this journey with me, for being my cheerleaders and my biggest fans. I'm proud to be your dad!

Chapter 1 INTRODUCTION

The purpose of Chapter 1 is to describe the motivation for the research, state the problem, research questions, objectives, and the solution approach. It will also be to introduce the contributions of the research to the literature in both process mining and transfusion communities.

1.1 Motivation

The provision of blood products by transfusion services is necessary to deliver modern healthcare. Modern transfusion services have needed to maintain inventories to ensure that there are no shortages between supply and demand, both of which are stochastic in nature [1]. The current dogma of transfusion inventory management is to provide blood products in a timely manner for clinical needs while balancing the potential for wastage [2]. Blood product inventory management is often a background process, classically out of the spotlight of front-line medical care, but important from a patient care and system stewardship perspective. The stewardship of blood products goes beyond a supply and demand function. Blood products are costly. For instance, data from the Cost of Blood Consensus Conference (COBCON) in 2003 estimated that each red blood cell unit costs between $522 and $1183 (mean, $761), based on all direct and indirect overhead costs from donation to delivery [3-5]. In Canada, the acquisition cost of a red blood cell unit is $419 and a unit of HLA matched single donor (apheresis acquired) platelets is $1250 [6]. The cost of a two unit transfusion in Western Europe (United Kingdom, Sweden, Switzerland, Austria, and France) is €877.69 [4, 7-12]. There are also opportunity costs to sequester capital in the form of blood products inside the inventory, preventing the financial resources from being spent elsewhere in the health care system[1]. Moreover, due to increasing stringency and screening technologies to ensure a safe blood supply, blood will become increasingly costly to provide [5, 13-15]. Blood products are also scarce. For example, in Canada, blood supply is derived from approximately 12.5 million eligible donors [16], with an overall per capita donation rate of 35.8/1000 (3.6%). Since the per capita issuance rate is

32.6/1000 (3.3%), the supply margin is very thin at a difference of 3.2units/1000. This supply scarcity is seen in other countries, such as Australia (Australian Red Cross Blood Service), whereby similar rates of donation and issuance are 40.47/1000 (4.0%) and 36.77/1000 (3.7%) respectively [2]. Blood products should also be treated with utmost respect in the context of ethical obligations. For instance, in countries such as the United States, England, Scotland, and Canada [6, 17], blood products are collected from voluntary unpaid donors who donate, with the expectation that their products are being responsibly stewarded by the transfusion system. Every single donated unit represents a voluntary and altruistic contribution to someone else in need, and the donor's time and effort, which they could have spent doing something else [17]. Therefore, the steward of these blood products is expected to manage this valuable resource properly and ensure the implicit wishes of the donors are fulfilled.

Despite, an extensive body of transfusion inventory management literature of over half a century, the description, and thus the understanding of real inventory performance has been simplified to key performance indicators (KPI) pertaining to inventory size and wastage. Size descriptors include 'days of inventory on hand'(DOH) [2] and 'issuable stock index' (ISI) [18] while wastage KPIs include 'outdate %' and 'wastage as a percentage of issuance (WAPI); [18-21]. These KPIs are ubiquitous and important in their use in day-to-day practice, and in benchmarking exercises [20, 22-24], forming the basis of the current inventory management paradigm. KPIs summarize the endpoints of a complex process involving many steps and potential modifications to products within the inventory. This is the only method currently used to describe and understand real inventories in the blood transfusion community. Although other techniques such as modelling and simulation exist and have a role in transfusion inventory management, they conceptually represent idealized inventories but are not used in day-to-day description of inventories. A review summarizing the International Forum on [blood transfusion] Inventory Management highlighted the state of practice and struggles as transfusion services looked to optimise inventory management [2]. In this paper, Devine *et al*. [2] indicate that there was only one nearly universal practice, the first-in first-out (FIFO)principle, which is practiced by transfusion services to minimize wastage. In addition, various practices are being used to minimize product discard, but there are no other universally practiced inventory management techniques. This is paradoxical, as there is a seemingly mature

and established body of literature available since the 1960s to inform current practice [25, 26]. For instance, concepts such as FIFO [27], reducing the inventory to a correct size [18, 28, 29], reduction of crossmatching rates and times [18, 30], ensuring fresher products are received into the inventory [29], recycling of inventory to higher turnover sites [31, 32], and improvement of supply chain transparency through information display [24, 32-35]are some examples of principles which are available to apply to inventory management. Devine *et al.* suggest a gap between the plethora of methods published in the literature and what is consistently implemented in day-to-day practice.

It is noted that blood product inventories possess a number of attributes which allow for their classification as processes. At a basic level, a product inventory consists of a continuous series of actions and operations which ultimately result in the delivery of a blood product derived from a donor to a recipient. Despite the mature body of literature, only recent papers have acknowledged the inventory supply chain as being a process and some papers illustrate that the inventory contains defined sub-processes. One of the first acknowledgements of the overall blood transfusion supply chain as a process was in Aledort's paper. In the paper, costs were attributed to various process steps within a decomposed process flow model of a blood transfusion supply chain, although the inventory itself was not addressed as a process itself [5]. One of the first diagrammatic acknowledgements of the inventory as a process was a simulation paper from Katsaliaki and Brailsford [36], whereby 'storage in the hospital blood bank' was explicitly annotated as multiple process steps including unassigned inventory, crossmatch, and assigned inventory 'storage in hospital bank' box in Figure 1, adapted from [36]). The only other paper referring to the blood inventory as a process is by Cheng *et al.* [37], suggesting that the RBC inventory can be viewed as a 'lifecycle' and suggesting that characterising sub-processes within that lifecycle may help improve overall inventory performance.

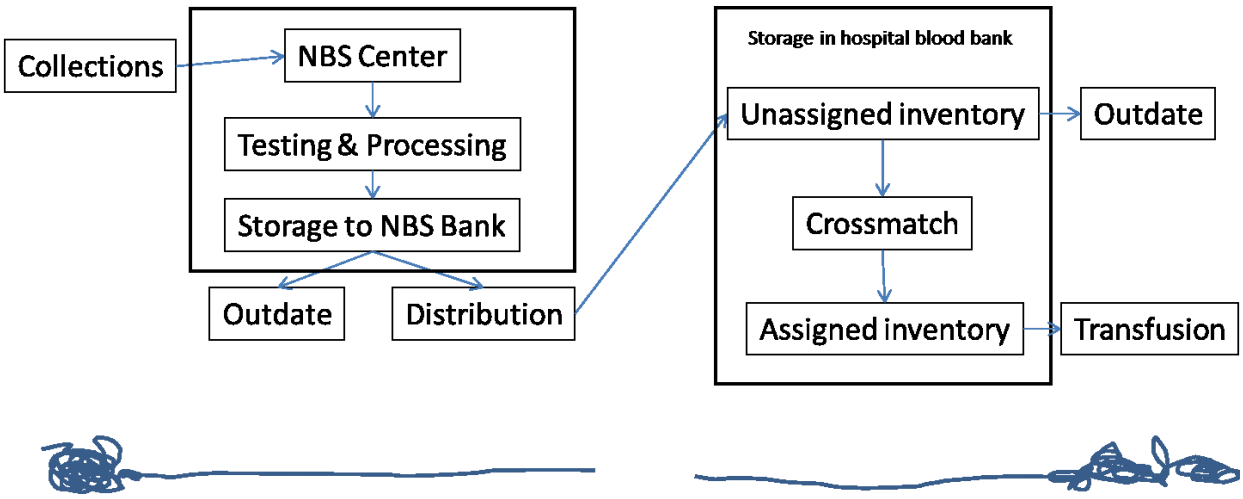


Figure 1: Flowchart of blood supply chain from supplier (left) to hospital user (right) with gap in 'rope' illustrating discontinuity of data between supplier and hospital (adapted from Katsaliaki *et al.* [36] and Devine *et al.* [2])

A handful of other papers this past decade have since described various elements of the blood transfusion supply chain as processes [22, 38-46], and this is alluded to, with the reference of the 'rope', in Devine *et al.* [2], the ends of which are drawn at the bottom of Figure 1.

If inventories are considered as a process, then they can be approached in two different ways. The first way is using expert opinion to describe the inventory and how blood products flow between receipt of the product into inventory and transfusion to recipient or other endpoints such as wastage or transfer. Manual process map creation is a subjective method for understanding a process and it may or may not ultimately result in a valid process map. If a process map is manually created, it can be fraught with oversimplification, an incorrect resolution or abstraction level, and can lead to erroneous conclusions about a process [47]. An alternative approach could be to use a process-data-driven methodology using process mining. The end-goal of process mining is to generate a process map from process log data derived from an information system that captures process steps. This methodology is less subjective and can be more efficient in generating a process map. If the process mining-based approach is used, subject matter experts (SMEs) would still be required to interpret the maps for maximum benefit [48], as they understand their business processes best.

There are potential similarities with transfusion services and financial services. For instance, auditing is a common practice in the financial sector as it is in transfusion service operations. In response to various accounting scandals at companies such as Enron and Tyco, the Sarbanes-Oxley Act (2002) and Basel II Accord (2004) have been implemented and necessitate more stringent auditing. The financial services industry has responded to this practice environment by shifting their traditional transactional value-based auditing techniques to include process-based ones [49, 50]. Prior to the use of new process-centric tools, such as process mining, value-based auditing techniques would have examined financial transactions based on the value of a transaction. Using process-oriented tools, auditors can now examine transactional violations such as signature control and segregation of duty (Section 404 of Sarbanes-Oxley), which traditionally would have been difficult to do [49]. The application of such process-centric techniques in transfusion inventory management should ideally be done with knowledge translation in real-life settings as the ultimate goal in the context of continuous process improvement.

1.2 Problem

Modern transfusion inventory management plays a crucial role in the provision of blood products in a timely manner, preventing shortages [1, 2], and minimizing the wastage of a costly [2, 4-15]and scarce resource [1]. Transfusion inventory management is now being challenged by an increasingly diverse menu of products and mounting pressures to achieve cost and performance efficiency [2]. In the past decade, the transfusion literature has begun the shift to use process-related terminology to describe the supply chain and the blood product inventory [2, 5, 22, 36-46]; however, the methods used by the transfusion community to understand and analyse inventories are still not process-centric. They are based on simple aggregated inventory size and wastage rate numerical descriptors, collectively known as KPIs or quality indicators [2, 18-21]. This method of describing inventories is an oversimplification and does not allow for maximal understanding of the inventory. Therefore, to progress the current value-based understanding of the inventory, a process-based analytical framework is required.

1.3 Research Questions

The topic of this research will be to create a process mining framework applicable to the transfusion inventory, apply existing inventory data from a laboratory information system and demonstrate the role of process mining to allow for further process-related understanding of a real transfusion inventory. Knowledge translation will occur in two areas: red cell wastage and policy conformance involving use of an automated blood fridge.

The research will focus on the red cell unit inventory within the Central Zone at Nova Scotia Health Authority and will not include patient details or Canadian Blood Services data.

1.4 Research Objectives

- To create a process-mining framework applicable to transfusion inventories
- To characterise any limitations of process mining as a technique for understanding blood transfusion product inventories.

1.5 Solution Approach

The approach will build on emerging understanding that the blood transfusion inventory is a process itself [36, 42] within the blood supply chain [2, 5, 22, 36-46], and thus should not only be understood from a value-based perspective but also from a process-oriented one. This is similar to the process-based shift occurring in financial auditing [49, 50]. The approach will be to develop a generalizable transfusion inventory process mining framework which can be translated into day-to-day transfusion community practice and will adapt components of the Process Mining Methodology Framework for real-life case studies as described by De Weerdt *et al.* [48], consistent with other analytical frameworks in process mining [51, 52].

The knowledge translation component of the research will apply the developed framework to examine two case studies using day-to-day transfusion problems, not yet documented or solved in the transfusion literature. The two case studies are chosen as they are difficult questions to answer in the current value-based paradigm and, in the context of limited human and financial resources at our institution, a process-based methodology may be a feasible alternative approach. The first case study is to examine the root causes or antecedent events relating to red cell inventory wastage. The second case study is to characterise how effectively the Hemosafe ™ automated blood dispensing fridge is being used in inventory management. These two case studies are important in the overall understanding of avoidable red cell unit wastage, especially the first case study which is an area of focus for global transfusion services [2]. This case study approach is similar to the approaches used in the financial literature related to process mining [48].

The starting point will be to use data that is already resident in the laboratory information system (LIS) and emulate the normal function of quality assurance initiatives in transfusion services. The rationale for using RBC unit data is that these units are the most frequently used cellular blood products in the transfusion inventory [2]. RBCs also have multiple transitions to fridges and inventory locations at our hospital facilities, and there is previously published research in this area describing the inventory as a lifecycle [37].

The analysis of the RBC data and generation of process maps will occur with process-mining tools used in the financial auditing literature, such as Disco (fluxicon.com) [49]. A focus group format facilitated by a process mining subject matter expert (SME) will be used to elicit feedback from transfusion SMEs while interpreting process-mining generated maps in a live environment. This mix of process mining and business experts is critical for success [48]. The feedback process for quality improvement data, typically informal at our institution, will be formalised for the purposes of the research.

1.6 Research Contributions

The main contribution of the research to the transfusion and process mining communities will be to define a framework to examine existing transfusion product datasets, with the end-goal of applying process mining to optimise transfusion inventories. This framework will define data preparation requirements, validation, and exploration steps required to generate valid process mining-derived maps. This will allow for a holistic understanding of the transfusion product inventory as the framework can provide a complementary view to the current value-based paradigm of the transfusion inventory 'rope' described by Devine *et al.*[2]. This holistic and deeper understanding could allow for further enhancement of inventory performance and possibly decrease avoidable wastage. The research introduces process mining software and vocabulary to the transfusion community, highlights its potential application, and opens the possibility for more process-based research in the area. This represents a new technique which has never been documented in the transfusion inventory management literature.

In the process mining community, this research will document another application of process mining to inventory management practices as there is only one paper in the area of application to inventory [180]. Moreover, De Weerdt *et al.* indicate there has been very little literature (less than a dozen studies) on the practical applications of process mining. The literature to date has focussed on technical contributions especially in the area of control flow discovery (defined as the creation of a process model reflecting causal event dependencies) [48]. This research will broaden the process mining literature by introducing its application to audit real transfusion product inventories.

1.7 Chapter Description

Chapter 1 provides the motivation for doing this research, an overview of the problem, research questions, research objectives and the solution approach as well as research contributions.

Chapter 2 describes the background of the transfusion business and inventory practices, what the practices are, best practices, and the research related to these topics. It highlights the various issues that transfusion services are encountering currently. This chapter also introduces the concept of process mapping and process mining, and the rationale for their use.

Chapter 3 describes the methodology, research environment, hypothesis, and how the hypothesis will be tested.

Chapter 4 discusses the results and is divided into two subsections, with one pertaining to red cell wastage, and the other pertaining to Hemosafe™ utilization. These results will highlight the descriptive and qualitative analyses, as well as highlight any discovered guidelines for process improvement.

Chapter 5 is the discussion and conclusion section, exploring the insight gained, validity of research design, limitations of research, and recommendations for further research.

Appendix A includes the literature search strategies.

Appendix B is an overview of the laboratory information system data table structures.

Chapter 2 Background

The background chapter will examine the transfusion business and how an inventory fits within this context. The background also examines the issues and gaps in transfusion inventory management practice and will introduce the concept of process mining, its application to other areas, and demonstrate a rationale for using this analytical technique.

2.1 The Role of Blood Transfusion Inventory and Its Management

The transfusion supply chain connects the blood supplier and the hospital provider. It is an ecosystem of collection, processing, testing, and delivery steps which ultimately results in the delivery of a blood product from the donor to the recipient [25, 36, 40, 42, 43, 53-57]. Due to the stochastic nature of the availability of donors and the unpredictability of demand, both blood supplier and hospital providers must maintain inventories to buffer this uncertainty [1, 25]. For instance, supply could be impacted due to various reasons such as the availability and willingness of donors to donate, permanent or temporary deferral of donors due to emergence of transmissible disease, unpredictable quality assurance issues, weather-related supply disruptions, and other unforeseen factors [57-70]. Likewise, demand is stochastic, and may fluctuate on the daily interventional (i.e. surgical, oncological) caseload and unforeseen events such as massive transfusion secondary to bleeding or traumas. Due to this supply/demand uncertainty, transfusion product inventories are a necessary aspect to ensure availability of blood products in transfusion medicine practice.

There are two aspects of properly managing a transfusion product inventory. The first aspect is to avoid shortage and be able to deliver a quality blood product to the recipient in a timely manner [2]. The second aspect is to ensure that the blood products are not wasted, due to their scarcity, perishability, cost, and other ethical factors.

Firstly, all blood products are perishable and have variable shelf lives, as noted in Table 1, with platelets being the shortest and plasma being the longest; red blood cell units are in between with a shelf-life of 42 days, though irradiated RBCs have a maximum shelf-life of 29 days [6]. Due to variable storage conditions (from room temperature to frozen), as noted in Table 1, this poses a challenge for transportation and storage of products, as different products must be separately transported and separately stored in inventory.

Table 1: Common Whole Blood Derived Blood Products, Product Characteristics, and Storage Conditions (Adapted from [6])

Components	Type of component	Volume	Shelf life	Storage temperature
Red blood cells	Cellular	280 mL	42 days	1-6°C
Buffy coat derived platelets (from 4 units)	Cellular	350 mL	5 days	20-24°C
Apheresis platelets	Cellular	300 mL	5 days	20-24°C
Frozen plasma	Plasma	250 mL	1 year -	-18°C or colder
Apheresis plasma	Plasma	500 mL	1 year -	-18°C or colder
Cryoprecipitate	Plasma	15 mL	1 year -	-18°C or colder

Secondly, blood products are costly, and have both an absolute financial cost (Table 2) and an opportunity cost to the health care system, the latter of which reflects financial capital that is tied up inside an inventory and cannot be invested elsewhere [1]. As indicated before, data from the Cost of Blood Consensus Conference (COBCON) in 2003 indicated that a red blood cell (RBC) unit costs between $522 and $1183 (mean, $761) [3-5]. This price was based on the cumulative attribution of costs associated with the entire manufacturing chain including donation, testing, storage, distribution, and delivery, and the associated staffing, consumables, direct, and indirect overhead costs. In Canada, the acquisition cost of a red blood cell unit is $419 and a unit of HLA matched (apheresis derived) single donor platelet unit is $1250 [6], as noted in Table 2. The Canadian unit costs are consistent with those in Western Europe, as studies from reported data from the United Kingdom, Sweden, Switzerland, Austria, and France showed a two unit cost of €877.69 [4, 7-12]. Moreover, with increasingly stringent pathogen screening technologies employed to ensure a safe blood supply, blood will become more costly to provide in the future [5, 13-15].

Table 2: Blood Product Acquisition Price (Canadian Dollars) Adapted from [6]

Blood product name	Acquisition Price (Canadian $)
Red blood cells*	$419
Autologous (whole) blood	$419
4 units buffy coat derived platelets	$286
1 unit single donor (apheresis) platelets	$619
1 unit HLA-matched single donor (apheresis) platelets	$1,250
Apheresis fresh frozen plasma	$360
4 units frozen plasma	$156
8 units cryoprecipitate	$1,080
IVIG per gram	$63
Octaplex® 1,000 IU	$720
Recombinant Factor VIIa/mg	$1,163
Anti-D 120 mcg	$33
Anti-D 300 mcg	$82

Thirdly, the blood supply is very scarce in absolute terms. For example, in Canada, blood product supply is derived from altruistic donations to two organisations – Canadian Blood Services and Héma-Québec - from approximately 12.5 million people [16], with an overall per capita donation rate of 35.8/1000 (3.6%). As the per capita issuance rate is 32.6/1000 (3.3%), the supply margin is very thin at a difference of 3.2 units/1000. Other countries, such as Australia (Australian Red Cross Blood Service), have similar rates of donation and issuance of 40.47/1000 (4.0%) and 36.77/1000 (3.7%) respectively [2]. The collection and utilization rates of other jurisdictions are outlined in Table 3.

Table 3: Blood Product Collection and Utilization Rates in a Subset of Economically Advanced Countries, (Adapted from [2])

Country	Blood product (annual collection rate/annual utilization rate per 1000 population)						
	Red cells	*Platelets, pooled*	*Platelets, PRP derived*	*Platelets, apheresis*	*Frozen plasma/fresh frozen plasma*	*Cryoprecipitate*	*Crow supernatant plasma*
Australia (Australian Red Cross Blood Service)	40.47, 36.77/1000 population	6.57, 5.65/1000 population			7.46, 7.16/1000 population	2.84, 2.73/1000 population	0.84, 0.78/1000 population
Canada (Canadian Blood Services)	35.8, 32.6/1000 population,	3.8, 2.3/1000 population	0.5, 0.4/1000 population	1.6, 1.5/1000 population	6.3, 5.6/1000 population	1.8, 1.0/1000 population	1.8, 0.7/1000 population
Austria	56.32, 51.39/1000	3.04, n/a		3.57, 3.21/1000	2.94, 2.63/1000 (quarantine) 17.72, 16.77/1000 (fractionation)		
Héma-Québec (2008-2009)	31.7, 29.9/1000			3.4, 3.2/1000	1.2, 0.9/10001.2, 0.9/1000		
Finland(2007)	51.9, 47.7/1000						
France (2008) (Établissement Français du Sang, EFS)	36.9, 35.4/1000			7.7, 2.9/1000			
Hong Kong Red Cross Blood Transfusion Service (2007) (HKRCBTS) 香港红十字会输血服务中心	29.06, 26.36/1000						
Italy (National Blood Centre)	40.9, 40.5/1000						
Singapore (2008)	19.09, 15.83/1000						

Fourthly, blood products possess an ethical cost, especially when they are altruistically donated, such as in Canada. The ethics of blood donation is multifaceted and will not be further explored in this research; however, it is noted that blood must be treated with respect, not only because

of its scarcity, but because a donor's intent should be respected. Blood donations in many countries such as the United States, Scotland, Canada, England and Wales are volunteer unpaid donors, and it is suggested that the voluntary nature of this makes it less likely that a donor would lie about their health status, and thus results in a theoretically superior and more rare blood product [6, 17]. For instance, a World Health Organization (WHO) study in the mid 2000s highlighted that only 49/124 (39.5%) countries investigated have 100% voluntary blood donation, with 17/49 (34.7%) being developing countries [71]. To highlight the advantages of voluntary donation, it is noted that in South Africa, HIV has a prevalence of 0.03% in regular blood donors, but the HIV prevalence is approximately 23.3% in the adult population [71]. It is this voluntary nature that results in low donation rates, but increases the safety of the blood supply. Every single donated unit therefore represents a voluntary and altruistic contribution to someone else in need, and the donor's time and effort, which they could have spent doing something else [17]. The units that are collected altruistically are therefore a much more privileged and precious subset of an already scarce resource. Therefore, these units must be handled and treated with the utmost respect, bearing in mind the donor's wish to have that donation given to somebody who needs it.

Finally, there is a biological complexity to the blood transfusion inventory which serves to amplify the issues of scarcity, and thus functional availability for recipient needs. There are physiological limitations that prevent universal transfusion as not all blood product types can be universally interchangeably transfused to all recipients. This limitation is due to naturally occurring antibodies (isohemagglutinins), which have the potential to cause intravascular red blood cell destruction (hemolysis), in the recipient when a blood product is transfused. These naturally occurring A and B isohemagglutinins are present in all humans to A and B antigens they do not have. For example, a person with a B blood group will have Anti-A in their plasma, and a person with an O blood group will have Anti-A and B. It is also typical practice that red cells bearing RhD positivity are not transfused to recipients having RhD negativity. Overall, there are different permutations and combinations of these compatibilities, and they are outlined in Table 4 and Table 5 for products that consist of red cells or plasma (platelets and plasma) primarily.

Table 4: ABO and RhD Compatibility of Red Blood Cells

Donor (RhD)	Recipient (RhD)
A-, O-	A+/-
B-, O-	B+/-
AB-, A-, B-, O-	AB+/-
O-	O+/-
A+, O+	A+
B+, O+	B+
AB+, A+, B+, O+	AB+
O+	O+

Table 5: ABO Compatibility for Plasma and Platelets

Recipient ABO group	Donor ABO group
O	O, A, B, AB
A	A, AB
B	B, AB
AB	AB

Therefore, even though units of red blood cells are essentially physiologically equivalent, the O+ and O- units are functionally more valuable, often referred to as the 'universal donor' blood group, as they can be given to any recipient blood group. This is in contrast to AB red cells, which are only limited to the AB blood groups respectively.

2.2 Best Practices in Transfusion Inventory Management

Inventory best practices and optimized inventory performance are ultimately determined by the measurement of quality metrics known as Key Performance Indicators (KPIs). These KPIs can be divided into two different types of values, with one measuring inventory size and the other measuring inventory wastage.

The values related to inventory size KPIs can be measured as absolute numerical quantities or expressed as a number of days that an inventory will last. With regards to the category of red blood cell (RBC) units, one of the main commonly accepted and understood descriptors of inventory is the inventory size, measured in the absolute number of units. At the Central Zone Blood Transfusion Service (CZBTS) at the Nova Scotia Health Authority (NSHA), a functional amount of inventory is measured based on the available number of units plus one half of the units that are allocated for patients (crossmatched). In the literature, other RBC inventory size KPIs include the issuable stock index (ISI), defined as the "measure of the days of unreserved (uncrossmatched) stock of all inventory blood groups" [18], a variation of the one used at our institution. Canadian Blood Services (CBS), for the purposes of benchmarking, has an Inventory Index (actual inventory count divided by the [sum of transfused + outdated + wasted units for three months/90]), which is used by CBS in their disposition reporting [72] to hospital end users. This Inventory Index number is nearly equivalent to another indicator that is used at the CZBTS - 'days of inventory on hand' -with the difference that outdated and wasted units are ignored for the purposes of our calculation. Overall, there are various KPIs used to measure inventory size, and not all of them are equivalent.

Inventory wastage KPIs are used to describe the net effect of blood inventory handling and management practices resulting in wastage. Wastage indicators are mostly expressed as a percentage, either of the total units transfused or total received, or they may be expressed as an absolute number. For example, in the Canadian literature if units are expired, or 'outdate', they are expressed as a percentage of the total units transfused (outdate %) [20]. In other literature, wastage is expressed as Wastage As a Percentage of Issuance (WAPI), which is the number of units that are wasted for a particular reason divided by the total number of units received as a percentage [18, 21]. Wastage can be sub-classified into units which expire (outdate) vs. those that are wasted through non-expiry reasons, such as product manipulations (i.e. bag spiked but subsequently not transfused), accident (i.e. broken bag), or other causes (i.e. visual inspection failure, temperature deviation out of range or non-conformance).

Table 6: Key Performance Indicators Related to Inventory, Calculations and Usage Cases

Category	Key Performance Indicator	Calculation	Usage
Inventory Size	Issuable Stock Index (ISI) [18]	Number of unreserved stock of red cell units of all age groups divided by the number of red cell units dispatched to the hospital in 1 year divided by 365	Determines the number of days of unreserved inventory in stock.
Wastage	Crossmatch: Transfuse ratio [19]	# crossmatched RBC-containing units/# transfused RBC containing units	Used to examine the efficiency of units that are held in reserve as a function of what is transfused. A high ratio indicates that too much blood is tied up.
	rate of RBC unit expiration [19]	(# expired nondirected allogenic RBC units/# expired and transfused Nondirected Allogenic unitsX100)	A rate of all units expired (due to due date being exceeded) as a function of essentially all units which exit the inventory permanently (excludes other causes of wastage)
	rate of RBC unit wastage [19]	(# wasted RBC units/# transfused and wasted RC unitsX100)	A rate of all units wasted (due to due date being exceeded) as a function of essentially all units which exit the inventory permanently (excludes expired units)
	Wastage as a percentage of red cell issues (WAPI) [18]	Hospital total red cell wastage/red cell issues from blood service X100	Rate of hospital wastage as a percentage of total units sent to the hospital from the supplier
	%RBCs outdated [20]	Number of RBC units outdated/Number of RBC units transfused+Number of RBC units outdated X100	Similar to rate of RBC expiry noted above [19].
	RBC Unit Expiration Rate (for nondirected allogenic RBC-containing Units) [19]	# expired Nondirected allogeneic RBC-Containing Units/(# of transfused nondirected allogeneic RBC-containing units + # of expired nondirected allogeneic RBC-Containing units)X100	A rate of all units expired (due to due date being exceeded) as a function of essentially all units which exit the inventory permanently (excludes other causes of wastage)
	RBC Unit Wastage Rate [20]	# of Wasted RBC-Containing Units/(# of Transfused + Wasted RBC-Containing Units)	Reflects rates of units that are wasted due to handling and storage errors (i.e. temperature), but are not limited to that.

With respect to KPIs, ideally, best practices achieve minimal wastage rates while ensuring the inventory is available for the needs of the eventual recipients in a timely fashion [2, 19-21, 38, 44, 57]. Transfusion inventory management best practices can be divided into internal and external practices and factors.

Within internal inventory factors, it is best practice to keep the correct inventory size according to the needs of the institution. The absolute transfusion inventory size is not directly discussed in the literature at all, it has been determined that there are too many variables to consider, and it is too computationally complex to be described by a mathematical model [73], though it has been attempted historically [28]. While the exact calculation of an ideal inventory size is not

possible, it is understood that the size of the inventory affects wastage. Perera *et al.* [18] describe a positive correlation between the size of a hospital's inventory level and the wastage rate, expressed as issuable stock index (ISI) and wastage as a percentage of issuance (WAPI). Conversely, there is an inverse relationship between inventory size and shortage rates [28]. The best practice of issuing the oldest unit of blood first, known as the First In, First Out principle (FIFO) vs. Last In, First Out (LIFO), has been long established to minimize shortages and minimize items reaching the end of their shelf life [2, 27]. Another best practice to reduce wastage is to ensure that the residual shelf life (RSL), or the freshness of the blood shipped from the supplier is as high as possible since RSL is inversely correlated with outdate percentages as suggested by Pereira [29]. The redistribution of blood stocks within an inventory is also another best practice in the literature. The theory is that sites transfusing smaller numbers of units can transfer inventory to larger blood centers with higher turnover and reduce outdate rates [31, 32].

Best practices in inventory management also include mitigating factors external to the inventory itself. Although it is difficult to control the environment to impact demand, there are factors that can be potentially influenced which can improve KPIs. A factor impacting inventory wastage and shortage rates is the coefficient of variation of daily transfusion (CVAR), measuring the variation of the number of RBC units transfused per day. The CVAR is known to be exponentially correlated to outdate rates according to Pereira [29]. Therefore, if there is increased variation of usage of red blood cell stocks, there will be a corresponding increase in outdate rates. The CVAR is a variable which cannot be controlled, though theoretically an institution could smooth out demand by distributing services which demand blood throughout the week and potentially impact wastage. Another best practice relates to the concept of cross matching, which allows for the assignment of inventory units for specific recipients, and makes them unavailable from the overall inventory. The concept of these assigned sub-inventories is what makes the transfusion inventory unique from all other perishable goods inventories [29]. These 'crossmatched' inventories are typically reserved for a specific period of time, known as 'crossmatch time' [29] and are unavailable for other recipients to use during this state [29, 74, 75]. They can only be released for others' use through manual release, or when the crossmatch period is over. To reduce wastage and reduce inventory size, it is best practice to reduce cross matching rates and cross matching time periods [18, 30]. Perera *et al.* [18] demonstrate that by

using intra- and post-operative cell salvage techniques, a reduction in crossmatching rates will result in reduction of inventory size in response to reduction in blood requirements. Modern day electronic crossmatching has similarly reduced the size of the inventory [29] through reduction of the need to allocate this subinventory.

Those best practices discussed above have existed since the beginning of transfusion inventory management literature nearly 50+ years ago. In the more recent literature of the past 20 years, other best practices have emerged. In a prescient article by Hirsch prior to computerisation in transfusion medicine practice, he notes that it could be possible that blood distribution systems could become reactive and predictive, and share information [76]. One of the most pervasive practices globally is the ubiquitous use of computerised systems, such as laboratory information systems, for transfusion product documentation [2]. An example of a regional information system for sharing information is the VANESA database (www.bloodstocks.co.uk) which stands for Vital Appropriate Knowledge (silent K) Empowers Stock Analysis, implemented in England, Wales, and Northern Ireland. VANESA collects and disseminates blood supply information (i.e. stocks, wastage rates, transfused units) via the Internet to various participating hospitals. Through the VANESA information system, information about the supply chain can inform demand forecasting, determination of appropriate inventory levels, and contingency planning [24, 33, 34]. More recently, there have been efforts to improve inventory management through the use of dashboards which are derived from real-time laboratory information system data for recycling units and possibly predicting blood product usage [32, 35]. At the CZBTS, we have also meaningfully reduced red cell outdate rates to near zero by creating and implementing a novel red cell inventory ordering algorithm as of June 1, 2015.

Another best practice is the application of quality improvement and process reengineering programs, such as Lean and Six Sigma. These have allowed enhanced monitoring and control/audit of internal and external processes and influences on inventories. In the area of process reengineering, Heitmiller *et al.* described targeting RBC product wastage, including in-date discards, with Lean Sigma techniques, resulting in $800,000 in savings over a four year period [38]. This study has limited applicability as it was noted that a factor related to the initial

drop in wastage was from the Hawthorne effect, which relates to the improvement in behavior of subjects from being merely observed, rather than secondary to the experimental manipulation. Heitmiller *et al.* indicated that there were other interventions which were temporally coincidentally introduced, and thus the impact of a particular intervention could not be fully quantified [38].

Finally, one of the most critical best practices is to document the performance of the inventory through the use of KPIs or quality metrics. These metrics are used to understand and describe an inventory and are used for benchmarking purposes against peer institutions as a crucial first step in improving process [20, 22, 23].

2.3 Issues Facing Modern Transfusion Inventory Management Practice

There are various issues facing the management of blood product inventories [1, 2, 57] and they can be broadly divided into inventory related issues or systemic issues.

Transfusion inventories have to cope with a trend of decreasing supply and increasing demand. As noted in Table 3, the margin between supply and demand is usually measured by a few percentage points, highlighting that there is significant supply tightness globally. An example of the tenuous blood supply is by Drackley *et al.* [65] in the context of the blood supply and demand in Ontario, Canada. The paper suggests there is an organic shift in the current donor base of older adults to one that is younger and more culturally diverse, resulting in less donations. This demographic change has been seen in other countries also and as the adult population ages, it shifts to become a user of blood [77-84]. Through modeling different scenarios, Drackley *et al.* suggest that in Ontario, the demand and supply curves could cross as early as 2012, and though they highlight different scenarios to forestall that crossover, the intent of the paper was to suggest a tenuous blood supply in the context of demand in Canada's

most populous province. Other supply shocks could involve emergent transmissible disease (i.e. Zika virus), unpredictable quality assurance issues, weather-related supply disruptions, unforeseen factors [57-70, 85-87], changes in donor selection policies [88, 89], and a blood supplier's response to changes in infectious diseases [66, 90, 91]. Demand for blood products is also increasing, with an increasing proportion of hospitalisations requiring blood transfusion [92]. In the case of platelets, increasing demand is due to population age and number, increasing rates of hematological malignancies and the changes in treatment of such malignancies [93, 94]. Finally, transfusion inventory management will have to keep up with a changing demand profile, where inventory may not be able to meet the demand in specific blood groups [95]. Supply and demand are not the only factors which are exerting pressure.

Transfusion inventory management must also cope with the financial costs of transfusion. Globally, there are numerous drivers of increased costs, including the potential adoption of technologies such as pathogen inactivation, in addition to the already-established costs of providing blood products [3, 4, 7, 57, 96-100]. In Canada, fluctuations in the USD/CAD and CAD/EUR exchange rates in the past 2-3 years have caused acute cost pressures on CBS, with rising blood product prices inflating budgets by at least 30-40% (CBS internal data, National Liaison Committee). This has likely impacted other countries as well, but it is not yet documented in the transfusion literature. Another impact that is not mentioned in the transfusion literature is the concept of inflation, which could be considered an issue which faces blood transfusion practice. For instance, the Bank of Canada (bankofcanada.ca) aims to keep its inflation rates at the midpoint of the 1-3% range, which is the year-over-year increase in the Consumer Price Index (a cost-of-living measure for Canadians). This would also translate into a blood product cost increases, as the increasing cost of living eventually translates to increasing wages and input costs, which are components of the blood product supply chain [5]. By increasing the cost of blood products, the opportunity cost of maintaining an inventory would also increase commensurately [1].

Another issue that faces transfusion service inventory management is an adequate response to the various financial, supply, and demand pressures. One way that transfusion services can

reduce costs is to reduce the input costs of the transfusion process itself [5]. As an illustration, if two transfusion services have similar testing methodologies or can make use of a single facility rather than two, combining the two services into one service in one facility could potentially be an economy of scale, and result in cost savings to the blood product supply chain. Inventory would be impacted as it would become more centrally managed. This trend to centralize transfusion services continues to be identified as a way to increase efficiency, and it is increasingly practised globally [2, 101, 102]. In addition to the centralization of physical and administrative infrastructure, transfusion service information technology infrastructure is also following that trend, with benefits in safety and better documentation by using more centralized transfusion databases or laboratory information systems [103, 104]. With centralisation, transfusion services and their inventories will become increasingly more complex entities to manage due to their size and geographical coverage, as a single or few hospitals will manage a cluster of hospitals, each with its own personnel, product inventories, clinical issues and consultations, and with transactional and bureaucratic interplay between them. Furthermore, there will be a need to assimilate and integrate new technologies and new ways of thinking, such as positive patient identification [105, 106]. As there will be increasing trends to automate the transfusion laboratory, other data sources from instrumentation, middleware, and other databases will increase the complexity and amount of data gathered [107-110].

Transfusion service inventory management is also facing barriers relating to communication of practice and data silos associated with separate information systems and management practices. One of the issues highlighted as a significant barrier to understanding avoidable wastage was a discontinuity in the data. Devine *et al.* [2] explains it by this quote:

> *"A strong recurring theme in these responses is the presence of gaps in the information required to really have an optimal inventory management system. Many providers of blood products have only one end of the rope: that which is attached to the inventory around the production and distribution processes. Very few respondents had information back from the hospital customers as to the disposition of the products provided in terms of when or even if they were transfused. This lack of information hampers the ability to drive down the rate of avoidable discards or to know if the manner in which the blood provider manages inventory is optimal for patient safety or ready access to transfusion products."*

While this 'rope' discontinuity has been identified, this issue continues to be unresolved in the transfusion literature after the Devine *et al.* publication in 2010. This discontinuity still persists (CBS Platelet Stakeholder Meeting, 2017, meeting notes). Another issue that has not been named, but that is data-related, is the variety of KPI definitions and calculations. If the goal of global transfusion services is to improve efficiency and communicate the most efficient best practices [2], then benchmarking is important in identifying and surfacing leading practices to follow [20, 22-24]. One of the issues facing the transfusion community is the variability in how KPIs are calculated. This will affect global efforts to translate best practices as it will be difficult to translate a methodology from one system to another system, as the measured KPIs may be different. Ideally, the transfusion community should come to a uniform consensus on how these KPIs (Table 6) are to be calculated.

The last issue facing blood transfusion services is abstract and philosophical, and involves a conceptual ceiling that has been reached in the transfusion community's ability to understand the inventory. This likely relates to the maturity of the tools that are being used and to the traditional thinking that is being applied to inventories currently. Despite the plethora of analytical techniques since the 1960s [25, 26], there are very few universally practised inventory management concepts other than FIFO [2]. There is no mention of universally applied techniques other than the pervasive and implied use of KPIs. One can speculate that the concepts illustrated in the literature have not been generalizable to other institutions, or that potentially some of the techniques are so complex that they are not applicable to real-life day-to-day inventory management practices. Such is the case of modelling or simulation techniques, which have required significant engineering or mathematical expertise to interpret and understand their theory and outputs, and suffer from their inability to truly describe a real-life process [25, 26]. With the ubiquitous computerization of the transfusion service, perhaps other techniques such as those associated with the 'big data' movement can help further understand these increasingly complex data sets [111]. The concept of thinking about an inventory, and the potential lack of evolution in that thinking, may be silent issue facing inventory management today. This current global thinking is noted in Devine *et al.* [2] as they consider the transfusion supply chain ideally informationally contiguous, and that the breaking of that 'rope' with discontinuous data between the supplier and hospital information systems, could be responsible

for the inability to prevent avoidable wastage. An alternative way of thinking about the blood supply chain is expressed by Beliën *et al*. [25], with the following quote, "…*that it might be worthwhile to investigate how the blood supply chain problem can be decomposed into different sub problems".* Rather than understanding the supply chain in its informational breadth, perhaps the supply chain should be understood in its informational depth. Another way of thinking about inventories, and which is not expressed in the transfusion literature, is the concept of complementing the current value-based thinking with process-based thinking by applying process-based analytical tools [49]. Currently, the paradigm in thinking and description of real inventories is through KPIs, essentially understanding inventories from a value-based perspective. The blood transfusion community describes the complexity of the 'storage in hospital' box from Figure 1 with a value describing the size of the box, how long the box will last, and provide a value of outdates. Conversely, if one uses a process-based perspective on the box, the smaller boxes of "unassigned inventory", "crossmatch", "assigned inventory" and "outdate" and the pathways that connect these smaller boxes could be characterised (Figure 1). Although this would not replace the traditional KPI-value-based view, it would deepen how inventories are understood. Conceptually, if transfusion services do not make a shift to thinking about inventory differently and apply different techniques to do so, inventory management will not evolve past the single universally practiced technique of FIFO distilled from over 50+ years of techniques in the literature.

2.4 The Concept of the Business Process Model/Map

A business process is defined as "*a specific ordering of work activities across time and place, with a beginning, an end, and clearly defined inputs and outputs. Business processes are the structure by which the organization physically does what is necessary to produce value for its customers and are broadly defined across functions/departments.*" Furthermore, a process model is defined to be "*a representation of one or more processes and their associations that an enterprise performs. Process modeling is a mechanism for describing and communicating the current or intended future state of a business process*", and a business process model to be "*a*

mechanism for describing and communicating the current or intended future state of a business process." (https://www.ca.capgemini.com/).

Figure 2 depicts a business process model (process map), based on a hypothetical business process of handling a loan application. In this diagram, loan applications begin with the registration of a client and can take two pathways: Check ID and Initiate Transfer to another institution. If the ID is checked, the client is rejected and the file is closed or, if not rejected, the transfer is initiated. Initiation of transfer can also occur after registration of client.

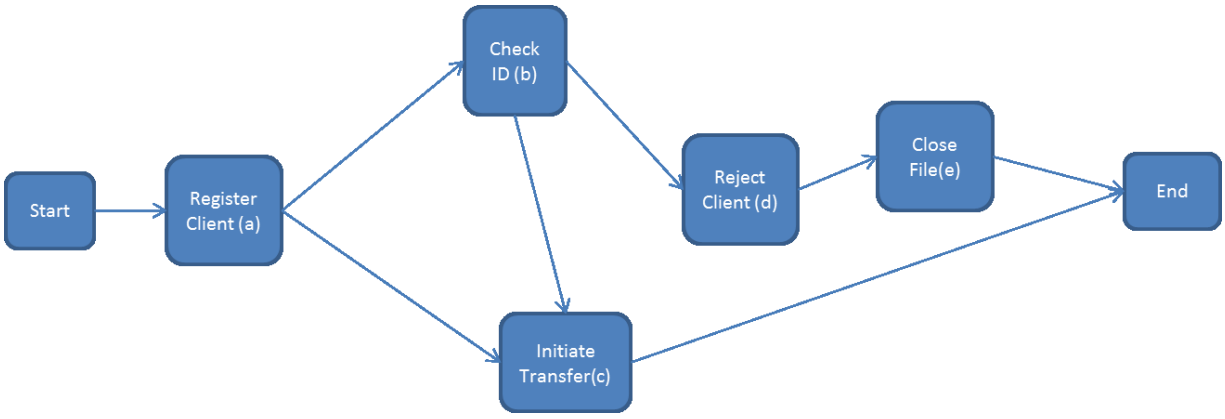


Figure 2: A business process model for a hypothetical loan application business process.

These activities are connected in a defined sequence and this sequence of activities and pathways is what defines this particular level of 'zoom', or abstraction, of the process model. The process model can have a different level of abstraction as various activities, such as the client rejection activity, can be composed of sub-processes (not shown). Overall, the process model is a flow of activities and sub-activities that have causal dependency on one another, may include time properties, and may also include resources that are used to effect those activities (i.e. *Jane Smith* initiates the transfer) [47].

The creation of a business process model is crucial to understanding a process, examining planning and control decisions, and allowing the process to be reviewed with the purposes of redesigning it. The creation of a process map is both science and art, is subject to error [47], but can be done either manually or through process mining. Creating a process model using manual methods requires experienced human resources and can have various pitfalls. The first includes oversimplification of reality, whereby a model would represent only common events but not the entire event space. The second potential pitfall is to minimize the impact of or underrepresent human behavior, which may cause the process map to be more variable than, for example, if a process was done by machines. The third pitfall is that the map is created and locked-in with the inappropriate level of granularity or abstraction, such that there isn't an ability to change the 'zoom' level if there isn't enough granularity (process map too simple), or there is too much granularity (process map too complex). If the resolution is incorrect, questions pertaining to another resolution or abstraction level of the map may not be answerable [47]. Finally, if a process model is to be derived through observation of a process, it may be subject to the Hawthorne Effect, which is a well-known psychological phenomenon where there is a change - often an improvement - in worker productivity due to simply being observed [112]. Therefore, in order to efficiently generate a process model, a less manual and more automatic/data-driven methodology is preferred.

2.5 Introduction to Process Mining

In the 1970s and 1980s, data was at the heart of information system design, and business process models were at the periphery; however, in the early 90s, process reengineering techniques became mainstream as did process modeling techniques. This necessitated a powerful trend in information systems design, which allowed the evolution from a data-orientated perspective to a process-oriented perspective [113]. The discipline of process mining arose at that time, with seminal work in 1999 by Van der Aalst in the Netherlands with a project entitled "Process Design by Discovery: Harvesting Workflow Knowledge from Ad-hoc Executions" [114].

Process mining is defined as a technique which is '*able to extract knowledge from event logs*' and '*sits between computational intelligence and data mining on the one hand, and process modeling and analysis on the other hand*' [115].Process mining allows real processes to be discovered, monitored and improved by extracting knowledge contained in event logs which are commonly available in modern information systems [115]. Since process mining uses real data about a real process for construction of a process map or model, it avoids the pitfall of manually created process maps behaving differently in reality or during their execution.

The core philosophy of process mining is to generate process knowledge from already-logged data, and relates to several other areas of business process management and business intelligence as noted in Figure 3. Process mining is a subset of business intelligence, defined as turning data into information and then into knowledge. Business intelligence allows enterprise data to be analysed and improve how decisions are made, thus increasing the competitiveness of businesses [116]. Process mining is a subset of a subdomain of business intelligence known as Business Process Intelligence [117], which includes tools which assist in process analysis, prediction, control, monitoring, and optimisation. Process mining also crosses over into Business Process Management, defined as techniques which allow for the design, enactment, control, and analysis of operational events involving humans, organisations, documents and applications [118]. Process mining overlaps with some aspects of Business Process Analysis, whereby it plays a role in *a posteriori* process diagnosis, and Business Activity Monitoring, where real-time process management can occur using dashboard-like tools [48]. These terms and their overlaps are summarised in Figure 3.

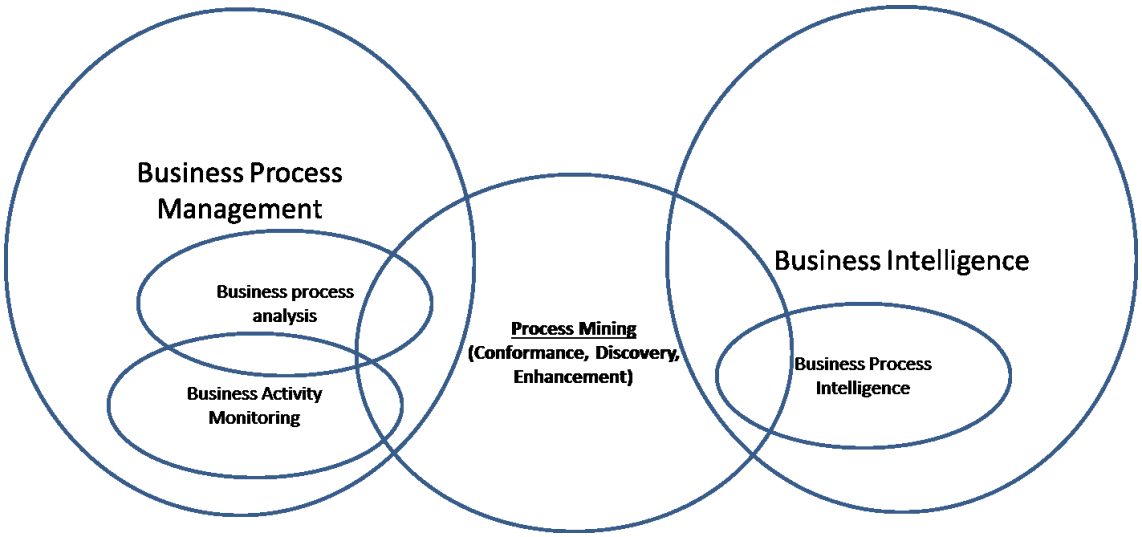


Figure 3: Process mining Venn diagram in the context of business process management and business intelligence (adapted from [48])

Data mining is data centric, while process mining is process-centric [119]. In contrast to conventional data mining, whose algorithms are generally purposed to associate events that are more frequently correlated with each other, these algorithms do not consider or visualise an end-to-end process. Conventional data mining algorithms, in the context of a process, can only identify localized variations within a larger process model, and cannot re-create that process model with parallel vents, loops, choices, and other variations [47, 119]. Due to such differences in philosophy, process mining is a specific discipline that exists outside of traditional data mining, as it examines the entire process, not just associations between events such as in conventional data mining.

2.5.1 Process Mining Methodology Framework

Process mining plays an integral part in many aspects of business process improvement. Firstly, it is used in process discovery, whereby a process map is created without a-priori knowledge of the process itself. Secondly, process mining is used is in regards to conformance checking. This is where an existing model of the process (or business rules and policies) is compared to the

reality visualised within the process map and vice versa, and is a measure of the adherence of the real-life process to the model. Thirdly, process mining is used is for enhancement. Enhancement refers to augmentation of the existing process model with useful information which can allow for better interpretation or possibly alteration of the model. This can be through the overlay of calculated and aggregated timestamp data to quantify bottlenecking and understand throughput times in a process [115]. This is outlined in Figure 4, and has been adapted from the Manifesto [115], which is a document that the Institute of Electrical and Electronic Engineers Task Force on Process Mining published to given an overview of the concept of process mining.

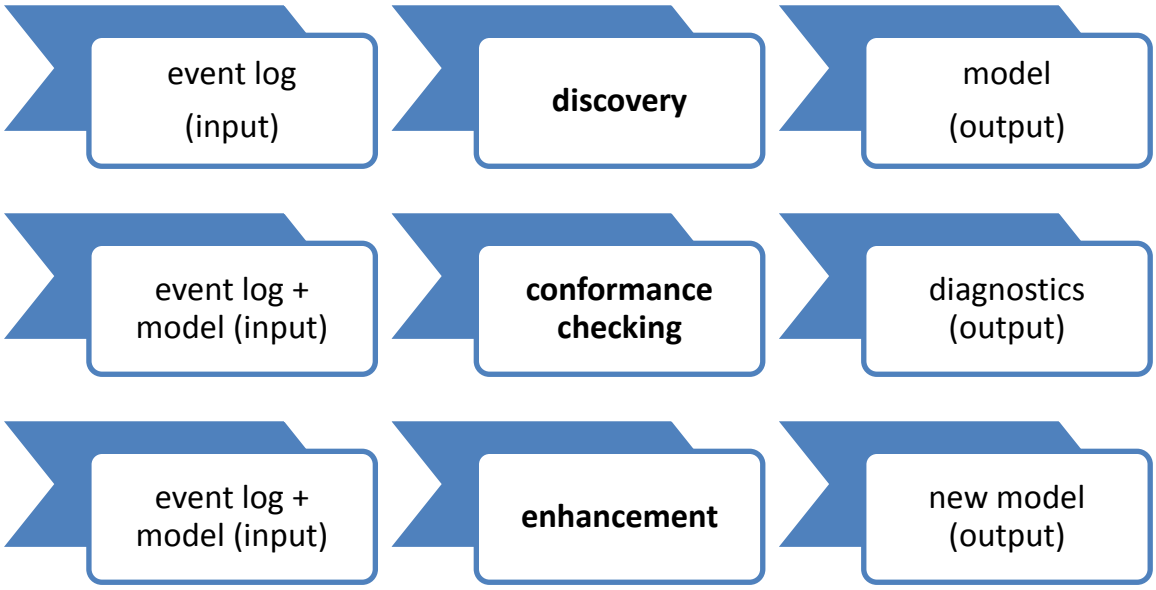


Figure 4: Comparisons of inputs and outputs in the context of general process mining use cases (adapted from [115])

In order to guide process mining research, a framework must be put forth on how to deploy the process mining methodology. A common way to integrate process mining into practice is to conduct a case study using real data. The Process Mining Methodology Framework (PMMF) put forward by De Weerdt *et al.* [48], is a general framework for the performance of a case study in

process mining, and is consistent with other approaches in the literature [51, 52]. Figure 5 has been adapted from De Weerdt *et al.* and outlines this framework [48].

Data extraction and preprocessing are the first two steps of the data preparation building block of the PMMF. The goal of this building block is to obtain the appropriate scope of data in terms of timeframe and quantity. There needs to be enough data to make an interpretation of the process; if there is too little data, the process may be oversimplified; if there is too much data, then computational times become an issue or the process map can be difficult to interpret. The scope for data extraction according to the PMMF should consider only data from one process, rather than obtaining data from overlapping or multiple processes. Another aspect of the data extraction is to ensure that an adequate timeframe is considered. If the time frame is too large, then the amount of data can be too expansive and increase computation times. Likewise, if the timeframe is too narrow, or of an incorrect time frame, then rare or seasonally-related events can be missed. The data preprocessing component is closely associated with data extraction, ensuring that the data contains the correct fields and format, in addition to validation and conversion to a correct file format if required.

The second building block of the PMMF is process data exploration. This further refines the data preparation portion and allows for a 'first impression' of the data that has been gathered, including generation of statistical data and preliminary process maps using different mining algorithms. There is a feedback loop between this exploration step and the data preparation step, as the scope and timeframes of the input data can be further refined to optimise the output. This feedback loop can occur many times until the data is suitable for further analysis. After process data exploration, there is a block which allows for the 'perspective' of the analysis to be chosen.

There are three possible analytical pathways, all contained within the process log data, which can be pursued: 1) control flow data analysis, includes examination of activity sequences and decisions; 2) case data involves examining the performance and frequency of business activity; 3) organisational analysis involves examining data from an organizational point of view, examining how teams work together. Typically, one of these analytical pathways is chosen and followed through to in-depth analysis along a compliance or performance pathway. Compliance verification validates whether reality, based on the data extracted, reflects what is expected

from a process. Performance analysis involves examining the impact of process execution over time, and potentially can identify inefficiencies.

The culmination and the output of these steps is the results building block. In this, 'guidelines for process improvement' should be the goal of the process mining exercise. De Weerdt *et al*. indicate that this PMMF is intended for the "*analysis of process inefficiencies within services organisations"* [48], and validate the use of the PMMF using a case study involving back office processes in a financial services context. This PMMF will be a starting point for this research, as another framework will need to be created that is tailored to the use of process mining for transfusion inventory data.

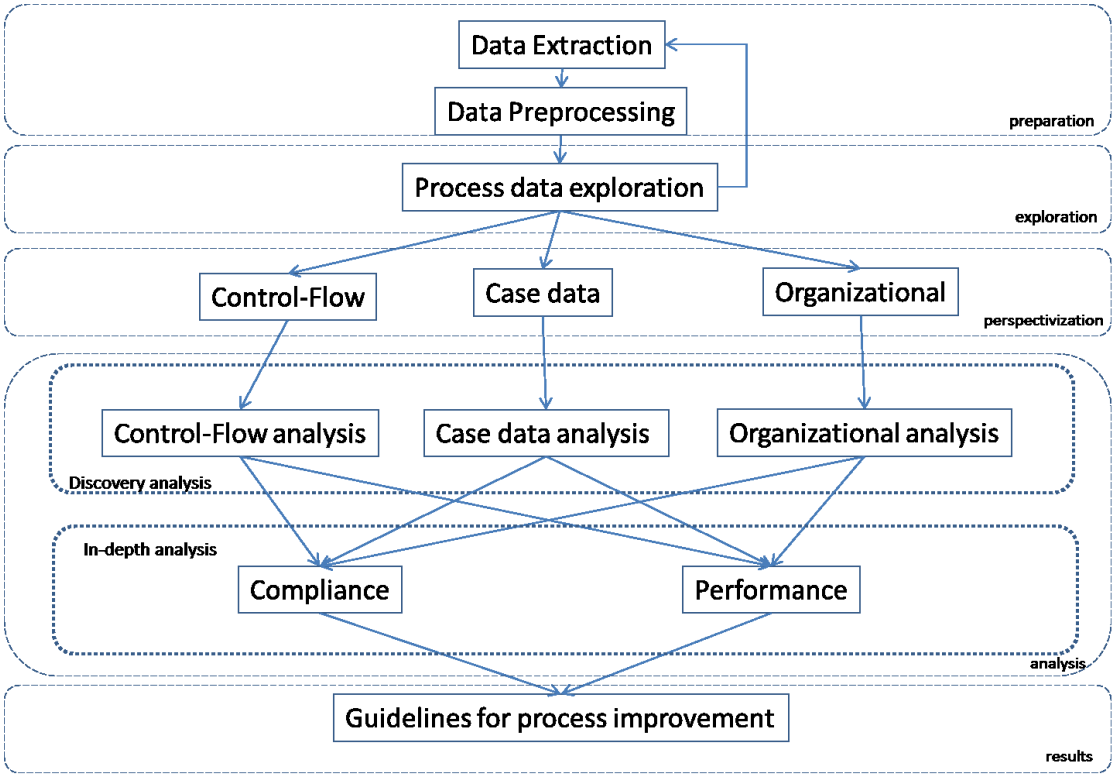


Figure 5: Process Mining Methodology Framework adapted from De Weerdt *et al*. [48]

2.5.2 The Process Log – Where Process Mining Starts

There are different data sources and different types of data quality that can be used as an input for process mining, and others that should not be, and these characteristics are summarized in

an event log (or process log) table adapted from [115] (Table 7). In order to fully benefit from process mining, the organization should have data which is of sufficient quality, corresponding to a minimum requirement of a 'level 3' standard according to the event log maturity table (Table 7).

Table 7: Event Log Maturity Table with Event Log Characteristics, Examples, and the Suitability of Use as an Input for Process Mining (Adapted from [115])

Level	Characteristics	Examples	Suitability as Process mining input (yes/no)
5	Trustworthy, complete data, well-defined events. Systematic automatic recording of events. Recorded events and attributes have clear semantics.	Semantically annotated logs of business process management (BPM) systems.	yes
4	Systematic automatic recording of events. Process instance (case) and activity are explicitly supported vs level 3 below.	Event logs of traditional BPM/workflow systems	Yes
3	Events recorded automatically; no systematic approach. Guarantee that the events match reality. Events may need to be extracted from different tables but the information assumed to be correct.	Event logs of customer relationship management (CRM) and enterprise resource planning (ERP) systems, transaction logs of messaging systems, event logs of high-tech systems.	Yes
2	Events recorded automatically, but no systematic approach to recording all events. Possible to bypass information system, therefore events may be missing or not recorded.	Document and product management systems, error logs, worksheets.	No
1	Poor quality event logs. Recorded events may not correspond to reality and may be missing. Manual recording.	paper-based charting in medical records, sticky notes.	No

Process mining begins with the process log, a sequential collection of events or activities, as the input. This data can be found in error logs, message logs, transaction logs, and other database tables. Moreover, if the data does not exist natively inside a log file it can be compiled or built from transactional data found in different data tables. The existence of this log data can be either intentionally captured or can exist as a byproduct, though the former is preferred. Process logs can originate from process aware information systems (PAIS) and process unaware information systems. A definition for a PAIS could potentially be "a software system that manages and executes operational processes involving people, applications, and/or information

sources on the basis of process models", as proposed in Dumas *et al*. [113]. The difference is that process aware systems have explicit process notation hard-coded into the information system and the system is actively involved in managing the business process. This can include workflow management systems (WMS), enterprise resource planning (ERP) or customer relationship management (CRM) systems [47, 113].

The core components of a process log must contain activities (well-defined process steps) and a case ID (i.e. unique process instance). Timestamp variables documenting when an activity/case ID combination occurred and other variables such as a resource (i.e. person or object that actioned that activity), cost, location, and other attributes may be helpful to separate two identical activity/case ID combinations [47]. Table 8 outlines basic definitions and terminology for data elements found in process logs.

Table 8: Process Log Related Terms, Synonyms, Definitions, and Examples Based on Table 9 and Figure 2 of a Hypothetical Business Process.

Term	Synonyms	Definition	Example
Event	Activity	A well-defined step inside a process.	Start, register client, reject client, close file, check ID, initiate transfer
Event ID	Activity ID	A unique identifer for an event, which is similar to a primary key in a database.	998374982374, which uniquely identifies a particular activity or event record in the process log.
Case	Process instance	The entity being handled by the business process being examined.	The entity of 1234, as it interacts with the loan application business process in Figure 2
Case ID	ID Case	A unique identifier for a case	1234
Timestamp		Time and date annotation of when an activity occurs.	July 23, 2014 13:10
Resource	Actor Perfomer Originator	A person, device, or entity which executes or initiates an activity.	Jimmy, Valin, Judy, James, Kiosk
Trace	Variant	A sequential set of activities pertaining to a case	For case ID 1234: register client → check ID → reject client → close file.

Applying this to the hypothetical loan application business process in Figure 2 and Table 8, and focusing on case ID 1234, one will notice that this process instance has four different event IDs

annotated by four different timestamps. There were three resources used, Jimmy, Alvin (twice), and Judy. It appears that the case has a unique trace (variant), not followed by other cases, in contrast to case ID 2678 and 3890, which have identical traces but different resources are involved. Case 4000 only involves the registration and initiation of transfer, and this is all done using a kiosk resource.

Table 8: Simplified Ordered Event Log of a Hypothetical Business Process

Resource	Case ID	Event ID	Timestamp	Event
Jimmy	1234	998374982374	23-Jul-14 13:10	register client
Alvin	1234	998374982375	24-Jul-14 12:45	check ID
Alvin	1234	998374982377	25-Jul-14 22:34	reject client
Judy	1234	998374982383	26-Jul-14 9:34	close file
Kiosk	2678	982734274321	20-Jun-14 3:40	register client
James	2678	982734274324	21-Jun-14 12:20	check ID
Jimmy	2678	982734274339	22-Jun-14 15:23	initiate transfer
Judy	3890	9884593845983	20-Jun-14 11:09	register client
Alvin	3890	9884593845992	20-Jun-14 11:34	check ID
Judy	3890	9884593845999	21-Jun-14 11:00	initiate transfer
Kiosk	4000	9234237428347	13-May-14 12:45	register client
Kiosk	4000	9234237428352	14-May-14 12:55	initiate transfer

2.5.3 Process Mining Algorithm Ecosystem:

The process mining algorithm ecosystem is important to know for reference, but this research will not be exploring the algorithms in computational or mathematical detail. Since process mining is very early in its development, most of the algorithms have been developed over the past two decades with only recent refinements to date [47]. The first foundational algorithms - RNet, Ktail, Markov – were developed by Schimm, Manila, and Meek, and were statistically and probabilistically-based and only able to compute a small portion of an entire process [120-124]. These earlier techniques will not be discussed in further detail, but are provided for reference. More recently developed algorithms include the alpha-algorithm and variants [125], heuristic

mining [125, 126], fuzzy mining [127], and genetic process mining [128]. These process mining algorithms are listed in chronological order of publication in Table 9 (adapted from [129]). De Weerdt *et al.* discuss the mathematical assessment of the quality of each algorithm in detail; however, that is outside the scope of the research [129].

Table 9: Process Mining Algorithm and Year of Publication

Year of Publication	Algorithm names
1998	General DAG [130], B-F(k,c)-algorithm [122], Rnet, Ktail, Markov [121]
2000	Global partial orders [131]
2002	Process Miner [123, 124]
2004	α/α+α/α+ [125], InWoLvE—splitpar [132]
2005	Multi-phase Miner [133], Workflow Miner [134]
2006	HeuristicsMiner [135, 136], DWS Mining [137], Rule-based approach [138, 139]
2007	Genetic Miner [128], DT Genetic Miner [140], Fuzzy Miner [141], α++α++ [142], DecMiner [143]
2008	AWS Mining [144]
2009	AGNEsMiner [145], ββ (or Tshingua αα) [146], Enhanced WFMiner [147], EM-approach [148], ILP Miner (Parikh) [149]
2010	FSM Miner/Petrify [150], FSM Miner/Genet [151]
2015	Multilevel Process Mining Algorithm (MLPM) [152]

2.5.4 Process Mining Software Packages

There are very few software tools that are available for process mining. The goal of this section is to outline the available freeware, educationally-licensed, open-source, and commercial process mining software available to use. These software packages are compared below in Table 10. This research will not present a detailed evaluation on these software solutions and services, and due to the quickly advancing nature of the field, this is a non-exhaustive list for reference purposes only.

Table 10: Process Mining Software Website and Licensure, Adapted from [47]

Name of Software	Website	Licensure
ARIS Process Performance Manager	www.softwareag.com	Commercial
Disco	fluxicon.com	Freeware, Educational, and Full License
Enterprise Visualization Suite	Businesscape.no	Commercial
Genet/Petrify	www.lsi.upc.edu	Academic
Icris webservices	icris.nl	Commercial webservice
Interstage BPME	Fujitsu.com	Commercial
Lexmark	Lexmark.com	Commercial
OKT Process Mining Suite	Exeura.com	Open source
Process Discovery Focus	www.iontas.com	Commercial
ProM 6.5.1a	promtools.org	Open-source
QPR Process Analyzer	qpr.com	Licensed software
Rapid-i	Rapidminer.com	Freeware and Licensed Rapidminer.
Rbminer/Dbminer	www.lsi.upc.edu	Academic
Reflect	www.futuratech.nl	Commercial
Reflect|one	www.pallas-athena.com	Commercial
ServiceMosaic	Soc.cse.unsw.edu.au	Academic

The open-source ProM framework is considered a tool with a reference library of process mining tools available, and is the grandfather of modern-day process mining tools. It is a framework of more than 90 plugins which pertain to the process mining workflow. The plugins include different algorithms, data import and export tools, analytical plugins, and converters. While ProM is not the earliest tool developed for process mining, as earlier tools such as EMiT [153], MiSoN [154]and Little Thumb [135]were available, the ProM framework allows the earlier tools to retain functionality as plug-ins in ProM.

The second pure process mining tool, Disco, is a Java-based software developed by Drs. Christian Günther and Anne Rozinat, and whose core algorithm is based on the fuzzy mining algorithm. The tool has been designed to be maximally useful for real-world process mining, with speed, non-destructive filtering, playback animations, detailed statistics, project management, and data

handling (importation/exportation) that is fully integrated and user-friendly (www.fluxicon.com). The fuzzy mining algorithm deployed within Disco allows a process map to be zoomed in and out to any abstraction level similar to a scalable map, like Google Maps.

2.5.5 Current Research on Approaches to Optimize Transfusion Inventory Management Practices

As indicated previously, modern blood transfusion services have an obligation to responsibly manage blood products - guarding against undersupply while minimizing wastage or oversupply – while operating in an environment of stochastic supply and demand.

There are numerous transfusion inventory management techniques in use today, and these are summarised by two major review papers in this field: Prastacos [26] in 1984 and Belien [25] in 2012. The two review papers have different approaches to evaluating the literature. Prastacos [26] uses an operations research-based approach to examine different methodologies that have contributed to blood transfusion policy and inventory management beginning with statistical studies, ordering policy determination using analytical and empirical approaches. Prastacos also examines various policy approaches implemented by different groups on indicators such as crossmatching parameters, freezing, shelf-life extension, and age of units or collection planning. Almost two and a half decades later, Belien [25] takes a slightly different multiaxial and functional approach, and is done categorically along different types of classifiers, such as blood product type, solution method, approach, type of problem, exact vs. heuristic, performance measures, and practical implementation or case studies. Some of the explored scenarios include evaluation/best practices, simulation, queuing models, statistical analysis, and 'what if' scenario analysis, etc. The paper also examines articles with respect to era of publication and the authors note that most of the papers involve red blood cell units, with two peaks of publication intensity: 1976-1980 and 2006-2010 [25].

While the category-by-category approach is appropriate for surveying the literature, this research will simplify the categorisation and group paradigm-defining literature into either direct or indirect methodologies.

2.5.5.1 Direct Methodologies Informing the Transfusion Inventory Management Paradigm

The concept of a *direct methodology*, as defined by this research, pertains to a technique that measures and represents a real-life inventory. This can include KPI-related methods, such as statistical analysis, or benchmarking, and is typically used in day-to-day practice due to the direct derivation from real data.

KPIs are direct methodologies commonly used as measures to demonstrate how effectively an organisation handles or achieves key business objectives. In transfusion medicine inventory management, the business objective is to optimise the blood inventory to maximise availability of quality blood products, reduce shortages, and reduce waste [2]. Table 6 (page 15) reflects KPIs which are commonly used in the literature, but these are non-exhaustive, as there may be KPIs used in day-to-day transfusion service inventory operations that are not documented in the literature.

The largest in-depth survey of key performance-based methodologies to date is a summary paper from the Q probes study of the College of American Pathologists. The goal of the paper was to document the experience of individual institutions in homogeneous and heterogeneous practice environments primarily in the United States (80%) [19]. Participants were benchmarked against each other using normalised rates of various KPIs such as crossmatch:transfuse ratio, rates of RBC unit expiration (# expired nondirected allogenic RBC units/# expired and transfused Nondirected Allogenic unitsX100) and rate of RBC unit wastage (# wasted RBC units/# transfused and wasted RBC unitsX100), and attempted to determine the effects of certain practices on those parameters. The study was a descriptive compilation of

three separate studies and used simple descriptive statistics (I.e. percentiles and p values) to arrive at the correlations between KPIs and practices [19]. One of the strengths of the study was the sheer size, as it involved 1639 hospitals and 12,288,404 red blood cell units. Devine *et al.* [2]approached a large-scale survey differently using a questionnaire method, with six questions pertaining to structure, product type, and they included KPIs which described donation rates (units per 1000 population), wastage rates, and age of red cells at transfusion. This was different than the Q-Probes study in that they also included blood suppliers in their analysis and there was less standardization of information as study participants were global in nature. Pink also used conventional statistical methodologies to report total inventory, ratio of average inventory to daily transfuse units, average age at blood issuance, hospital size, outdating rates, cross matching rates, but also examined the correlation with increased outdating in hospitals [155]

There were also papers using descriptive statistics to determine correlation between policies and various quality metrics [156, 157] or the reason for outdating of certain unit types [158]. Since those earlier papers, two papers by Chapman [24, 33] examined how the Blood Stocks Management Scheme (BSMS) was implemented in England, Wales, and northern Ireland, and how it served to collect and disseminate data on red cell stock and wastage rates (wastage as a percentage of issue), as well as providing performance-related feedback to participants through a web-based delivery mechanism. A limitation of these papers was that they were an overview of the method, but did not provide any conclusions on the efficacy or practical implications of the BSMS as an intervention, and they did not analyse real hospital data over time, despite five years elapsing between the two papers (2002-2007). In 2009, Perera [18] published a study using 2006 BSMS data and performed a statistical analysis to examine relationships between wastage and various KPIs such as Issuable Stock Index (ISI) defined as the 'number of days of unreserved stock of all blood groups held in inventory' (# unreserved RBC units of all groups/# RBC dispatched to hospital in 1 year/365) and Wastage as a Percentage of Issues (WAPI), (#RBCs wasted/total red cell issuesX100) and hospital practices, the latter of which was obtained through a questionnaire. These practices included many factors related to inventory, including how blood was ordered, how orders were calculated, crossmatching policies, inter-hospital transfers of units (i.e. stock share agreements), blood transfusion education, and blood

conservation measures. The paper concluded with correlations between lower inventory, lower ISI, shorter crossmatching times, and the use of computer-assisted ordering with lower wastage, as examples of best practices. One drawback of the paper is that the BSMS data was itself not used as the source of information, but that a survey needed to be sent out to elucidate best practices. Heddle *et al*. [20], during the same year, published a benchmarking study involving 156 hospitals. They used logistic regression techniques to identify factors affecting RBC outdating and found three factors: distance from blood supplier, transfusion activity, and month of year that influenced RBC outdate rates. They examined Ontario hospital transfusion services in terms of hospital characteristics, as well as distance from blood supplier and outdate rates, and ultimately compared hospitals of similar size (transfusions per month) and distance to the blood supplier using a bubble plot technique to visualise outdate rates and the aforementioned variables. The authors note that benchmarking is an important tool for continuous quality improvement, which is underutilized in the field of transfusion, and they note that this is mostly due to limited data or lack of comparison data [20]. A derivation of this tool is being used by CBS and by other provincial blood conservation programs to benchmark transfusion services (Blood Disposition Report, Canadian Blood Services, internal data).

There are other papers which examined inventory KPIs; however, they were highly specialised and non-generalizable, providing marginal additional insight into how direct inventory measurement techniques have been used. One small study examined the reasons for outdating group A units compared with group O units over a 6 month period and determined that neonatal transfusions and blood importing policies influenced outdating [158]. Another paper from Cheng *et al*. examined the red cell lifecycle and described the inventory through visualisation of day to day inventory, and highlighted how red cells transitioned through the blood system on their way to being discarded or transfused. This study was primarily descriptive in nature [37] but did acknowledge that there were sub-processes that needed to be further understood.

Overall, KPIs are critical for understanding and managing day-to-day transfusion inventories, as they are direct indicators of inventory performance and simple to understand. Although there is

an understanding of what the KPIs mean, there are still subtle differences in the literature as noted within the wastage category in Table 6, as different institutions calculate wastage differently. Although technically it is possible to benchmark institutions against each other, it is difficult to accurately benchmark processes behind those KPIs, as the comparisons are not equivalent measures. KPI measures also are limited in their ability to describe an inventory over time and space, as measures of time are not typically considered.

The second group of direct methodologies involves policy changes based on active interventions or technologies driven by KPIs with KPI-based outcomes. While policy changes are not analytical methods, they do occur in the context of a real inventory and are an important aspect of the transfusion inventory management literature. Policy changes are not used to understand inventory, but rather they are used to influence inventory practices and are based on themes such as information flow, information systems, tracking, storage, and other management-related practices designed to impact KPIs. Several papers relate to information flow and information generation policies which have directly impacted inventories. These papers highlight the use of laboratory information systems and related information exchange in hospital blood banks [76, 159, 160], such as computer assisted crossmatching, commonly used in modern hospitals [161], and radiofrequency identification to improve platelet management [162]. Another paper outlines an information exchange process using committee meetings surrounding the blood supply chain [63, 163, 164]. Another series of papers highlights blood recycling and rotation, either to or from small hospitals [31, 165]. Another broad category focussed on red cell preparation and storage media [166] , such as red cell freezing [55]. There is also broader discussion in these papers about donation and donor management [89, 167-169]. Poisson *et al.* describe the use of active management of RBC age through the use of a smaller inventory and the requesting of fresher blood from the supplier by the transfusion service [170]. Zoric *et al.* highlight a two period study over ten years whereby seven blood wastage mitigation strategies, including auditing, transportation, storage, and changes in transfusion practices, and found that there was a reduction in blood wastage in the second period from 4% to1% [46]. Collins *et al.* demonstrated a low-cost methodology of using multiple interventions on the wastage rates of platelets, red cell, and plasma units using a combination of educational initiatives, messaging (print and digital), component transportation improvements,

and component identification, with net savings of approximately $131,520 excluding intervention costs [21].

2.5.5.2 Indirect Methodologies Informing the Transfusion Inventory Management Paradigm

The research will define an *indirect methodology* as an approach that represents a real-life inventory in an idealized state, even though it may possibly use real data. Indirect methodologies are usually more mathematically or computationally complex thus rendering them 'opaque' for average human understanding and day-to-day use. There is a role for these methodologies during the inventory management process, as they typically can be used for initial planning purposes, and may be used to predict what would happen if certain policies were followed or not [25]. Methodologies in this category involve, for example, mathematical modeling and simulation. The scope of this section will not examine each methodology in detail, but to acknowledge its role in the context of transfusion inventory management.

Modeling involves constructing a working representation (i.e. a model) of a system in question. The model will always be simpler than the system it represents, and allows the operator to examine the effect of changes on the actual system. To be effective, the model must closely approximate the real system by including its most relevant features, but balance being overly complex which can render it impossible to understand, or to manipulate either experimentally or computationally. Whenever models are utilized, the question of validity must be answered – "how close is this model to approximating this real life system?" To validate such models, the real environment is simulated within the model by using known input variables while comparing the model output with known actual system outputs. In addition to validating a defined model, simulation also has another role. Simulation allows for the evaluation of system performance under different inputs, configurations, and allows this to happen over time. Simulation is a well-known technique for assessing the operational impact of changes to facilities, processes, and systems, and also to experiment with "what-if" questions prior to any live implementation. Simulation is known to be used in systems which have been altered or created to quantify

bottlenecks and can reduce the chance for failure, and prevent inappropriate resource utilization [25].

Modeling-based Inventory Management Techniques

There are different classes of models discussed in the transfusion inventory management literature. Although there is a technical distinction between deterministic models (where input and output variables have fixed values) and stochastic models (where at least one of the input or output variables is probabilistic), this distinction is irrelevant according to Belien [25]. There are also dynamic and static models, which may or may not take time into consideration respectively [171].

Some of the earliest papers on computer modeling and simulation were written in the 1960s. Elston [172] suggested that daily blood usage could be approximated by a lognormal or a Poisson random variable; however, this was later refuted by Brodheim and Prastacos using a larger dataset [173, 174]. Jennings [175] used a mathematical model that was opaque, and in the 21 day shelf-life era, which determined that the age of received units increased outdating, and that there was a tradeoff between shortage and outdating.

Data envelopment analysis is a non-parametric technique which analyses efficiency based on various economic inputs and outputs, allowing for efficiency determination and productivity benchmarking. It is an indirect measurement of institutional efficiency as efficiency scores are calculated and reflects how inputs (i.e. employees, volunteers, budget, population serviced), are converted to outputs (i.e. blood products - red blood cells, plasma, platelets) and these are based on the performance of other organisations [176, 177].

Another modeling category involves queuing models. Mathematically these involve Markov chain models [178], and have been used to assess issuing policies on average age of blood transfusion and inventory levels [179]. Other uses of Markovian models include modeling blood supplier collection plans to reduce imbalances and wastage resulting from rescheduling collections [180]. Broadheim and Prastacos [174]combine a Markovian model to schedule blood delivery with the statistical need of individual hospital blood bank, forecasting blood surpluses and shortages, allowing for blood distribution control to various regions. The authors estimated

that there was $500,000USD saved per year, and reduced delivery charges by $100,000 per year (1979 dollars; $1,734,209.74; $346,841.94 respectively; inflation adjusted with dollartimes.com; 3.52% annual inflation rate).

Stochastic dynamic programming has been used for blood product distribution [181-183] and due to computational complexity, has also been simplified in the platelet inventory space [1].

Kendall deployed a goal programming model to attain multiple goals based on constraints related to inventory level, fresh blood availability, outdating, age, and collection costs. They examined an unlimited rotation model, limited rotation model, and a non-rotational model for inventory and studied the interplay between shortage and outdate rate [184]. Kendall also studied the Long Island Model [174], suggesting that the specificities of the parameters of Long Island might not be consistent with their own and that their own model allowed for their own priorities and a regional administrator's wishes to be incorporated, implying that another model could not be translated to their institution or other institutions.

One of the oldest papers in statistical modeling involves Elston and Pickrel's work which determined that blood inventory follows a negative binomial distribution, whereby weekdays are different than weekends, and they used a UNIVAC 1105 to run the analysis [185]. Statistical modeling has also been used to relate shortage rates, inventory level, and mean blood demand at a hospital level [28] or at the blood collection stage [186]. In the platelet transfusion setting using time series modeling to determine the total number of platelets needed in a tertiary care hospital on a weekly basis [187]. Cohen *et al.* used multivariate regression analysis to describe regional blood bank system performance in light of other variables [188]. Logistic regression and survival regression was used to examine when blood donors would arrive [189] or return [169, 190]. Logistic regression was also used to predict factors for blood donation [191, 192].

Other complex mathematical models, including optimal myopic allocation models, have also attempted to describe issuing policies from a hospital or a central blood bank [27, 193, 194]. Inventory and allocation setting models have been used to minimize shortage rate based on CT ratio, daily demand and length of crossmatch period [174], and examine the impact of inventory on some of these parameters [28]. The complex interaction between optimal inventory level with daily demands, transfusion to crossmatch ratio, release period and age of arriving unit has

been mathematically described and used to predict shortage and outdate rates [195], allowing an operator to set their inventory optimally.

There is only a single paper using machine learning methodologies – k-means cluster analysis– which examined factors predicting dropout vs. commitment for first-time donors[196].

Simulation-based Inventory Management Techniques

One of the earliest simulation papers was by Jennings, where a complex computer model was used to calculate the desired beginning inventory level (all unassigned units in the morning). This model included many variables and it highlighted that increasing frequency of ordering, reducing the average inventory age of blood, and limiting 'non-useful time' whereby units would be spent in crossmatched status would be optimal [175]. Dumas modeled and simulated multiple crossmatch situations whereby units are available to multiple users (i.e. double crossmatching) and examined a policy which considers Rh negative blood approaching the end of transfusable life for Rh positive patients and a combination of the two policies. Dumas found that the most effective reduction in wastage of both Rh positive and negative blood is achieved by combining both double crossmatching and Rh negative to positive policies [74]. This was based on earlier studies simulating the impact of different policies on wastage and workload [197]. Cohen and Pierskalla used a regional simulation model to study issuing and crossmatch policy combinations and another model to examine policy inputs and performance indicator outputs, with the conclusion that units that are crossmatched should be made available as soon as they are not needed [30]. Cumming studied frozen blood during times of shortage (in the 21 day shelf-life era) using a simulation model for the Midwest blood region, but did not find any major benefits to freezing red cells [198]. Friedman examined computer simulation to determine the impact of inventory reductions on shortages and outdate rates, and the extent of using group O inventory as a buffer for shortages of other blood types. This study occurred at the beginning of the computerisation era and assessed the effect of inventory reductions on shortage and outdate rates using 21 day and 35 day shelf life RBCs. The study was a simulation examining the extent of group O inventory to buffer against periodic shortages of other blood types. The other modeling was done to examine the impact of shelf-life extension of red cells

from 21 to 35 days on the inventory. Friedman found that lower inventory levels and rapid return of crossmatched units were important variables affecting outdate rates [199].

Simulation methodologies have been used in platelet inventory management to examine shortage and outdate rates based only on mean daily demand [200]. Using the stochastic dynamic programming method outlined above, van Dijk *et al.* applied simulation and asked 'what if' questions based on shelf-life extensions to six or seven days, changes in issuance policy (LIFO and FIFO), and what would lead to decreases in shortages [201-203]. Katsaliaki examined the effect of various policy changes on the percentage change in various performance measurements for approximately 26 weeks and for 630 orders of RBC units per month [36, 204]. These authors also published using a game-based approach, based on their model, to teach students and healthcare service professionals about how to steward the blood supply chain [42]. At the Finnish Red Cross, modeling and simulation studies were carried out to examine the impact of eleven operational policies and their effects on cost and service level metrics such as outdating and backorder-related costs and percentages across all blood products including rare blood types. The authors discovered that better inventory control results from using the patient's blood type, that fresher blood from the blood distribution/supplier beneficially impacts KPIs in addition to having a smaller platelet inventory [56].

Katz *et al.* [205] used Monte Carlo simulation methods and reconstructed probability distributions based on generation of random numbers, with application to two years of platelet orders to maximise availability while minimizing waste. Custer *et al.* used a cohort simulation model (@Risk 4.5 Software) using data from year 2000 of allogeneic whole blood donation to evaluate new donor and donation deferral requirements on the supply and cost of blood, and used it to evaluate threats, safety, and adequacy of the blood supply [167]. Kamp *et al.* simulated the impact of epidemic and pandemic situations using a computer model which allowed for the analysis of deficits in the blood supply. The authors indicate that during pandemic scenarios, a triaging scheme for the release and use of blood products should be done before any critical events occur, and demonstrated that improving blood supply occurs through relaxing donor deferral time [206]. Pereira employed a stochastic model that simulated routine hospital blood bank inventory operations and determined that the coefficient of variation of daily transfusion is the major parameter in inventory performance, in that smaller variations

perform better with older stocks, and that frequency of shipment, as well as the remaining shelf-life of units was important [29]. Kopach acknowledged that inventory modeling is a complex task, and used queueing simulation to obviate the need to track the individual age of items in the inventory. Their model considered a fixed-life perishable commodity problem with lost sales, and allowed for modeling trade-offs between multiple demand levels (i.e. in emergency and discretionary transfusions), and service costs/levels and attempted to match supply and demand [207]. Asllani used a simulation-based decision support system to reduce waste while meeting demand for apheresis platelets, but it was a black-box model [208].

Process simulation has also been used by Baesler *et al.* to address twelve different scenarios through the supply chain, and determined it was possible to reduce wastage and unsatisfied demand by 2.5-3% [209]. Vaquero *et al.* also described a mathematical simulation model in Spain to optimise platelet concentrate stocks [45].

In summary, although there are many simulation and modeling methods which use transfusion inventory data, Jagannathan summarise the limitations of simulation and modeling studies by indicating that there is limited generalisability, inability to integrate crossmatching, or the difficulty in arriving at simple decision rules [210].

2.5.6 Results and Gaps of Current and Available Methods for Informing Inventory Management Practices

The available indirect and direct methodologies to date have resulted in reductions in wastage, improvement in blood product age, reduction of shortages and improvement of cost variables according to the literature. Factors such as inventory size, age, crossmatching rates, times and the availability of a recycling program are important inputs that affect inventory performance. In addition, improved visibility of the transfusion inventory through benchmarking or newer electronic methods, communication, and education have also been factors that have contributed to the understanding of inventories in the context of inventory management practice.

Despite the foundational aspects of these methods to managing the transfusion service, there exist significant gaps and shortfalls to these techniques. The first shortfall is that these techniques are fundamentally antiquated. Most of these techniques were introduced prior to the 1980s, with most of the techniques being introduced in the mid-1970s, approximately 40 years ago [25]. Many of these papers originated in the pre-computerised era, involving blood transfusion products that were not modern in their constitution (i.e. twenty-one day shelf life, different preservatives and anticoagulants, prior to leukoreduction), and existed in the environment where there was no electronic crossmatching, significant electronic auditing capabilities, and unit documentation within the LIS. Anecdotally, the focus of the recent literature is on adaptation of previously established techniques with no new paradigm-altering tools. A second gap that exists is functional in nature. Much of the techniques available for modeling or simulating the inventory are 'black boxes', and although they demonstrate an output that is understandable in terms of KPIs, the inner workings of the algorithms contains complex mathematics and can only be understood by SMEs. These modeling and simulation techniques reflect idealised inventories rather than real inventories, with the pitfall being that these techniques tend not to be generalizable to other transfusion systems.

Although the methodologies are seemingly complete and mature, there are some other blindspots that appear when looking at the literature body from a conceptual perspective, suggesting that there is still more insight that can be gleaned from the transfusion inventory.

One area where there are potential deficiencies is in the area of benchmarking and best practices. Typically, the goal of benchmarking is to determine which hospitals or suppliers have the best KPIs, find out what they do to achieve such good performance indicators and attempt to emulate them. Another goal of benchmarking is to identify those hospitals or suppliers which perform poorly, attempt to diagnose problems in their policies or procedures, and improve them in order to improve their KPIs. The transfusion benchmarking literature only documents external and environmental variables associated with hospital KPIs, and there are very few internal inventory variables of any significance. For instance, current inventory benchmarking practices in Canada are based on a study from Heddle *et al.* which found that distance from blood supplier, mean monthly transfusions, and month of the year were variables affecting RBC outdates [20]. This study found significantly discrepant results between some hospitals that are

of similar transfusion activity and distance from blood supplier, but there was no explanation as to why inventory performance was so disparate between these hospitals. This raises important philosophical questions: Can one truly benchmark different transfusion services by only looking at external and environmental parameters? Are these parameters the correct determinants to examine inventories? The follow-up question would involve the portability or generalisability of techniques used in one setting for another setting. The area of expressing inventory performance using numerical KPIs is not perfect since there are many ways to express similar concepts. As noted previously (Table 6), there are many ways to express the concept of wastage, and different formulae can be used to derive various percentages. Due to this variability, there is always a question of how various KPIs are calculated. When comparing more than one institution and depending on the formulas that are used, it may not be possible to equivalently compare KPIs without querying extra input variables or using conversion factors. Are there other KPIs that exist or that could be created, which are truly equivalent between institutions, and which can be easily understood, interpreted, and generalised?

Another issue and gap with the available methods to understand inventory is that all of them scratch the surface of the workings of the inventory and currently the inner workings of the real transfusion inventory are largely undescribed and remain opaque. All of the direct and indirect methodologies in the literature have stopped short of truly describing what is inside the opaquw 'black box'. Could this be contributing to the lack of understanding regarding avoidable wastage? How do units interact with surgical services or the emergency department? If there is recycling, how does this happen and does it perform well? How long do units persist in certain states or are delayed for various reasons? How does an inventory function or look like if it is distributed over multiple sites? None of these questions are able to be answered with conventional techniques in the real inventory realm.

Another gap or constraint with the available techniques is the resource intensiveness of the available methods. For instance, it is also not economically feasible to continue hiring consultancy services for one time assessments or snapshots of how the inventory is performing as suggested by some of the literature. Likewise, it is not feasible to have mathematical or

engineering support to help create or interpret mathematical models from a continuous quality improvement standpoint. The issue of resources is very important in terms of sustainability, as with any continuous process improvement, such as inventory optimisation, the tools must be accessible and available on a continuing basis so that changes can be evaluated and assessed for benefit. The resource constraints on the transfusion service thus impede it from continuous quality improvement.

One significant issue that has been stated in the literature is the criticism of the data itself. There is a stated goal in the literature to understand the transfusion supply chain from end-to-end, particular to understand reasons for avoidable wastage. Devine *et al.* [2] state this plainly as follows: 'a *strong recurring theme in these responses is the presence of gaps in the information required to really have an optimal inventory management system. Many providers of blood products have only one end of the rope: that which is attached to the inventory around the production and distribution processes. Very few respondents had information back from the hospital customers as to the disposition of the products provided in terms of when or even if they were transfused. This lack of information hampers the ability to drive down the rate of avoidable discards or to know if the manner in which the blood provider manages inventory is optimal for patient safety or ready access to transfusion products* [2].' No matter what techniques are used, a full understanding of the entire transfusion process cannot happen unless the data is either linked or that it originates from a source that captures the entire transfusion supply chain. This is one view of the problem. Another view is that the lack of connection is not the problem, but that the individual rope ends are not fully understood to begin with. This relates to the inability of the currently available data analytical techniques to mine the full depth of the available data. Despite the computerisation of transfusion inventories and the advancement of information systems that house transfusion inventory data, the techniques that are being used merely scratch the surface and do not capitalise on the potential capability of the data. Why are other data mining, artificial intelligence, or machine learning techniques in other industries not being used to infer associations between events in transfusion inventories? Can the use of big data analytics, social network mining analysis, and other newer tools be used? Why do we only still aggregate variables and calculate them as simple percentages or rates?

Overall, whatever the gaps and deficiencies in the methodologies or data are, they have prevented blood transfusion services globally from fully understanding their inventories and the causes of avoidable blood product wastage.

2.5.7 Process Mining and Its Applications

Process mining is a very young field, and the literature regarding its application to real-life industries is very scarce. A very recent article by De Weerdt *et al.* illustrated eleven process mining articles from 2007-2011 inclusive [48] and a literature search revealed a few more not listed in that review. Process mining has been applied to telecommunications industry, insurance sector, healthcare, silicon wafer manufacturing, and is evolving the way financial auditing is performed.

Goedertier [211] described an application within the telecommunications industry, highlighting three possible algorithms –genetic miner [140], AGNEs [145], and the Heuristics miner [136]. The paper examined the event logs from a European telecom provider, how cases were being handled by different operator seniority levels, and the validity of these generated process maps was confirmed by a Lean Sigma black belt expert [211]. The paper itself was a validation of the methodology rather than a detailed study on how it impacted the telecom provider. Mahendrawathi [212] also examined an unstructured customer fulfillment process within a telecommunications company. Process map generation was using Disco and ProM tools. The authors discovered that there was a low completion rate, and this may have been secondary to manual recording or recording on different systems, and stresses the importance of data integration.

De Weerdt [48] described the use of process mining to highlight operations using process data from a Document Management System of a Belgian insurance company. They were able to visualise how teams functioned, and examine confusion in document forwarding and returned documents being sent to the wrong team. They used this to create a benchmark of how long a document resided in the system prior to termination and the throughput of documents in working days. The case study demonstrated that process mining can help to identify and explore areas of inefficiency, and provide business insight.

Within the health care setting, Mans *et al.* [213] applied process mining to gynecological oncology patient flows at a Dutch hospital, and examined them from control flow, organizational, and performance facets. Relevant event logs were analyzed on the ProM framework after extraction from the hospital information system. In another health care environment, Rebuge *et al.* [52] described process mining within a healthcare environment using a hospital emergency service as a case study in Portugal. The authors [52] used different process mining tools to examine processes such as emergency radiology workflow and handover using the ProM platform with hospital electronic patient record system data. They indicated that they can use the knowledge gained by examining the radiology workflow to assess for adherence to guidelines and performance [52]. In another health care environment in Taiwan, Yang and Hwang [214] used a process mining approach to examine health care fraud in a real-life data set from the National Health Insurance program relating to a gynecology department of a hospital between July 2001 and June 2002. They noted that there were some fraudulent cases that are detected by the detection model and missed by the manually constructed detection model.

Process mining has also been applied to audit for fraud in other settings. The earliest conceptual application of process mining was to complement the work of an auditor, termed *Auditing 2.0*, and was highlighted by van der Aalst [50]. The authors describe the natural evolution of auditing in the context of more strict legislation, such as the Sarbanes-Oxley Act (2002) and Basel II Accord (2004), drafted in response to various accounting scandals (i.e. Enron, Tyco). Jans *et al.* [215] examined internal transaction fraud, a subset of corporate fraud, specifically in the area of procurement and creation of purchase orders. They applied process mining to history logs at a case company that was in the top 20 European financial institutions. They compared the use of descriptive data mining techniques and suggested extending transaction fraud detection to process mining. The methodology Jans *et al.* [215] used consists of five steps: log preparation, inspection, control flow analysis, performance analysis and role analysis, and used the final state machine (FSM) miner algorithm on the Petrify tool, fuzzy miner algorithm, and social network analysis. They concluded that a lot of the tools that are available are still premature, but that this could provide another view into transaction fraud. In another paper, Jans *et al.* [216] used the analogy of transactional data within an Enterprise Resource Planning System, such as SAP™, being similar to video documentation generated from surveillance cameras. Jans indicated that by knowing that these event logs are being mined, it

may deter individuals from deviating from policy. Also, the article highlighted that if there is no electronic trace of a process, fraud can still occur [216]. In a subsequent article, Jans *et al.* [49] examined the approach of using process mining in the financial industry, and highlighted that the approach did not focus on the value of the transaction, typically examined by conventional auditing techniques, but the actual transactional processes. This paper discussed the value-added benefit for accounting, allowing for the reduction of process logs to core processes and the ability to see less frequent flows using Disco. They also highlight the potential use of rule analysis and social network analysis, and highlight cases where individuals violated segregation of duty controls and signature control violations with regards to Section 404 in the Sarbanes-Oxley Act. Jans *et al.* highlighted that process mining helps to uncover areas where auditing can occur and that are not detected by conventional audit procedures [49].

Process mining can also be applied to public utilities. Van der Aalst [154] described process mining an industrial event log from an operational workflow management system used for invoice handling at a National Public Works Department (the *Rijkswaterstaat*). One lesson learned was that for the information to be relevant, the people involved with the organisation needed to be involved with the analysis. The author concluded the analysis by indicating that process mining could be applicable for healthcare, Web services, and case handling.

Within the semiconductor industry and manufacturing, Rozinat *et al.* [217] described a process mining application in the area of wafer scanners, instruments that print microchips. The wafer scanners are complicated instruments, composed of 1000 building blocks, and are required to have performance and reliability checks. The authors were able to examine logs of various tests that were executed, such as instrument stabilisation (PYWZ), reliability of a wafer scanner subsystem (DSNA), and another unknown test (IWOW) and were able to characterise their execution patterns. Paszkiewicz described process mining a warehouse management system (WMS) and examined conformance of inventory management processes using the ProM and Disco tools in the context of a Polish mattress company employing approximately 500 people (Dendro Poland Ltd), which is an exclusive supplier for IKEA™ in Western Europe. This paper examined how process mining can be used to answer real-life organizational process questions

from management in the context of conformance checking against a pallet life-cycle model *de jure* (defined as a process model that is normative, which indicates how a process should be done). They examined processes surrounding lack of delivery of pallets to clients, total process from production to shipping of pallets, and pallets that were damaged and the particular employees that were involved. They also examined the unequal distribution of work among various employees. Their analysis led to scheduling changes and more precise information on the FIFO policy for various product families. The authors' main conclusion was that to be an effective conformance testing process, examination of the process and the social relationship structures is required [218].

As the process mining field is in its infancy, there is an annual process mining workshop held for the global community of process mining experts (Process Mining Camp, Eindhoven, Netherlands). The 2016 process mining camp featured companies such as Telefonica (Spain), DHL Group (Germany), ALFAM Consumer Credit (Netherlands), and Sparq (Australia). There were talks on process mining and dispatching, internal audit, contract and quotation, and housing allocation. Most of these process mining activities have not been documented in the literature but yet are still being performed.

2.5.8 Conclusion for Background

Transfusion is an important part of modern-day medicine and managing transfusion product inventories is a crucial aspect to delivering quality blood products in a timely manner. Due to the stochastic nature of supply and demand, it is necessary for blood suppliers and hospitals to buffer that uncertainty by maintaining inventories. Having an inventory carries risk, as the inventory represents significant capital and opportunity costs, and must be managed adequately to maximise availability but minimise wastage of blood transfusion resources.

The practices surrounding the management of blood transfusion inventories have been rooted in various approaches such as mathematical modeling and simulation studies, qualitative and

descriptive studies on KPIs, inventory transparency, and studies outlining various policy changes. This literature has yielded several major paradigms in transfusion inventory management practice:

- FIFO - using the oldest blood first
- Inventory size should be kept to the minimally appropriate size – 'rightsizing' inventory
- Reduction in crossmatching rates should occur.
- Reduction in crossmatching times.
- Recycling of inventory to high-turnover sites
- Improvement of inventory supply chain transparency through information

The daily management of the transfusion service must also occur in the context of being able to describe the inventory and that is typically done through KPIs, such as wastage rates and inventory size. Using these indicators the performance of the inventory can be documented, communicated, and assessed on a continuous basis.

There are significant issues that are facing modern transfusion medicine practice. These include increasing costs and efficiency pressures, increasing product diversity and indications for use, increasing centralization and complexity of transfusion services related to economies of scale, and increasing supply and demand pressures as a result of the aging population (less donations and increased usage). Inventories on both the supplier and hospital side have had to work in this context, and those issues and pressures do not appear to be reducing.

Despite having methods and knowledge available for the last 50+ years to manage inventories, the current literature highlights a paradox, and this is best summarised by Devine *et al.*:

> *'while there is clearly no singular right way to handle inventory management and the choices involved therein, there is a need for constructive dialogue amongst blood suppliers and users of blood products to identify a set of best/better practices that can be implemented without prohibitory levels of cost or complexity to increase the availability, safety and quality of blood products for the patients who need them while at*

the same time increasing the efficiency of this important aspect of blood transfusion management. [2]'.

A critique of the current literature on transfusion methodologies highlights several gaps that are underappreciated by the transfusion community. Most tools which are being used for inventories either deal in an idealised space (i.e. modeling and simulation), or are very crude and of low resolution (i.e. KPI-based methods, benchmarking); policy interventions are usually limited in their applicability to other institutions and are resource intensive. There also exists an informational discontinuity between supplier and user, highlighted by Devine *et al.*[2], whereby data for the entire process of blood collection at the supplier through to final disposition at the hospital is not readily available, as the data typically rests in different silos. Moreover, there is a significant gap between the information captured by modern information systems and the methods that the transfusion community uses for analysis. Currently, day-to-day transfusion inventory management relies solely on the use of KPIs, which are rolled up and aggregated numerical descriptions of how the inventory is performing. This is the only method by which a real inventory can be described, as other techniques are abstract and deal in an idealised environment. This low-resolution description of inventory performance is happening, despite high resolution data being captured daily in modern transfusion medicine information systems, and the transfusion literature has not addressed this issue.

The rationale for using process mining is that this technique has not yet been applied to transfusion inventory data, despite the inventory being alluded to as a process in the recent literature, and despite the financial auditing industry leading this evolution already. Another rationale for using process mining is that the transfusion inventory is inherently a process whereby blood products are collected in the donation process at the supplier, move through processing and testing, and finally are released into the supplier's inventory. These units will then be shipped to the hospital, transacted internally, and then will have a final disposition. All of these activities (or actions) have interconnected dependent pathways, and thus lend the transfusion inventory to be visualised as a process. Until now, this complexity has only been summarised using a value-based KPI-centric view, rather than using a process-base view using process-centric tools. In addition to the technique itself, the culture of auditing in transfusion

medicine is analogous to auditing done in financial services, which has already been nicely complemented by process mining. Furthermore, in modern transfusion information systems, all of the activities related to blood products are tracked, time-stamped, and most blood products are already uniquely coded to ensure traceability from donor to recipient. Although these steps are not explicitly termed a process log, the data elements are present. The final rationale involves the need to find something that can be practically implemented and generalizable to transfusion services globally. Process mining is a low cost and low overhead technique that can be used to visualize already-available data within transfusion information systems, and can potentially be universally applied to any modern transfusion inventory.

Chapter 3 Research Methodology and Methods

This chapter will outline a generic approach to identifying if the study subject is a process and examine the environmental factors defining the process. The chapter will also discuss the specific development of the Transfusion Inventory Process Mining Methodology, and examine how it is applied to a red cell product inventory. This chapter will also discuss the evaluation framework and knowledge translation aspects associated with the research.

3.1 Problem Identification

The first step of this research is to examine if a generic problem or entity can be understood as a process. The term '*process'* in the Merriam-Webster Online Dictionary is "*a series of actions or operations conducing to an end; especially a continuous operation or treatment especially in manufacture*" [219]. Dictionary.com similarly defines a process as a "*systematic series of actions directed to some end*" or "*a continuous action, operation, or series of changes taking place in a definite manner*" [220]. Therefore, if a generic situation or circumstance was to be considered as a process it must satisfy all of the following three criteria:

1. a series of actions or operations
2. continuity between these actions or operations
3. an endpoint or purpose

In light of this, does the blood transfusion inventory and its associated business practices meet these criteria? This discussion is represented by the following evaluation matrix in Table 11.

Table 11: Evaluation Matrix to Determine If a Blood Transfusion Inventory Is a Process

Criterion	Blood transfusion inventory business practice	Criterion met (yes/no)
a series of actions or operations	Blood products are handled in discrete steps [36]	Yes
continuity between these actions or operations	Each discrete step is connected to another step [36]	Yes
an endpoint or purpose	To deliver blood product from a donor to a recipient [2, 36].	Yes

As the blood product inventory meets the basic evaluation criteria as a process, it can therefore be analysed as one. The next section will deal with how the problem can be evaluated and further specified.

3.2 Problem Specification

If the problem or entity in question fulfills the basic definition as a process, then it is important to determine if the local environment or business practice contains parameters to support and enable a process. This is summarized in a quote from Michael Hammer, in that the fundamental aspect of process [design] is "*the specification of what tasks are to be performed, by whom, when, in what locations, under what circumstances, to what degree of precision, with what information…* " [221].

The examination of the local environment for these process specifying parameters should be broad and include examining available documentation located in standard operating procedures, performing observations of staff and audits, and conducting interviews with SMEs and employees, among other methods. This documentation ideally would occur in the context of an electronic environment so that the documentation is robust and reproducible. Examples of process specifying parameters could include examples from Table 12.

Table 12: Process Specifying Parameters Adapted from Hammer [221].

Process Specifying Parameter	Examples
Tasks	Do this action; run this task; achieve this state
By whom	Human resource; machine; kiosk; self
When	Timestamp; at night
Process conditions	Location, circumstances (i.e. unsupervised); with what information (i.e. input variables); degree of precision (i.e. cursory initial scan, in-depth analysis)

The ultimate goal of the parameters of process specification is to obtain the breadth of descriptors associated with a business process to enable robust visualisation of it as a process map, as this is a methodology which is commonly used to execute, examine and improve business operations [221]. This process of constructing a model to the right resolution and detail is termed 'process model abstraction' [221]. Proper process model abstraction will allow for essential details of a business process to be visualised and minimize irrelevant details. Abstraction in this research will be automatically generated using the process mining software.

3.2.1 Research Environment

The overall environment in which this research is occurring can be summarized by the global survey on blood transfusion inventory management published by Devine *et al.* [2], indicating the need to increase efficiency from a cost perspective in the context of a growing blood product menu. The study was published in 2010, but those trends continue to persist at the time of the writing of this manuscript, as CBS continues to introduce initiatives to improve inventory transparency, benchmark product utilization, and hopefully decrease the gap in the 'rope' that is referred to by Devine *et al.* (CBS e-mail communications, Summer 2016). These initiatives are also in the context of ballooning budgets, as CBS is facing increased costs due to the weakness in the Canadian currency versus the US dollar and the euro in the past 3-4 years (National Liaison Committee documents, 2016). Therefore, this environment persists and continues to be an important issue in Canadian blood provision.

The research occurs within the inventory at Blood Transfusion Services (BTS) of the Central Zone at the Nova Scotia Health Authority. The transfusion service has always operated in a publicly-funded healthcare environment with limited human and material resources. The Central Zone transfusion service accounts for just over 50% of blood product utilization in the province of Nova Scotia, and like other hospital zones in the province, has always been monitored and kept accountable by the Nova Scotia Provincial Blood Coordinating Program, a practice and utilization standardization body in Nova Scotia. The NSHA is ultimately accountable to the Department of Health and Wellness (Provincial Government) and the taxpaying public for the utilization of this resource even though the cost of blood is not borne by the health region but is paid for by the Province, thus ultimately from the taxpayers. The Central Zone region is one of four zones within the Nova Scotia Health Authority. The Central Zone contains several hospital sites which transfuse blood: Halifax Infirmary (HI), Victoria General (VGH), Dartmouth General (DGH), and Hants Community (HC), geographically spread out over an area that is within a 90km radius from the hub of the transfusion service at HI site. There is a mixture of tertiary (HI/VGH), medium-sized (DGH), and community (HC) hospitals with services including solid organ and stem cell transplantation, a trauma service, and other cardiovascular and vascular services. CBS is the main blood supplier and is located in Dartmouth, Nova Scotia. The CZBTS also has close collaborations with the Pathology Informatics Group (PInG), which is responsible for the oversight and day-to-day maintenance of the LIS, Cerner Pathnet Millennium™ (www.cerner.com). Within PInG, there are data business analysts who are able to query the information system, write reports, and develop code to maximally utilise the information contained within the LIS. In addition, there are database analysts whose role is to ensure data quality and operational relevance.

The research follows a theme of inventory-related research and operational efficiency optimization dating back to 2006 [222]. This included the creation of the database-derived maximum surgical blood ordering schedule, resizing of the inventory, and the auditing of red blood cell age from CBS. These initiatives have resulted in a net savings of approximately $1,000,000 CDN dollars per year and reduced red cell wastage secondary to expiry from 20% to approximately 2% and below (internal data). Further to those initiatives, there has been creation of product recycling dashboards, a thrombocytopenia dashboard tool [32, 35], and most recently, a red cell inventory ordering algorithm. The paper which began the thinking of the inventory as a ‘lifecycle’ at our institution was published in 2010, prior to the authors’

knowledge of process-related terminology/software, and was done using data from Cerner Pathnet Classic, the previous LIS [37]. This research was also done prior to the authors' awareness of process mining tools, and was done with the notion that if there is cycling of RBC units inside the inventory, a reduction in unwanted RBC unit latency can improve inventory performance by making more of the inventory available, and therefore reduce wastage. All of the above initiatives were done in collaboration with the PInG and have been endorsed by the Division of Hematopathology and the Director of Blood Transfusion Services at the CZBTS. Various members of the PInG will also be involved in data extraction and initial validation of the query related to this research.

The impetus for this research was that the CZBTS had purchased an automated blood fridge in the 2012-2013 time frame, and was realizing that it was being significantly underutilized. It did not have the personnel to audit the transfusions coming from the fridge. In addition, the BTS was faced with increased outdate % rates, increased transfusion rates, and increased frustration with being unable to find the cause of the undesirable KPIs that were being used (outdate % and days of inventory on hand). Although there was a request to have a CBS inventory expert assist in this process, it was noted that recommendations from previous CBS inventory consultants to help model and simulate red cell and platelet inventory stocks (internal data, not published) have not been applied, as it was felt that they were not applicable for day-to-day use as they have been too generic or too complicated to use without specialized training (internal communication, NSHA). Therefore, the management welcomed any internal attempt to further understand the inventory. The research was therefore accelerated, with management and medical leadership support, as they needed a technique that was not too complicated to understand and that would assist their day-to-day understanding of the inventory.

Although different blood products within a transfusion inventory can be studied, such as platelets, the red blood cell component has many characteristics that make it favorable for study. The first characteristic of the RBC unit is that it has an intermediate maximum shelf life of cellular blood products: 5 days (platelets), 42 days for RBC units, and up to 1 year for frozen products, such as plasma. This intermediate shelf life allows one to monitor the impact of any

blood bank inventory policy on inventory metrics immediately, without having to wait for a significant period of time for the inventory to be affected. Studying the platelet inventory would be possible, but platelets are subjected to just-in-time inventory dynamics and careful management of platelets still does not result in reducing platelet outdating (typically around 20-30%) [2, 200]. The red cell unit can also be routinely manipulated without significant loss of function in various ways including irradiation, and washing or saline replacement (i.e. breaking of the sterile seal), with a respective reduction of maximum expiry to 29 days and 24 hours respectively [6] and rarely frozen at less than -65 degrees Celsius, which can prolong shelf-life to 10 years from collection. RBCs also account for the bulk of most cellular products within any transfusion inventory and are issued at a rate of 32.6/1000 population, 5-fold more than frozen plasma/fresh frozen plasma (FP/FFP) at 5.6/1000, and platelets (pooled, apheresis, buffy coat) at 4.2/1000 [2]. They also spend a significant portion of their life cycle within the inventory of a hospital transfusion service. Practically, RBC units are the only types of blood products that are crossmatched and stored at different inventory locations within the NSHA CZBTS and are the only blood products stored inside the Hemosafe^{TM} automated blood fridge. One important consideration for studying the RBC inventory is that the mechanism for reimbursement to Canadian Blood Services of whole blood derived products (i.e. plasma, platelets, red cells), is based solely on an equation reflecting the proportion of red blood cells used by a province divided by the total number of blood cells used in the country, regardless of how the RBC components are used (transfused, expired, discarded, etc.), multiplied by a total cost. Thus, due to the diversity of transactional states that can occur with RBCs, their usage rates and diverse interactions with the health care system, this was an ideal product to study.

3.2.2 Equipment and Software

As indicated before, this research was not going to individually address the different process mining algorithms, but the objective was to choose a flexible algorithm embedded within a user-friendly and easily available software package allowing for real-time manipulation and filtering of process data. Disco (http://www.fluxicon.com), available through educational licensure, was chosen as it had the capabilities for live process map generation in front of an audience at

different abstraction levels, while allowing for exploration of common and uncommon process events.

Process map generation was done using Disco 1.9.5. The PC was an Intel i5 with quad core processor with Windows 7 Home (64 bit). Disco's system control center indicated the software was running on Windows 7 6.1 x86, Java HotSpot Client VM 1.8.0_71 (32 bit) and processing power was 4 cores running at a clock rate of 2.6 GHz (single core: 916 MFLOPS; All 4 cores: 2108 MFLOPS) with 15.89 GB usable RAM and a disk read speed of 19.6MB/s and write speed of 342.5MB/s.

The Disco software was divided into numerous workflows: importation and data manipulation, process mapping, simple statistical analysis, and examination of variants and cases. Most of the functionality was accessed using the 'map', 'statistics', and 'cases' tabs at the top of the software interface.

Using the 'maps' tab of the interface, maps were automatically generated and started at minimum detail with minimum activities and paths on the detail slider at 0% (detail slider on left of Figure 6). This was done to minimise computational time and provide the starting and most basic process map of the data. The Frequency and Performance visualisations, as an example, showed Absolute Frequency for each of the activities and pathways as a primary variable, and mean duration as a secondary variable. Absolute Frequency was shown as bolded numerical data on the pathways or within activities (red, top arrow, Figure 7), and mean duration is shown in grey font underneath the Absolute Frequency and denoted with time units (green, bottom arrow, Figure 7). As the activities and paths sliders were adjusted, the process map would change in resolution (abstraction or zoom), similar to how a geographical map zooms in and out on websites such as Google Maps. Performance and Frequency data were modified to properly annotate the process maps as needed.

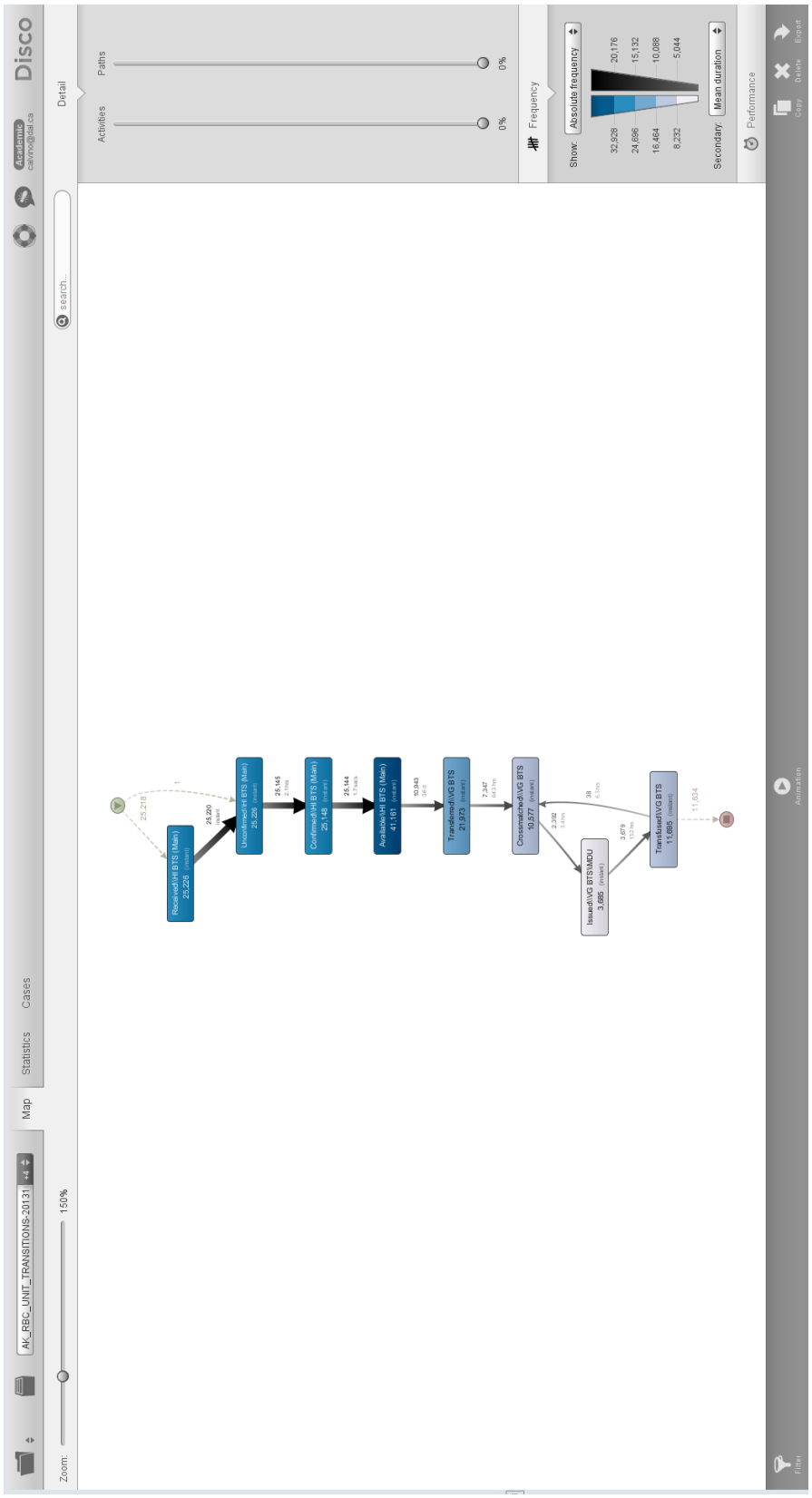


Figure 6: Interface window of Disco 1.9.5 process mining software('map' tab)

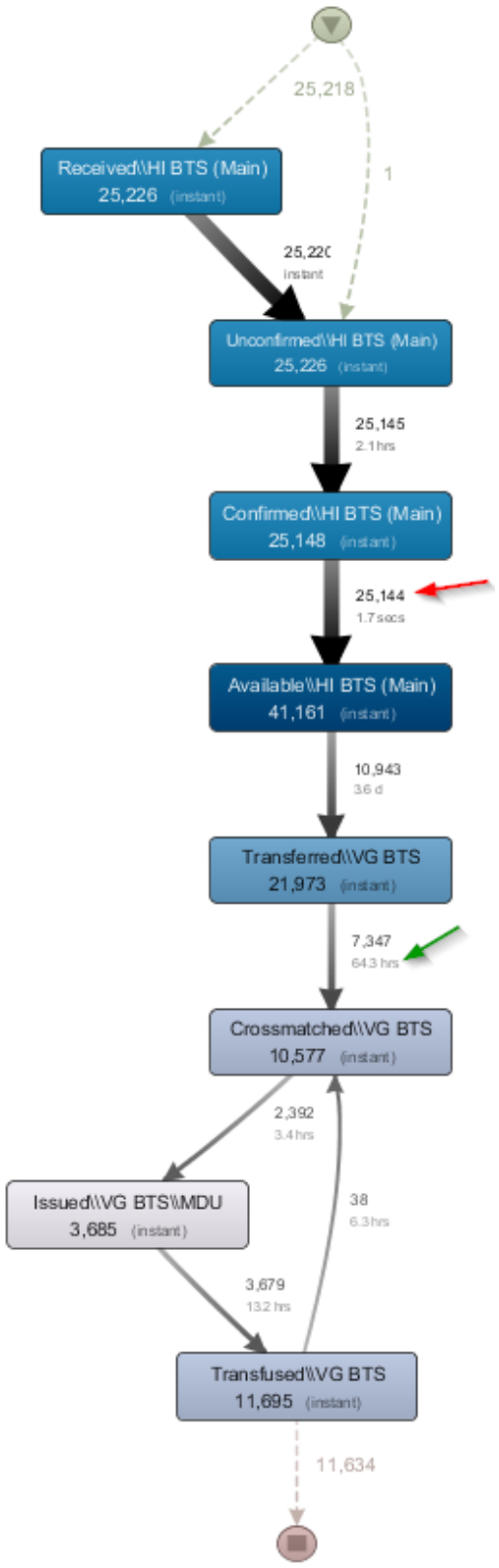


Figure 7: Process Map of AK_RBC_UNIT_TRANSITIONS-2013-201509 at 0% activity and 0% paths zoom resolution (red arrow denotes case frequency; green arrow denotes mean time between activities)

3.2.3 Hypothesis

In the past decade, the terminology used to refer to the red cell product inventory has become increasingly process-oriented, diagrammatically outlined in Katsaliaki *et al*. [36] as a process map, and described by Cheng *et al.*[37] as a 'lifecycle'. Despite this subtle shift in terminology, the red cell unit inventory (red box, Figure 8) is still being understood using traditional value-based descriptors, such as wastage and inventory size KPIs. Other industries, such as financial auditing [49, 50], have already transitioned from value-based towards process-based thinking and analytical techniques, by integrating process mining into financial process auditing.

The research hypothesizes that a transfusion service's red cell unit inventory (red box, Figure 8) can be visualised through process mining techniques using LIS data. The hypothesis also is that by analysing the process maps, process improvements can be made in day-to-day inventory management practice. This hypothesis will be tested using data already resident in the LIS and will apply the TRUISM Framework in Figure 15. Two case studies, using SMEs and a focus group methodology, will be performed and will assess if further process-related understanding of the red cell unit inventory can occur, and if guidelines for process improvement can be generated. The first case study will examine if red cell inventory wastage root causes can be understood. The second case study will characterise how effective the Hemosafe™ automated blood dispensing fridge is being used in the context of the transfusion service.

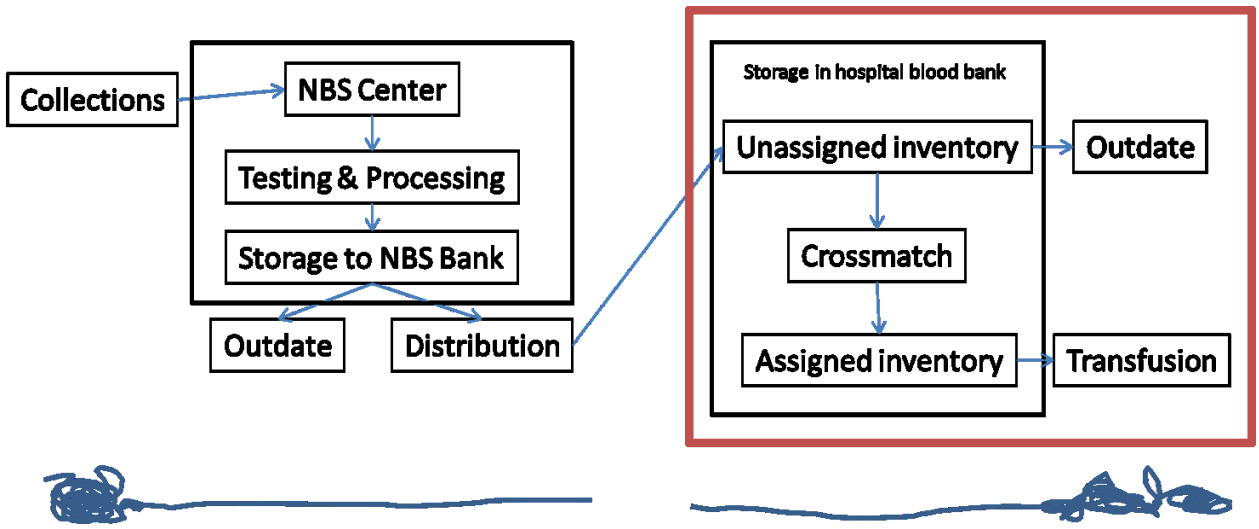


Figure 8: Process Flow Diagram of Blood Units from Supplier to Hospital, with the 'Storage in Hospital Blood Bank' Inventory Sub-processes Highlighted within the Red Box (Right); Adapted from Katsaliaki *et al*. [36],

3.3 Process Mining Methodology

As the process map will be generated using a process mining methodology, it is important to understand the anatomy of a generic process log, the input for process mining [47]. The hypothetical business process of handling a loan application noted in Figure 2 can be represented and decomposed into individual *Case IDs* in a compact form as indicated in Table 13. Each *Case ID* is unique and associated with a *Trace* (or variant), which consists of an ordered series of connected activities *(a, b, c, d, e)* pertaining to the *Case ID* where: *a = register client, b= check ID, c=initiate transfer, d=reject client, e=close file*, over the time dimension, *T*.

Table 13: Generic Log File with Data Elements of Case ID, Activity ID

Case ID	Trace
A	*{a, b, d, e}*
B	*{a,c}*
C	*{a,b,c}*
D	*{a,b,d,e}*
E	*{a, c}*
F	*{a, b, d, e}*
…	…

The same generic log file can be expressed as a tuple, where *process log (L)* contains multiple *traces {a, b,…}* over *time (T)*:

$$L = [\{a, b, d, e\}, \{a, c\}, \{a, b, c\}, \{a, b, d, e\}, \{a, c\}, \{a, b, d, e\} \dots]$$

The log file element of activity *(a, b, c, d, e)* can be further decomposed into verb-objects which allow for clarity of the activities and are useful for data interpretation [48]. If an activity is only expressed as a verb, there is less clarity than if it is expressed as a verb-object (Table 14). In this hypothetical example (Figure 2), the *verb* is an action and the *object* is that to which the action is done.

Table 14: Activity with Verb Object Decomposition Using Hypothetical Loan Application Business Process in Figure 2

Activity	Verb-Object	Verb	Object
a	Register client	Register	Client
b	Check ID	Check	ID
c	Reject client	Reject	Client
d	Initiate transfer	Initiate	Transfer
e	Close file	close	file

In order to use tools that will visualize the process log through process mining, the minimum elements of *Case ID, activities (a, b, c, d, e,…), and Time,* must be represented.

Generically, the process mining methodology output is a process map. This map will contain activities that are linked through pathways, and which visually represent the business process. The process map that is produced can be abstracted at different levels of detail to maximise the information gained and minimize extraneous details. From this point forward in the research, the rest of the steps associated with the generation and evaluation of the process map will be discussed in the context of the blood product inventory.

3.3.1 Development of the Transfusion Inventory Process Mining Framework (TRUISM)

This research methodology will outline the framework by which a transfusion inventory can be considered as a process. This methodology will be a high level overview of how to conceptually think of an inventory in generic terms, determine if process mining can be applied, and if the output of the process map is valid for further interpretation.

3.3.1.1 Process Data Assessment

This step encompasses an environmental scan of the business lexicon, or vocabulary, being used to describe the inventory and assesses whether these basic building blocks are being captured in the LIS. This process data assessment phase of the framework involves understanding the vocabulary of the transfusion inventory business process to determine if *activity* verbs and objects are being used at all in the BTS lexicon (Figure 9). One possible source of the lexicon of these activities is to examine the standard operating procedures (SOPs) of the BTS and determine if verbs and their respective objects can be identified in the body of the SOPs associated with inventory handling. In this example, the document management system folder housing the SOPs used by CZBTS was searched, but any generic document management or SOP directory can also be searched. Table 20 highlights examples of quotations taken directly from inventory handling-related SOPs and extracts the verbs and objects mentioned in the SOP. By demonstrating that inventory handling is described by the transfusion environment as verbs and

objects, the fundamental *activity* variable can therefore be reconstructed from them. This is one of the first steps to ensuring input data quality. Table 20 also highlights the use of timestamping and a unique barcode identifier for units within the inventory thus satisfying the *Case ID* uniqueness requirement and *Timestamp* variables. Based on the SOPs, the most fundamental elements of the process log are available. If the BTS lexicon cannot be located by examining the SOPs, then it may not be possible to apply process mining to the inventory, or it may be necessary to build such a dataset. Alternatively, it may be possible to do a manual audit of transfusion inventory handling events from receipt to final disposition of the blood product, including any intervening steps; however, this will lead to suboptimal datasets which are not amenable to process mining as was discussed in the Event Log Maturity Table on Table 7.

Another necessary step in characterising the lexicon is to determine if this BTS inventory handling lexicon is formally recorded in the LIS (Figure 9). The SOPs in Table 20 indicate that inventory is documented in a LIS (764- PLM BT Receipt of Red Blood Cells Procedure), and suggest that there is a requirement to at least receive RBC units into the inventory and document those within the LIS. The extent of the documentation of this receipt process, the existence of the log data elements and any processes following the receipt process documented in the SOP, needs to be verified using tools which are natively available to the LIS. For example, the institution used Cerner Pathnet Millennium™ platform and thus Product History Review and Complete Product History tools were used. Figure 10 highlights analogous variables which can be mapped to *Case ID, activities (a, b, c, d, e,…),* and *Time*. There are also potential candidate data elements which could be mapped to objects, verbs, and the resource/originator.

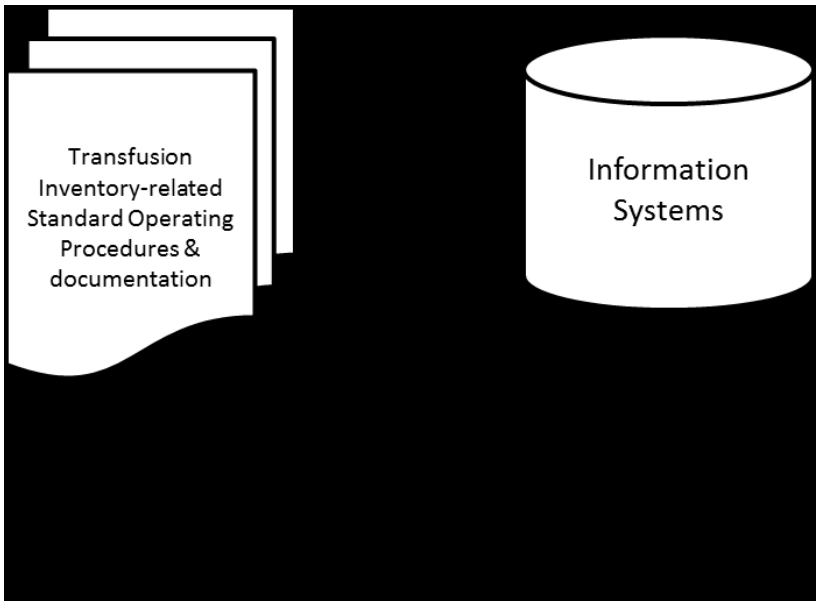


Figure 9: Examination of lexicon for process activities captured in both transfusion-related documentation and information systems

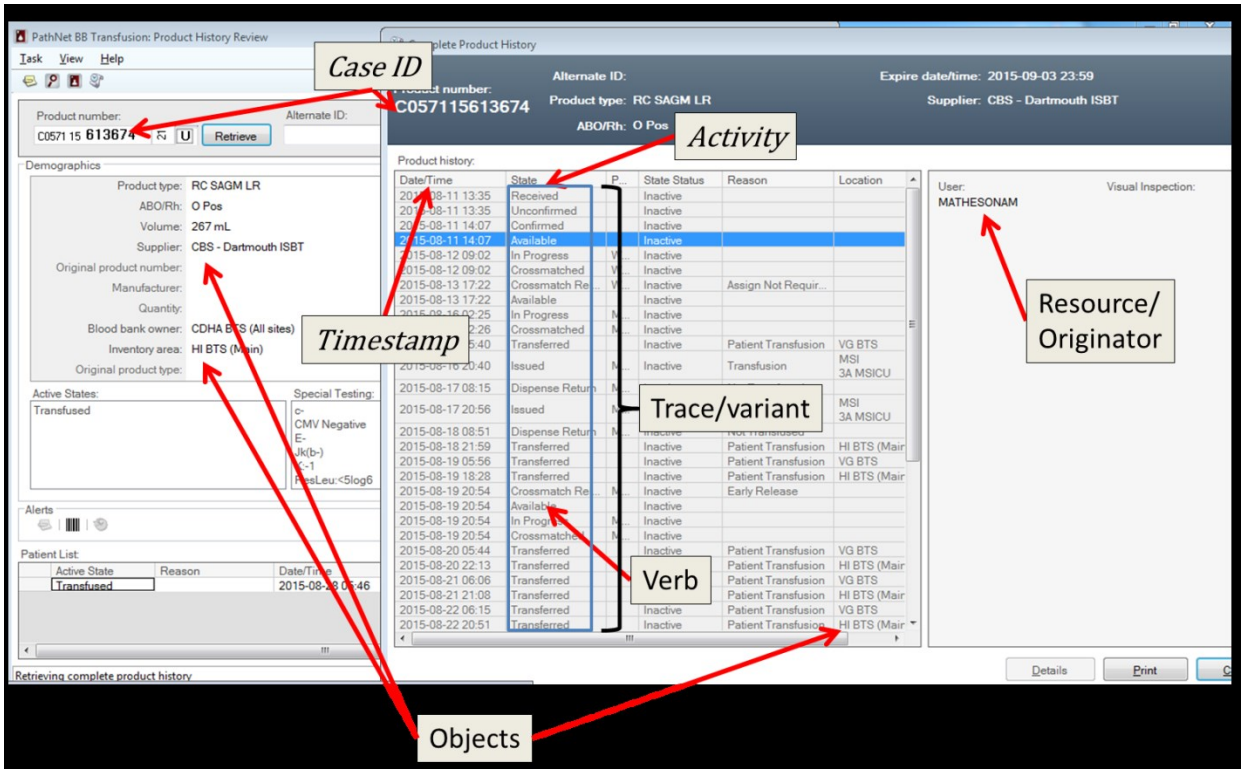


Figure 10: Annotated Product History Review and Complete Product History windows (Cerner Pathnet Millennium™) for unit C057115613674

If these variables are not readily available in native tools in the information system, then an extensive examination of the raw LIS data tables for the variables should occur. There will be two illustrations involving two different LISs, involving Cerner Pathnet Millennium™ (implemented in 2009) and Meditech Magic™ (implemented in 1989). The *Case ID, activities,* and *Time* were found in Cerner's Product and Product Event tables as shown in Figure 11 and Figure 12. Variables analogous to *Case ID, activities,* and *Time* were also found in Meditech Magic™'s *lab.b.unit.main* table as annotated in Figure 13, suggesting that it is possible to apply process mining to this older information system as well.

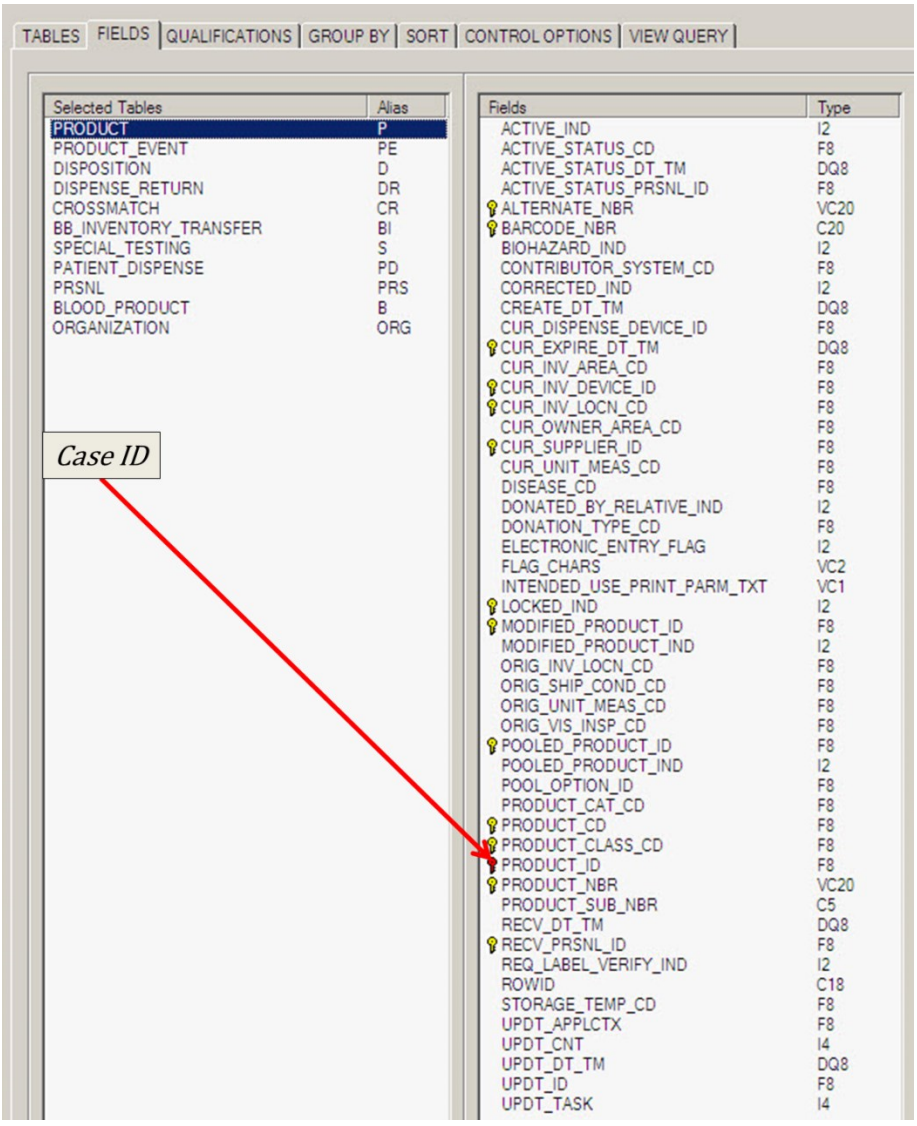


Figure 11: Case ID mapped to Product ID in Product Table (Cerner Pathnet Millennium™)

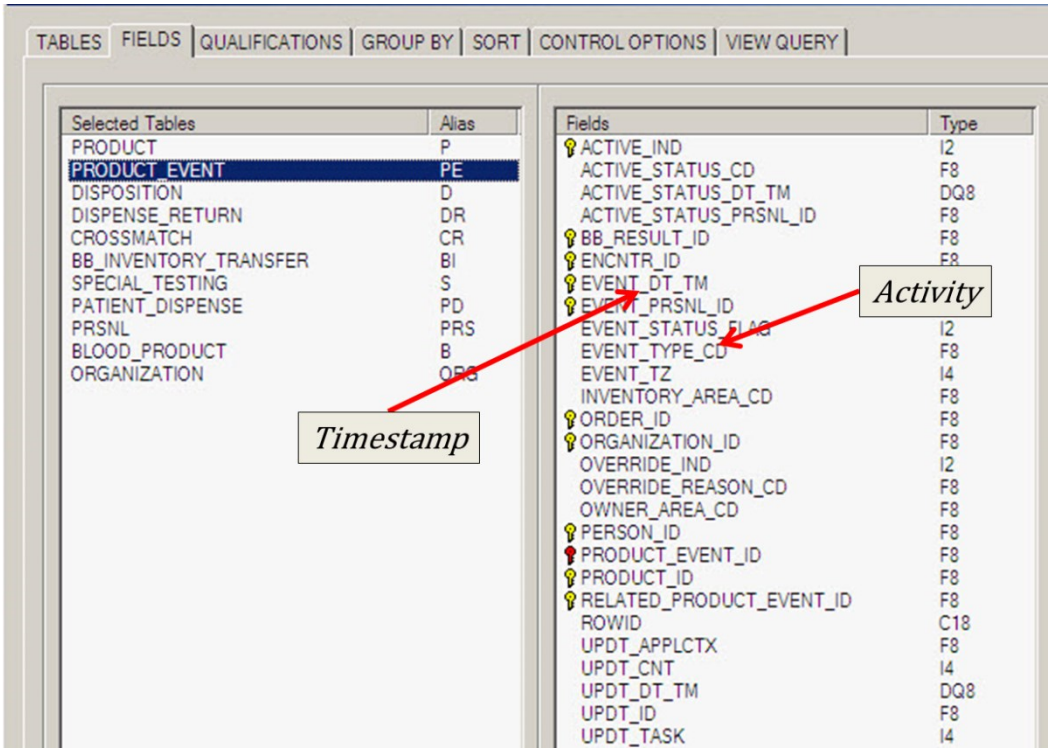


Figure 12: Timestamp and Activity mapped to EVENT_DT_TM and EVENT_TYPE_CD in Product Event Table (Cerner Pathnet Millennium™)

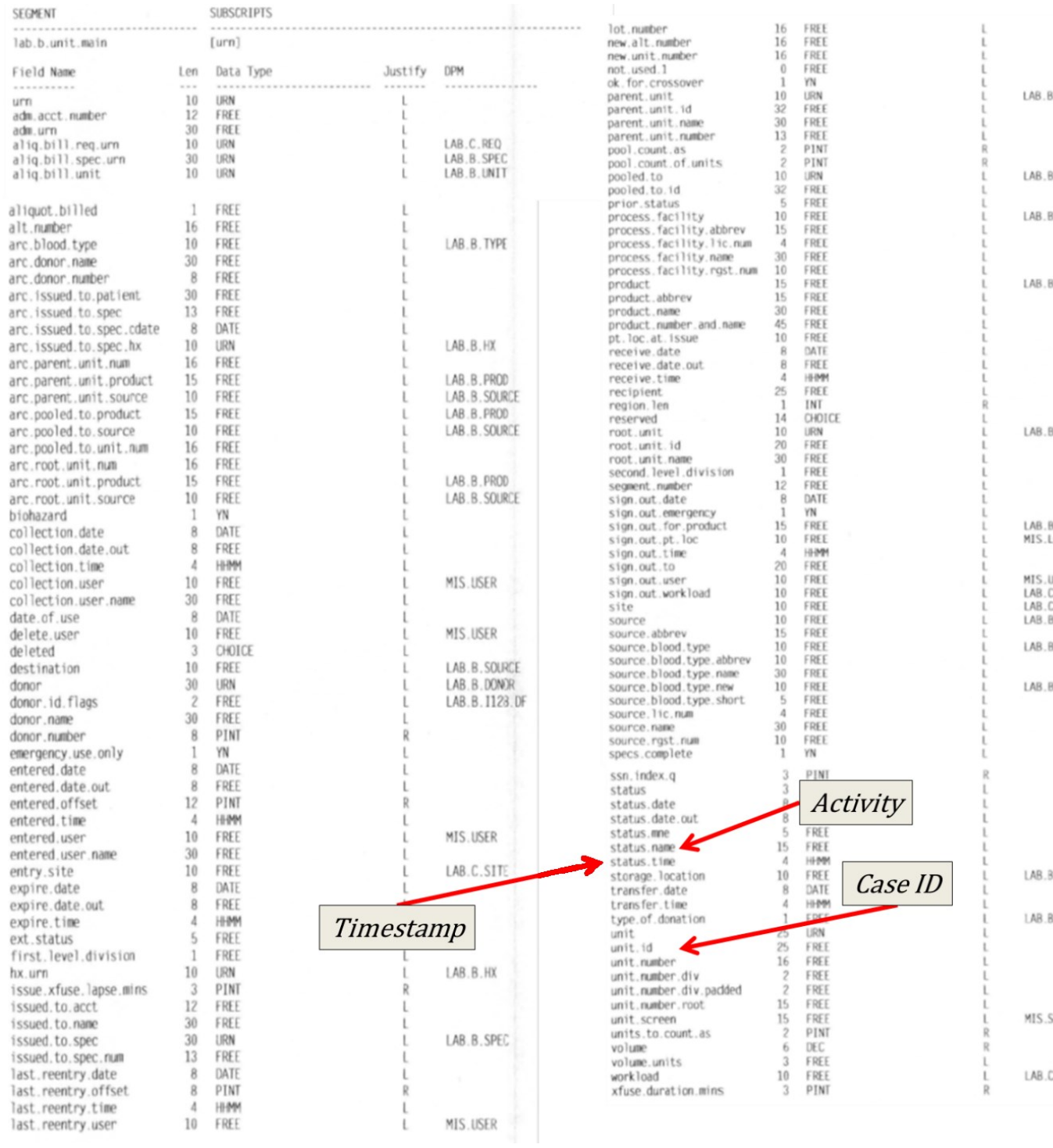


Figure 13: lab.b.unit.main (Main Meditech Table Pertaining to Red Cell Unit) Fields Mapped to Timestamp, Activity, and Case ID in the Meditech Magic™ LIS.

The only checkpoint in this data assessment phase is if the blood product inventory is manually captured, then one cannot proceed to the data preparation phase. This acknowledges that manual data capture would satisfy a level 1 or 2 on the event log maturity table, which is not sufficient as a process mining input (Table 7).

3.3.1.2 Data Extraction Phase: Iterative Process Log Query and Validation

The data preparation stage of the framework conceptually involves the query of a log file from the LIS, including iterative validation and preprocessing steps.

The first aspect of the extraction part of the framework involves the query and validation of a product process log, which may need to be done as this may not be natively available in the system. The query step is an iterative process where data elements which are analogous to *Case ID, activities (a, b, c, d, e,…),* and *Time,* are queried and compiled into a blood product process log. Since different data tables may be queried to construct the product process log, careful examination to ensure that timestamps, unit identifier, and activities are present on the extract if they are found in the Product History Review module. This is the validation step. Validation involves determining if the activity is present, and if the *Case ID, activities (a, b, c, d, e,…),* and *Time* variables are present and correctly associated with each other by comparing to the source system trace. For example, in Table 15, both query #1 and #2 have to match the source system trace. The issue with query #1 is that there is an incorrect activity that has been associated with $CaseID_1$ and in query #2, there has been transposition of timestamps for $CaseID_2$ and missing *activities* and *Time* elements for $CaseID_4$. These incorrect combinations of *Case ID, activities,* or *Time* compared to the source system trace are defined as *artifacts* in this research. The rationale for this step is to ensure the completeness and correctness of the data set. If there are data elements that are not queried, a search should be done to include those data elements on the subsequent query and the validation step should be re-initiated.

Table 15: Generic Validation Scenarios to Compare Source System Traces with Process Logs Generated from a Query of the Information System

Source system Trace (gold standard)	Example query #1 Trace	Example query #2 Trace
$CaseID_1, \{a, T_a\}_1, \{b, T_b\}_1, \{d, T_d\}_1, \{e, T_e\}_1$	$CaseID_1, \{x, T_a\}_1, \{b, T_b\}_1, \{d, T_d\}_1, \{e, T_e\}_1$	$CaseID_1, \{a, T_a\}_1, \{b, T_b\}_1, \{d, T_d\}_1, \{e, T_e\}_1$
$CaseID_2, \{a, T_a\}_2, \{b, T_b\}_2, \{d, T_d\}_2$	$CaseID_2, \{a, T_a\}_2, \{b, T_b\}_2, \{d, T_d\}_2$	$CaseID_2, \{a, T_k\}_2, \{b, T_a\}_2, \{d, T_z\}_2$
$CaseID_3, \{b, T_b\}_3, \{d, T_d\}_3$	$CaseID_3, \{b, T_b\}_3, \{d, T_d\}_3$	$CaseID_3, \{b, T_b\}_3, \{d, T_d\}_3$
$CaseID_4, \{b, T_b\}_4, \{d, T_d\}_4, \{a, T_a\}_1, \{b, T_b\}_1$	$CaseID_4, \{b, T_b\}_4, \{d, T_d\}_4, \{a, T_a\}_1, \{b, T_b\}_1$	$CaseID_4, \{d, T_d\}_4, \{b, T_b\}_1$

The practical validation of the generated process log can be done either on a random selection of units or can be done by examining units with the most complex traces, with the latter being preferred as there is likely a high probability of identifying an incomplete data extraction. Overall, the iterative process log query/generation and validation process should not stop until traces completely match the source system traces and/or there is a minimization of artifact generation. An example of this is shown in Figure 14 with the output of a generated log file validated against output from the Product History Review for a blood product with complex trace.

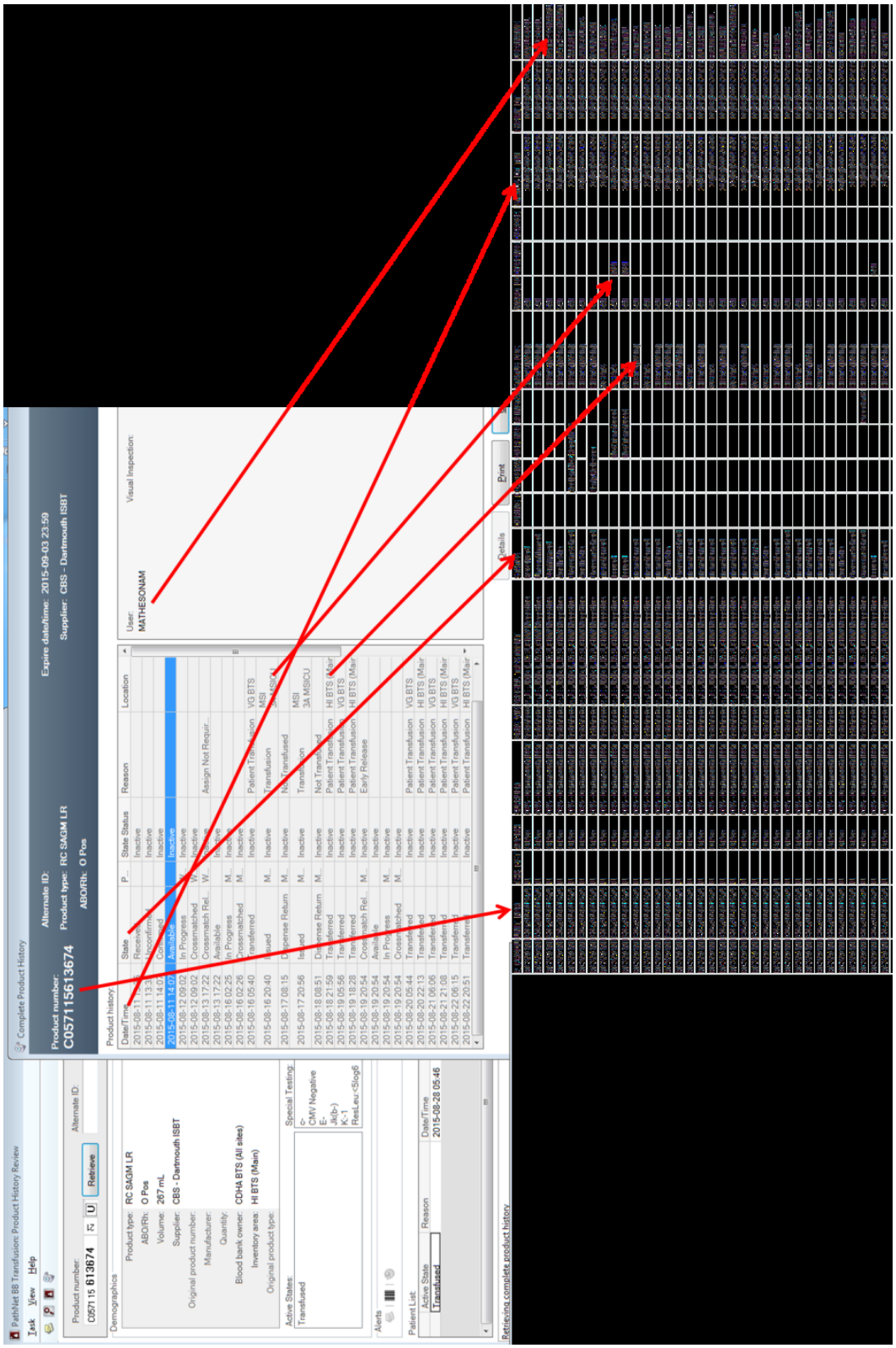


Figure 14: Validation of generated product log file (lower right corner) with output from Product History Review module of Cerner Pathnet Millennium™ (upper left-hand corner).

If there is no Product History Review-like module/function to display the gold standard trace (Figure 14), then this validation step cannot be done and there is a possibility that the data could be incomplete, resulting in an incorrect process not being generated, and potentially leading to incorrect conclusions. Conceptually, it may be possible to manually audit and follow units through their activities, but due to human factors it may not be possible to capture all possible transactional activities associated with individual blood products, as there may be transactions that are invisible to observation but that are still captured in the information system as an activity. Likewise, such a manual auditing process would be physically impossible, especially if a unit has transactions and activities over significant lengths of time and geography. If this entire validation step is not performed then process mining could still continue, but the downstream process mining results should be interpreted with caution.

Once the iterative validation step is completed, the file should be converted to a format which can be imported into the selected process mining tool(s). File formats can include XES (eXtensible Event Stream), MXML (Mining eXtensible Markup Language), and CSV (Comma Separated Values) files (processmining.org), but these formats are process mining tool dependent.

3.3.1.3 Data Preprocessing: Increasing Activity Descriptiveness and Specificity through Verb Object Concatenation and Validation

As the activities within the validated data set consist of discrete verbs and objects, the objective of this step is to concatenate verbs and objects into verb-objects to increase the specificity and descriptiveness of the activity field [48]. This step is iterative and concatenates activity verbs and activity-related objects while ideally maintaining a human readable format, and is repeated until all possible verb-object combinations have been exhausted as noted in Table 16. Each iterative concatenation must be validated against product data traces directly from a product history-type module noted above to ensure that no artifacts are introduced in the concatenation process. This is done using a selection of products featuring differing trace complexity and doing a comparison as noted in the generic examples in Table 16 and Table 17. If artifacts are introduced (*), such as in Table 17, with the joining of a verb and the incorrect object, another

iteration must occur with another object and the validation step repeated. The rationale behind increasing the descriptiveness of the activity field is to maximise the descriptive capacity of the resultant process map of the business process, similar to how using adjectives and adverbs in language increases the specificity and descriptiveness of a written language. This concatenation can happen in a database, spreadsheet, or directly in the process mining software, if the latter has those capabilities.

Table 16: Ideal Verb Object Iterations without Artifact Generation for Case ID.

Iteration Number	Verb and Object Combinations
Iteration 0 (starting point)	*Verb, object1, object2, object3, object4, object5 ….*
Iteration 1	*Verb-object1, object2, object3, object4, object5 ….*
Iteration 2	*Verb-object1-object2, object3, object4, object5 ….*
Iteration 3	*Verb-object1-object2-object3, object4, object5 ….*
Iteration 4	*Verb-object1-object2-object3-object4, object5 ….*
…	*…*

Table 17: Ideal Verb Object Iterations with Artifact Generation*for Case ID_x

Iteration Number	Verb and Object Combinations
Iteration 0 (starting point)	*Verb, object1, object2, object3, object4, object5 ….*
Iteration 1*	*Verb-object1, object2, object3, object4, object5 ….*
Iteration 2	*Verb-object2, object3, object4, object5 ….*
Iteration 3*	*Verb-object2-object3, object4, object5 ….*
Iteration 4*	*Verb-object2-object4, object5 ….*
Iteration 5	*Verb-object2 -object5 ….*
…	*…*

This step of verb-object generation does not need to occur strictly in the data pre-processing stage, as it can also occur in the data exploration and inspection stage, especially if it is unknown how a process map may look to the user with increasingly complex verb-object activities.

3.3.1.4 Data Exploration and Inspection

The data exploration step has been previously described by De Weerdt *et al.* as a 'first impression', or a first look, of the data using the process mining tool [48]. It is a step where the time-dimension and the scope of the process should be limited or expanded according to the context of the process mining tool being used. For example, if too much data has been extracted that is of too broad a scope or time dimension, the process mining tool may not be able to visualise the data due to limitations in computation or the process mining may result in a complex process map that is not human-interpretable. With an overly broad scope or dimension, the extracted data would have to be limited on extraction, or using post-processing techniques and then reimported. Likewise, if scope or the time dimension is defined too narrowly, then there is a potential to exclude rare, episodic, or seasonal events. Therefore, data extraction would need to be broadened with respect to the time dimension.

The first step of data exploration is to select the correct tool to suit the business requirements and the data. Selection of the process mining algorithm can be done in different ways. The first way is functional and tool-centric, beginning with an understanding of how the tool is to be used by the blood transfusion service users and how the tool is to be licensed. If resources are not a limiting factor, a tool that has the simplest user interface will be most readily used by non-expert users. This can be done on a trial and error basis and is a combination of art and science. The second approach is to understand the strengths and limitations of individual algorithms and make a selection based on that. The mathematical analysis of the algorithms is outside the scope of this research and has already been extensively performed by De Weerdt *et al*. These authors compared the qualities of algorithms in detail, examining accuracy and comprehensibility using F-score analysis as an evaluation approach [129]. A third approach is a hybrid of the first two and considers the business process of product inventory management, in the context of the nonexpert user, and selects a software package that is easy to use and has an algorithm which is consistent with the business process. The most fundamental construct of a transfusion product inventory is that, apart from potentially the receipt process, where inventory is handled in a structured way, the remaining process is unstructured, as units are free to interact and respond to an environment that has stochastic demand. In contrast to a

business process, where there is an ideal way of handling a process for optimal results, transfusion products do not conceptually obey a typical structured process model. The Fuzzy Miner algorithm can deal with minimally structured and unstructured processes, and due to its ability to visualise at different levels of abstraction, it is an ideal data exploration algorithm for visualizing the transfusion inventory product flow[129, 141]. The serendipitous attribute of the Fuzzy Miner algorithm (Table 9) is that it is packaged in a user-friendly interface in the Disco (www.fluxicon.com) process mining software. Therefore, Disco is an ideal software package to use for the transfusion inventory analytical environment.

Data exploration can also consist of generating statistical or value-based information about the dataset which has been queried, and provides an entry-point to allowing the value-based and process-based methods to be complementary. This portion of the framework has been adapted from De Weerdt *et al*.[48].

The next step in data exploration is to use the selected process mining tool to generate the first map and do a soft qualitative evaluation with the end-user, an inspection at a subjective visual level. Table 18 is a list of possible questions to informally ask the end-users and obtain their feedback. Ideally, a selection of end-users involved in different aspects of the blood product inventory handling process at the institution would be asked to provide informal initial feedback. This list is not exhaustive and is intended to act as a checkpoint. If issues have been identified which need to be addressed then a return to a prior stage in the analytical framework should be done prior to proceeding to the analysis stage. The rationale for doing this final step prior to proceeding to the analysis stage is to ensure that the process maps are valid and also usable, and that time will not be wasted in subsequent analyses using either incomprehensible or invalid maps.

Table 18: List of Questions to the User about Their First Impression of the Generated Process Map

Questions relating to Qualitative visual inspection of Process Map
Is the process map human readable?
Does the process map make sense?
Does the process map appear correct?
Is the process map consistent with what you know about your operational practices and processes?
Does the process map appear to be complete, and is there anything that is missing?
Are there aspects of the process map that need clarification?
Are there outright errors of the process map?
Does the process map cover the correct time scope?
Does the process map contain the data to answer the business process question?

The subsequent steps in the framework incorporate the remaining steps of the PMMF by De Weerdt *et al.* [48], beginning with the *perspectivization* step. This describes the concept that the process log data can be examined from a case standpoint, where sequences of activities are examined; from a performer viewpoint, where interactions between sequences of teams can be examined; and from an attribute standpoint, where different cases are analysed according to their attributes. These different perspectives are discussed in Table 19.

Table 19: Application of Different Perspectives ('*Perspectivization*'[48]) on Process Data to Transfusion Product Inventory Analysis.

Perspective from De Weerdt[48]	Example applied to transfusion product inventory
Control-flow (business activity)	- Basic flow of products from receipt to destination - Recycling of blood products from other sites in institution - Use of blood fridges - Product handling at different institutions
Case data (document types)	- The use of blood products that have special attributes, such as Cytomegalovirus negativity, or irradiation - The pathways that are attributed to wasted units
Organisational (team routing)	- What personnel are involved in handling blood products, such as during product receipt, dispense, or disposal?

3.3.1.5 Summary of the Transfusion Inventory Process Mining (TRUISM) Framework

The overall goal of this Transfusion Inventory Process Mining Framework is to output a valid human-readable process map of the blood product inventory and allow for subsequent analysis to be conducted with this output in any form. The framework involves determining the intersection of process-based elements of the transfusion inventory management lexicon with the transfusion inventory management system, extracting and increasing the specificity of the activity-related data elements (verb-objects), and ensuring that there is validation of all data manipulation steps. This framework will be called the **TR**ansf**U**sion **I**nventory proce**S**s **M**ining Framework (TRUISM Framework), which represents the mechanism by which transfusion product inventories can be understood from a process-based perspective (Figure 15).

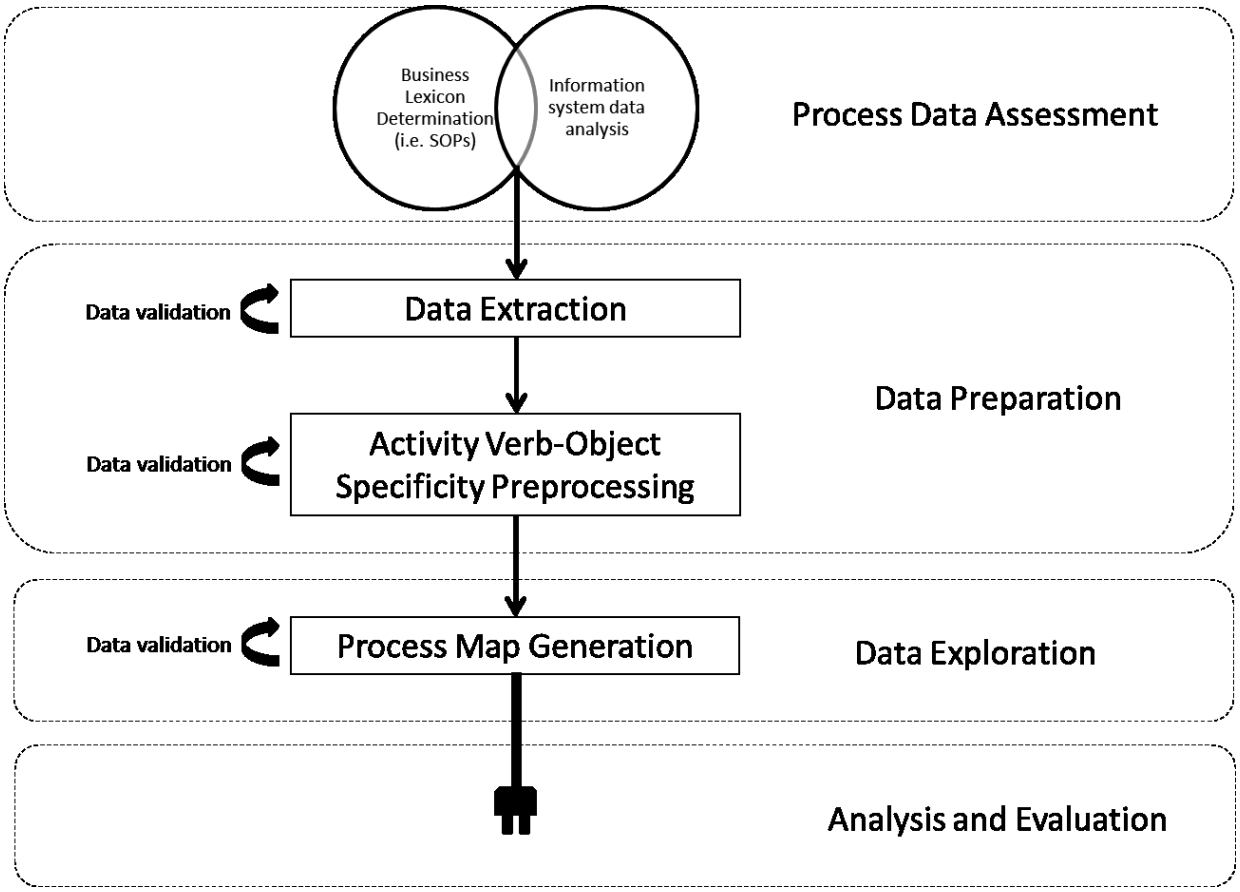


Figure 15: TRansfUsion Inventory proceSs Mining Framework (TRUISM Framework)

The objective of the TRUISM Framework is to be able to 'plug' the validated output (i.e. process mining-generated map) with confidence into any evaluation process and be able to generate further process-based insight. For instance, in De Weerdt [48], the authors describe a case study-based framework where the success of the process mining exercise is ultimately based on the creation of guidelines for process improvement. The TRUISM Framework increases the rigor and confidence in the data which would be subsequently imported into a process mining tool.

3.4 Application of the TRUISM Framework to RBC Inventory Data

The subsection will describe the application of the TRUISM Framework described in the prior section to RBC inventory data of the CZBTS and demonstrate how each step is conducted.

3.4.1 Process Data Assessment

A process data assessment was carried out by analysing the SOPs of the NSHA, CZ (Table 20) and it was determined that the terminology existed in the lexicon and in the LIS, and was of Level 3 quality according to the Event Log Maturity Table (Table 7). Information was being entered by the laboratory technologists at all transactional stages using a barcode scanner and there was full electronic documentation into the laboratory information system. Therefore, the data preparation step was able to continue.

Table 20: Standard Operating Procedures, Quotations Related to Transfusion Inventory Process, and Actions Pertaining to Products

Standard Operating Procedure at NSHA, CZ	Direct Quotations from Inventory SOP	Verbs/actions and attributes relating to inventory process	Objects mentioned
763: PLM BT Inventory Management and Ordering From Canadian Blood Services Procedure	*"Receive red cell units returned from OR, and release units from crossmatch"* *"release units from crossmatch"* *"Return units of blood in Emergency department blood fridge at HI site, VG 11BF Fridge, MSI (3A) Fridge, and CCHC Emergency that are 14 days from expiry, to HI BTS."*	- Receive - Return - Release	- Red cell units - HI site - VG 11BF Fridge - MSI (3A) Fridge - CCHC Emergency - HI BTS
752 PLM BT Storage of All Blood Components and Derivatives Policy	*[Technical staff] "Moves blood components and derivatives to alternate storage areas during a storage temperature deviation"* *"Quarantines and/or disposes of blood components and derivatives when applicable"* *""Rejected, outdated and quarantined blood components and derivatives are put in a designated area separated from stock and crossmatched blood components to prevent accidental use of these products."*	- Move - Quarantine - Reject - Outdate - separate	- blood components and derivatives - alternate storage areas
777: PLM BT Release of Crossmatched and Assigned Blood Components Procedure	*"Batch Release Schedule [Release of Red Blood Cells-Crossmatch expire]* *HI Site- three times daily- 05:30, 13:30, 21:30* *VG Site- three times daily- 05:30, 13:30, 21:30* *DGH Site- three times daily- 07:00, 15:00, 23:00* *HCH Site- twice daily- 07:00, 19:00"* *"24 Hour Release of Crossmatched Red Blood Cells"* *"Early Release of Crossmatched Red Blood Cells"*	- release - crossmatch release - crossmatch expire - crossmatch early release	- HI Site - DGH Site - HCH Site - Red Blood Cells
750 - PLM BT Packaging Blood Components and Derivatives For Transport to Outside Facilities Procedure	*"Visually inspect the units for the following:* *a. Appearance: not black or discolored, no clots, etc.* *b. Expiry date* *c. The blood bag closure is undisturbed"*	- visually inspect	- Units [blood] - Blood bag
761 - PLM BT Transfer and Receipt of Blood	*"Transfer of Blood Components "* *"Select blood component(s) to be transferred.* *NOTE: Components must be in a transferrable inventory*	- Transfer [initiate] - Available - Crossmatched	- Blood Components

Standard Operating Procedure at NSHA, CZ	**Direct Quotations from Inventory SOP**	**Verbs/actions and attributes relating to inventory process**	**Objects mentioned**
Components and Derivatives Between Sites Procedure	*status* *such as:* *a. Available* *b. Crossmatched* *c. Assigned"* *"Receive Transferred Blood Components"*	- Assigned - Receive transfers	
6398 -PLM BT Visual Inspection of Blood Components and Derivatives Procedure	*"Visual Inspection of Red Blood Cells"* *"inspect the unit of RBCs for signs of leakage and ensure the port covers are intact and do not have blood or plasma in them."* *"Disposition of Unacceptable Blood Components or Derivatives"* *"Dispose of visually unacceptable component or derivative using the* *Final Disposition Application in LIS, entering NCE number as a product comment."*	- Inspect - Dispose - Ensure	- Red Blood Cells - unit of RBCs - port covers - Blood Components or Derivatives - product
772 - PLM BT Dispense of Blood Components and Derivatives Procedure	*"Dispense of Blood Component and Derivatives"* *"Retrieve blood component or derivative from designated storage area.* *Visually inspect product per PLM BT Visual Inspection of Blood Components and Derivatives Procedure."* *"Enter blood component by barcode or manually enter derivative lot number in product field"* *"d) Ensure patient location is correct on Transfusion Tag, or correct if applicable.* *e) Time stamp derivative Transfusion Tag at time of dispense."* *"The MLT will retrieve the appropriate unit from the Blood Bank refrigerator and visually inspect the product."* *"The MLT will barcode the unit number of the product in the LIS.* *d) Verify with the RN/LPN/ACP the component/tag information"* *"Red cells for transfusion must be transfused before the expiration of the crossmatch and/or expiry date of product (within 96 hours from the time of collection of the specimen) per Policy CC75-005 Blood Transfusion: Administration of Blood, Blood Components and Blood Products."* *"Red Cell Transfusion Tags are not time stamped if they*	- Dispense - Inspect - Retrieve - Enter barcode of unit - Enter lot number - Ensure transfusion tag correct - Time stamp - stamped - retrieve unit - verify transfused - ensure	- Blood Component and Derivatives - storage area - product - patient location - Blood Bank refrigerator - refrigerated storage location

Standard Operating Procedure at NSHA, CZ	Direct Quotations from Inventory SOP	Verbs/actions and attributes relating to inventory process	Objects mentioned
	are going to a refrigerated storage location. The tag will be stamped when personnel remove them from those locations at time of transfusion."		
764- PLM BT Receipt of Red Blood Cells Procedure	*"Time-stamp the CBS packing slip with the time received."* *"Remove the Tamper Indicator Device from one shipping box at a time."* *"Remove the RBC units from the shipping crate."* *"Perform the following steps if the Tamper Indicator Device is missing:* *a. Enter RBCs into LIS inventory* *b. Quarantine and dispose of RBCs in LIS using Final Disposition* *application with the Final Disposition reason "Security Tag* *Broken or Absent"* *c. Initiate an NCE* *d. Complete a CBS Customer service letter and fax to CBS if the* *supplier is CBS"* *"Enter RBC units into inventory in the LIS using the Receive Products application. Enter the correct supplier name according to Table 1. LIS Supplier Names in the Procedural Notes section. Enter attributes such as CMV Negative.* *NOTE: Units must be entered in LIS by bar code. Do not manually enter the information."* *"Scan the LIS Products Received Report to ensure all RBCs have the* *correct number of digits, correct product code and an ABO/Rh."* *"RBC units are routinely received into the BTS LIS inventory at the HI, DGH, HCH sites. RBCs are not routinely delivered to the VG site or the Tri-facilities."* *"The RBC units must be entered into the BTS LIS inventory using the BAR CODE wand, as this will reduce the clerical error entry. Do not enter the RBC units manually."*	- Timestamp - remove tamper indicator device RBC units - quarantine - dispose - receive products - enter units by barcode - take out units - routinely received	- Tamper Indicator Device - RBC units - supplier name - inventory locations of HI, DGH, HCH, and VG.

3.4.2 Data Preparation (Data Extraction):

The time period (scope) for the data extraction was chosen to be June 2009 to June 2015 inclusively. This was to reflect current blood transfusion service inventory handling workflows which are commonly documented through standard operating procedures and understood by the technologist, manager, and physician community associated with blood transfusion services. Prior to May 2009, the LIS at Capital Health Authority (now Nova Scotia Health Authority) was using Cerner Classic and since there is no currently used formalised documentation from that era, the time period prior to May 2009 will not be considered.

Data from all hospital sites which are able to hold a red blood cell inventory was queried, including Hants Community Hospital (HC)[Windsor, Nova Scotia], Dartmouth General Hospital (DGH)[Dartmouth, Nova Scotia], Victoria General Hospital(VGH)[Halifax, Nova Scotia], and Halifax Infirmary(HI)[Halifax, Nova Scotia]. All possible endpoint locations and reasons for product disposition and transfusion were included from inside the formal health region system and other health regions in Nova Scotia.

Ethics approval necessity was sought in an e-mail dated May 30, 2012 to K.J. of the Research Ethics Board at Capital Health Authority, but deemed to not be applicable as it was quality assurance in nature. The data set contained quality assurance inventory data, no patient involvement, no patient identifiable data, and any data collection and analysis was done within routine laboratory quality assurance practices.

Data was extracted from the source system (Cerner Millennium™) using Cerner Command Language™ (CCL) and the extraction was facilitated by PInG staff members (A.K-M, S.M) knowledgeable in CCL and the Cerner blood transfusion service module at Nova Scotia Health Authority, as it required complex CCL query and knowledge of in-depth table structures in the LIS. The final report originated from a complex query of data pieced together from various Millennium™ tables. The query was placed in Discern Explorer (a tool for running the CCL-based

query) and could be executed in an ad-hoc fashion. The output from the query was in a comma separated variable (CSV) file in one-month blocks and with the appropriate column headers noted in the leftmost column (Field Name on final query of table) on Table 21. Each CSV file generated was named according to the convention of AK_RBC_UNIT_TRANSITIONS-YYYYMM.csv, whereby YYYY was the four digit year and MM was the two digit month, and the prefix before this was constant. These files were combined using the 'copy' function in Windows 7 command line, into the master CSV file which contained all the data from the whole time period. The query was validated using the data TRUISM Framework (Figure 15) with random red cell units. To ensure that most or all available product transitions captured during day-to-day red cell unit handling workflow are included in the query, a random audit of approximately 20 RBC units of varying transactional complexity was done, from products with a few transitions to products having numerous transitions. Validation of the product STATUS codes and timestamps were done using a comparison of the output generated by the Product History Review tool and Complete Product History function of Cerner Millennium™ and a test query of a subset of data from a random month of output in a CSV file imported into Microsoft Excel 2010. During the first data validation phase, the "crossmatch release" and "dispense return" STATUS states were present in the Complete Product History but not present on the CSV output. These states could not be found in a search of the Millennium™ tables, but 'CROSSMA_RELEASE_REASON' and 'REINSTATE_REASON' were subsequently added to the query, as they were the closest match to the missing states and matched in terms of timestamp also. During that process, it was discovered that the 'in progress' status had a similar time stamp as the other STATUS variables. This variable was dropped prior to importation into Disco as it was considered a redundant variable which would increase the complexity of the process map generated but would not yield useful information. The data validation process was done in conjunction with the transfusion database coordinator S.W. with the PInG at NSHA.

The core process log (Table 13) consisted of the activity (STATUS, Table 21), date and timestamp (STATUS_DT_TM ,Table 21), and unique product identifier (UNIT_NBR, Table 21). In terms of a hypothetical inventory (Figure 16) process map adapted from [36], the STATUS or activity would represent verb-objects referring to the smaller rectangles, a count of activities of UNIT_NBR would represent the # of units in that STATUS, and timestamp calculations (STATUS_DT_TM)

would allow for the determination of time between STATUS activities. As the objective was to eventually obtain user feedback, additional information was queried, including ward locations, inventory locations, product attributes, which allowed for post-query filtering of the dataset if required by the eventual users (Table 21).

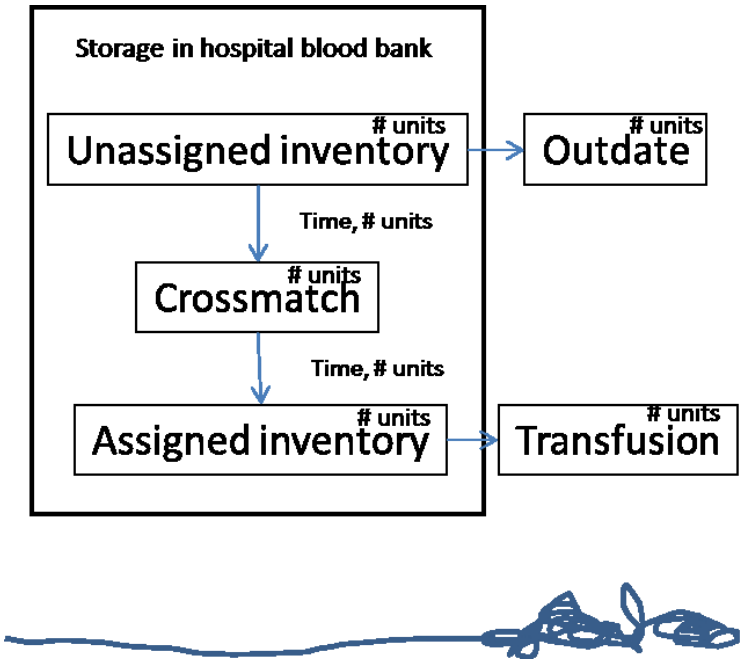


Figure 16: Hypothetical process map of a blood product in transfusion services (adapted from [42])

Table 21: Queried Field Names, Attributes, and Original Cerner Millennium™ Database Table References

Field Name for CSV file	Millennium™ Table	Millennium™ Field Name within Millennium™ Table	Attribute Description	Type of Attribute	Example data
P_PRODUCT_DISP	PRODUCT	PRODUCT_CD	Short form of description of blood product, derived from blood product label.	Textual	RC SAGM
UNIT_NBR	PRODUCT	PRODUCT_NBR	Unique unit identification number assigned to each product, directly derived from the label on blood at the time of receipt in blood bank, or if it is a plasma protein product, it is the lot number.	Alpha numeric textual	C057115236669
CHK_DGT	N/A		Checksum digit generated based on Donation Identification Number; used to verify manual entry of DIN.	Alpha numeric textual	1
ABORH	PRODUCT_ABORH	PRODUCT_ABORH_CD	Blood group of product	Textual	O Neg
SUPPLIER	PRODUCT	CUR_SUPPLIER_ID	Identifies the unique organisation who sent the blood product to the transfusion service.	Textual	CBS - Dartmouth ISBT
ORIG_VIS_INSP	PRODUCT	ORIG_VIS_INSP_CD	Description of the satisfactory or unsatisfactory nature of the product visual inspection at receipt.	Textual	Satisfactory
SPECIAL_TREATMENT	SPECIAL_TESTING	SPECIAL_TESTING_CD	Defines any special testing or attributes (CMV status, antigens) that a product possesses at the time of receipt into blood transfusion inventory or at time of product verification.	Textual	c-, E-, CMV Negative
STATUS	PRODUCT_EVENT	EVENT_TYPE_CD	The type of event or action that occurred with the product at a	Textual	Available

Field Name for CSV file	Millennium™ Table	Millennium™ Field Name within Millennium™ Table	Attribute Description	Type of Attribute	Example data
			specified time. Usually, expressed as a verb.		
DURING_DOWNTIME	From a custom table		Variable used to ensure that all available states, including those in downtime, were queried for a given product.	Y(or null)	Y
CROSSMA_RELEASE_REASON	CROSSMATCH	RELEASE_REASON_CD	The reason that a crossmatched product is released from the patient.	Textual	Assign Not Required
DISPENSE_RETURN_REASON	DISPENSE_RETURN	RETURN_REASON_CD	The reason the product was returned to the blood bank and not transfused to the patient.	Textual	Not Transfused
REINSTATE_REASON	CROSSMATCH	REINSTATE_REASON_CD	If this crossmatch was released and then later reinstated, it is the reason for that reinstatement.	Textual	Preadmit Patient
STATUS_LOC	PRODUCT_EVENT	INVENTORY_AREA_CD	The last known inventory location associated with the last STATUS field	Textual	HI BTS (Main)
WARD_LOCN	PRODUCT_EVENT	INVENTORY_AREA_CD	Contains the last inventory area associated with product event.	Textual	52
DISPENSE_LOCN	PATIENT_DISPENSE	DISPENSE_TO_LOCN_CD	Location at which a product is dispensed to.	Textual	11A
DISPOSE_REASON	DISPOSITION	REASON_CD	The reason this product was disposed of.	Textual	Product Expired
STATUS_DT_TM	On many tables	ACTIVE_STATUS_DT_TM	Date and time of status	Date and time	20/02/2015 14:53
EXPIRY_DT	PRODUCT	CUR_EXPIRE_DT_TM	Time of expiry entered at time product is received, from the label, or if product is modified or corrected.	Date and time	01/04/2015 23:59

Field Name for CSV file	Millennium™ Table	Millennium™ Field Name within Millennium™ Table	Attribute Description	Type of Attribute	Example data
USERNAME	PRSNL	ACTIVE_STATUS_PRSNL_IDRECV_PRSNL_ID	Represents the unique identifier of the user or personnel that cause the status to be set or changed.	Textual	CHENGC

3.4.3 Data Preparation (Increasing Activity Verb Object Specificity) Combined with Data Exploration

As the process tool had been selected prior to this stage of the TRUISM Framework and was found to have this function native to the importation interface, it was functionally possible to combine the iterative verb-object specificity stage of data preparation with the data exploration and inspection stage. This involved multiple iterative cycles of creating verb-objects through a brute force trial and error technique as within Disco, activity verb-objects can be constructed and re-constructed by mapping them multiple times to Disco's activity attribute(Figure 17). UNIT_NBR was logically mapped to ID and was not part of the verb-object iterative process (Figure 18).

Figure 17: Example of multiple mapping of verb-object fields pertaining to activity field in Disco

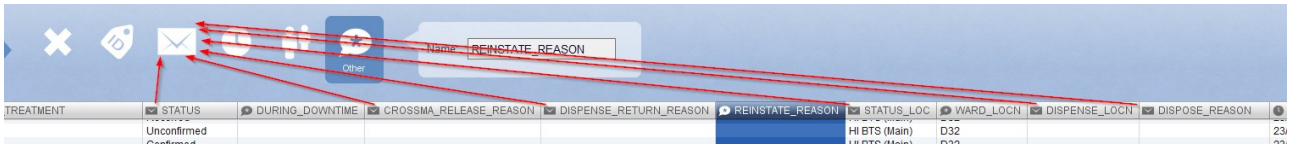
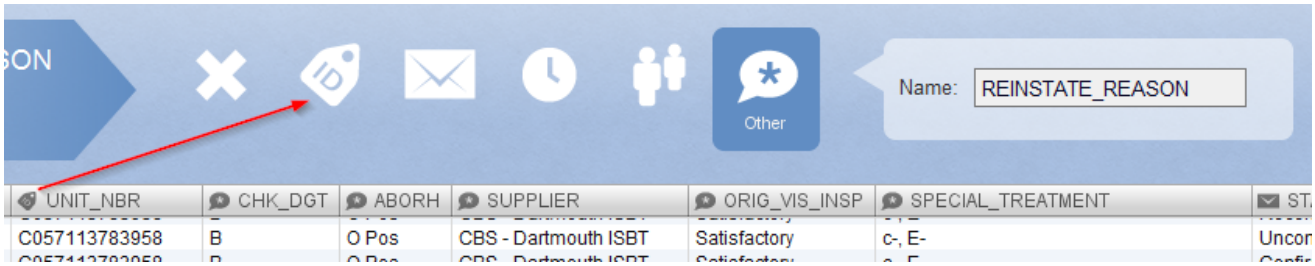


Figure 18: Example of regular mapping pertaining to data importation into Disco highlighting UNIT_NBR mapping to 'ID' field

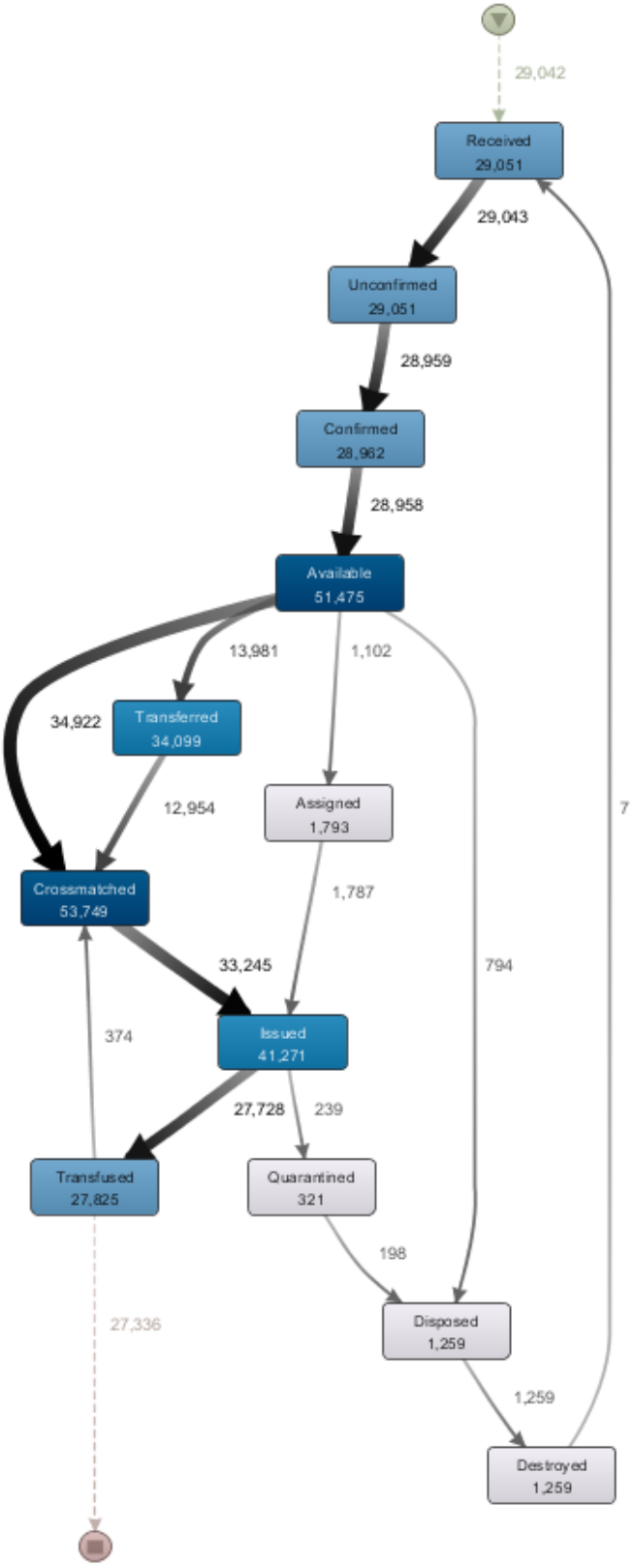


There were three phases of exploration, each of which were combined with the data preparation stage due to the inherent functionality of the process mining software. The first phase was a general exploration to optimise the 'activity' verb-object specificity. The second major phase included optimizing the mapping from the first phase for the purposes of using it in the RBC wastage focus group and the third major phase optimised the mapping from the second phase for use in the Hemosafe^{TM} fridge focus group.

The first exploration cycle involved the direct importation into Disco 1.9.5(Fluxicon.com). The data importation involved mapping the core attributes of the process log, UNIT_NBR and STATUS to 'Case ID and "Activity" respectively(Figure 18). "Timestamp" and "Resource" available with the correct formatting, and were mapped without modification to STATUS_DT_TM and USERNAME respectively (not shown). All other variables were mapped to 'Other'. The output generated using this level of mapping was a generic process map which was not informative, as the nodes were not descriptive enough to allow the process to be resolved(Figure 19). Assessment for location and specificity of the activity attribute could not be done, as equivalently named activities happening at different sites could not be distinguished on the process map(Figure 20) or using the output of the Cerner Complete Product History window. Known missing fields, such as Dispense Return, were not present in the map, but there were no newly-generated artifacts in the process mining software. It was noted that the generation of this first-exploration cycle process map was visually similar to the hypothetical process map adapted from the 'storage in hospital blood bank' box in Figure 16, adapted from Katsaliaki *et al.*

[42] Validations are following each iteration are summarised using Table 22 with the final activity verb-object combinations in Table 26.

Figure 19: First iteration of data exploration with a process map from Disco with core process map variables mapped.



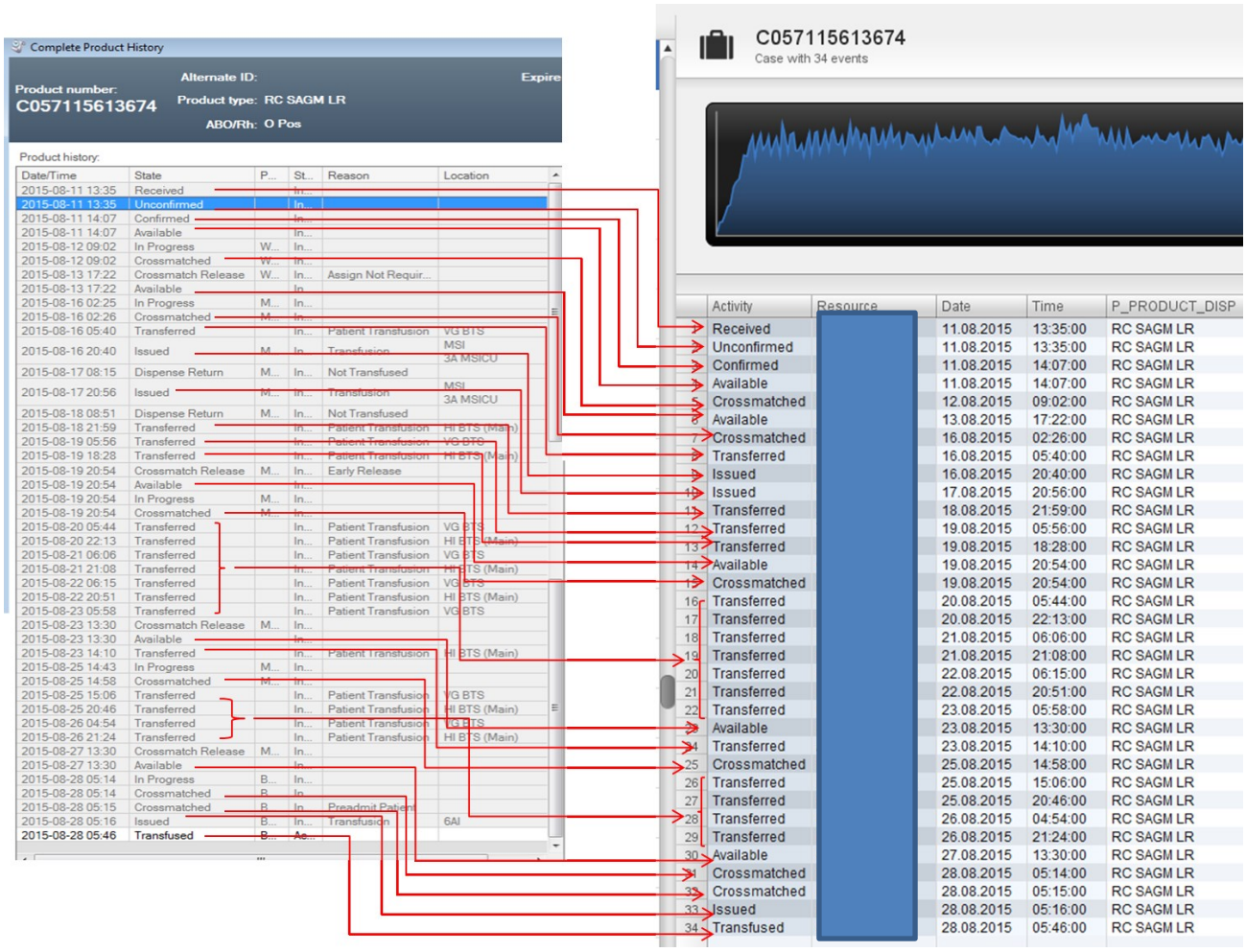


Figure 20: Example of first iteration/exploration cycle of unit C057115613674 with comparison between Complete Product History and Disco 'Cases' tab outputs (red lines indicate concordance).

The second iteration cycle involved an incremental increase in the activity mapping to include both STATUS and STATUS_LOC, as noted in Table 26. This yielded increased separation between the activities and allowed them to be attributed to location; however, transfusion destinations and crossmatch destinations were a challenge to visualise on the process map (Figure 21) and on the Complete Product History/Cases comparison(Figure 22). There were no artifacts introduced.

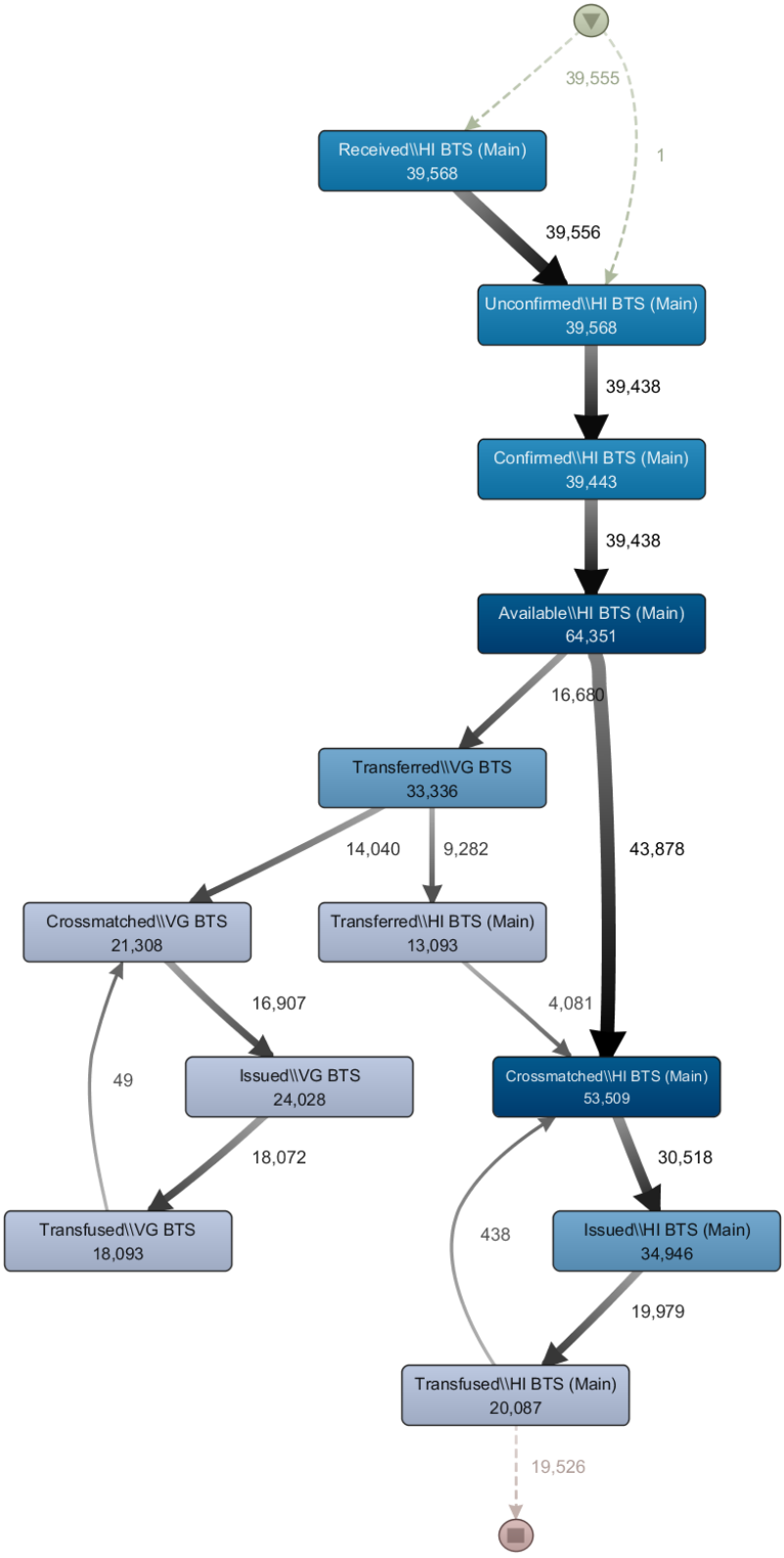


Figure 21: Second iteration of data exploration with process map generated from Disco with additional STATUS_LOCN mapped

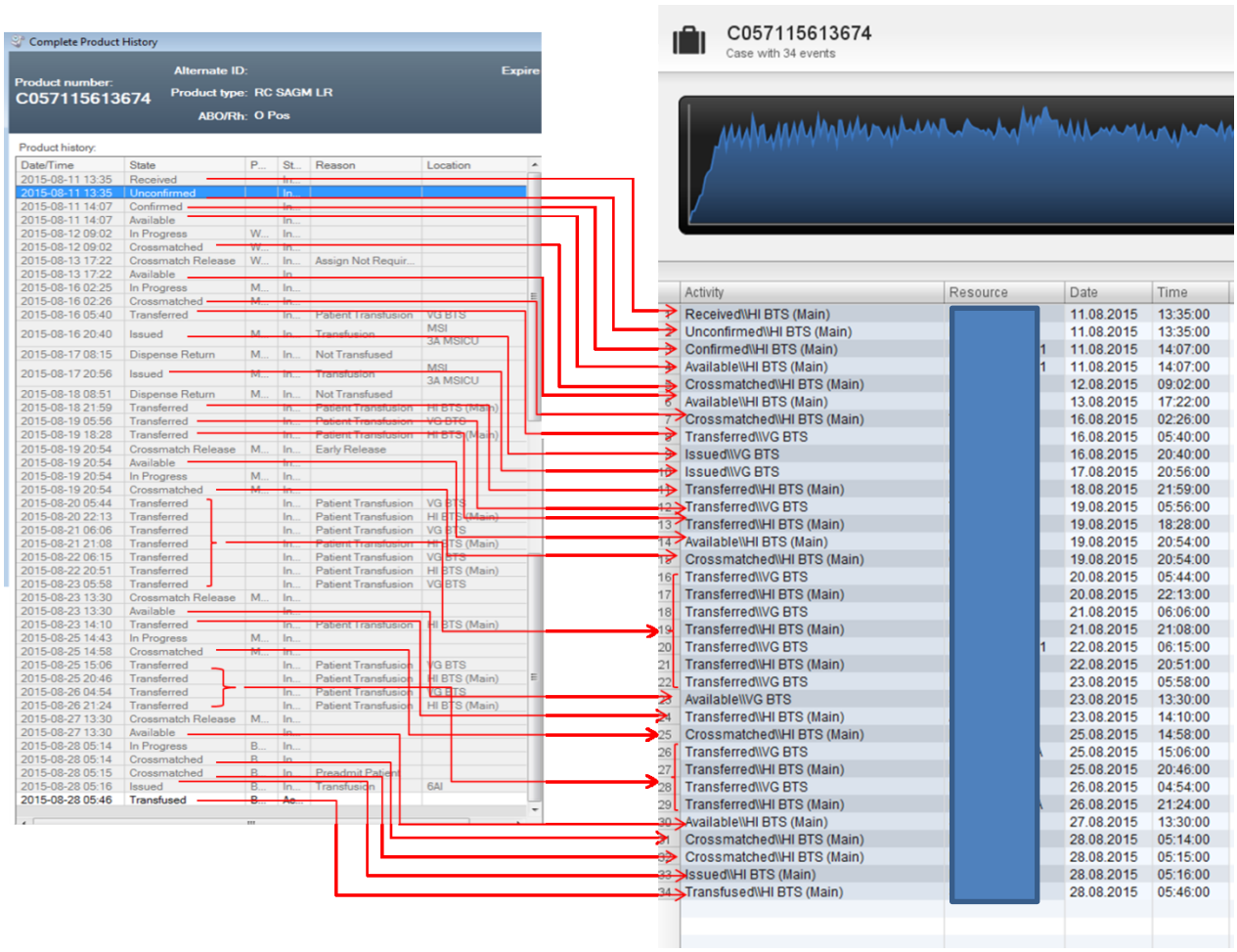


Figure 22: Example of second iteration/exploration cycle of unit C057115613674 with comparison between Complete Product History and Disco 'Cases' tab outputs (red lines indicate concordance)

The third iteration cycle involved an incremental increase in the activity mapping to also include DISPENSE_LOCN, as noted in Table 26. This yielded increased separation between the activities and allowed them to be attributed to location, especially at the transfusion destination, but crossmatch destinations were still a challenge to see on the process map (Figure 23) and on the Complete Product History/Cases comparison(Figure 24), necessitating a fourth iteration.

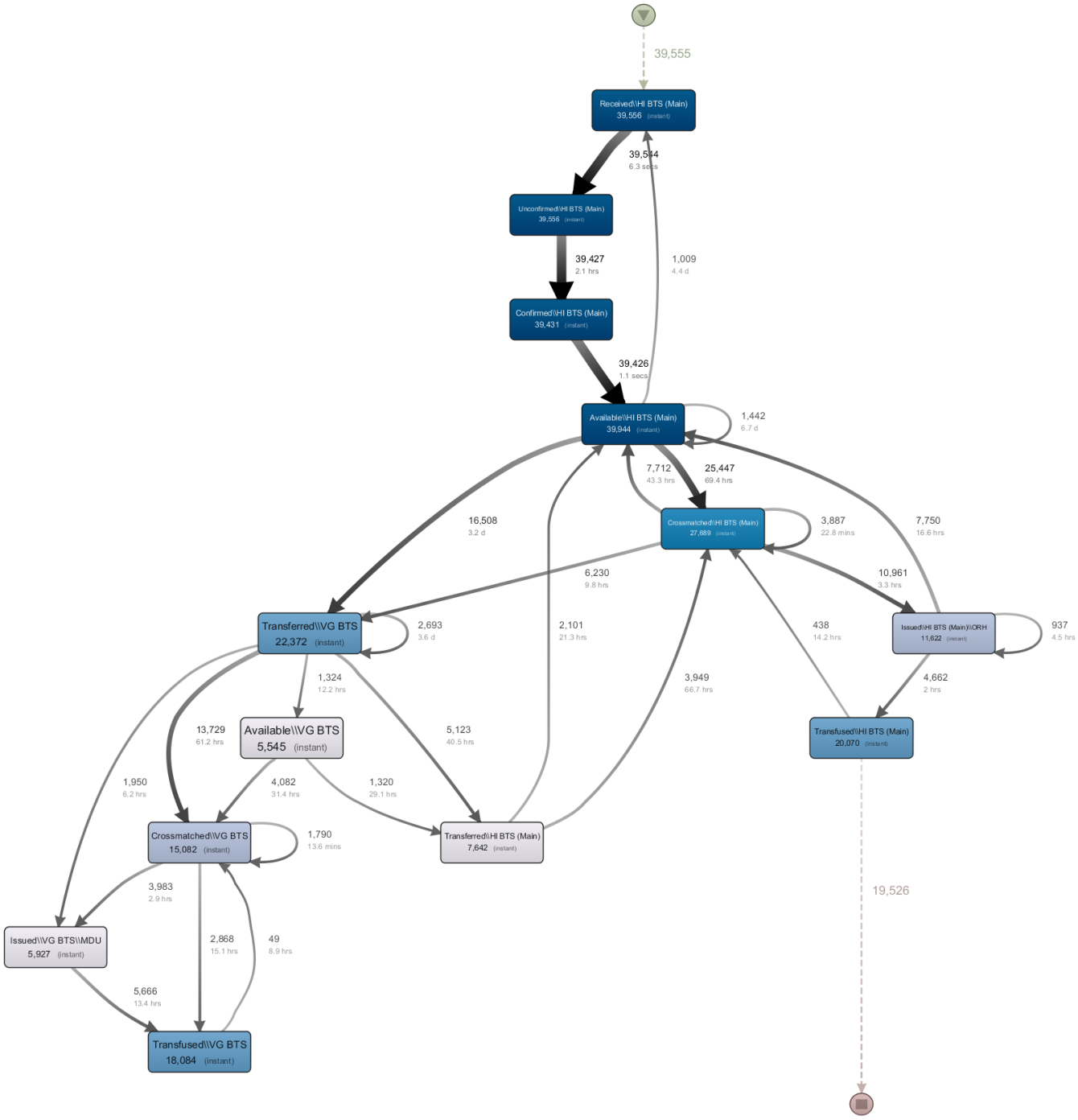


Figure 23: Third iteration of data exploration with process map from Disco with additional DISPENSE_LOCN mapped

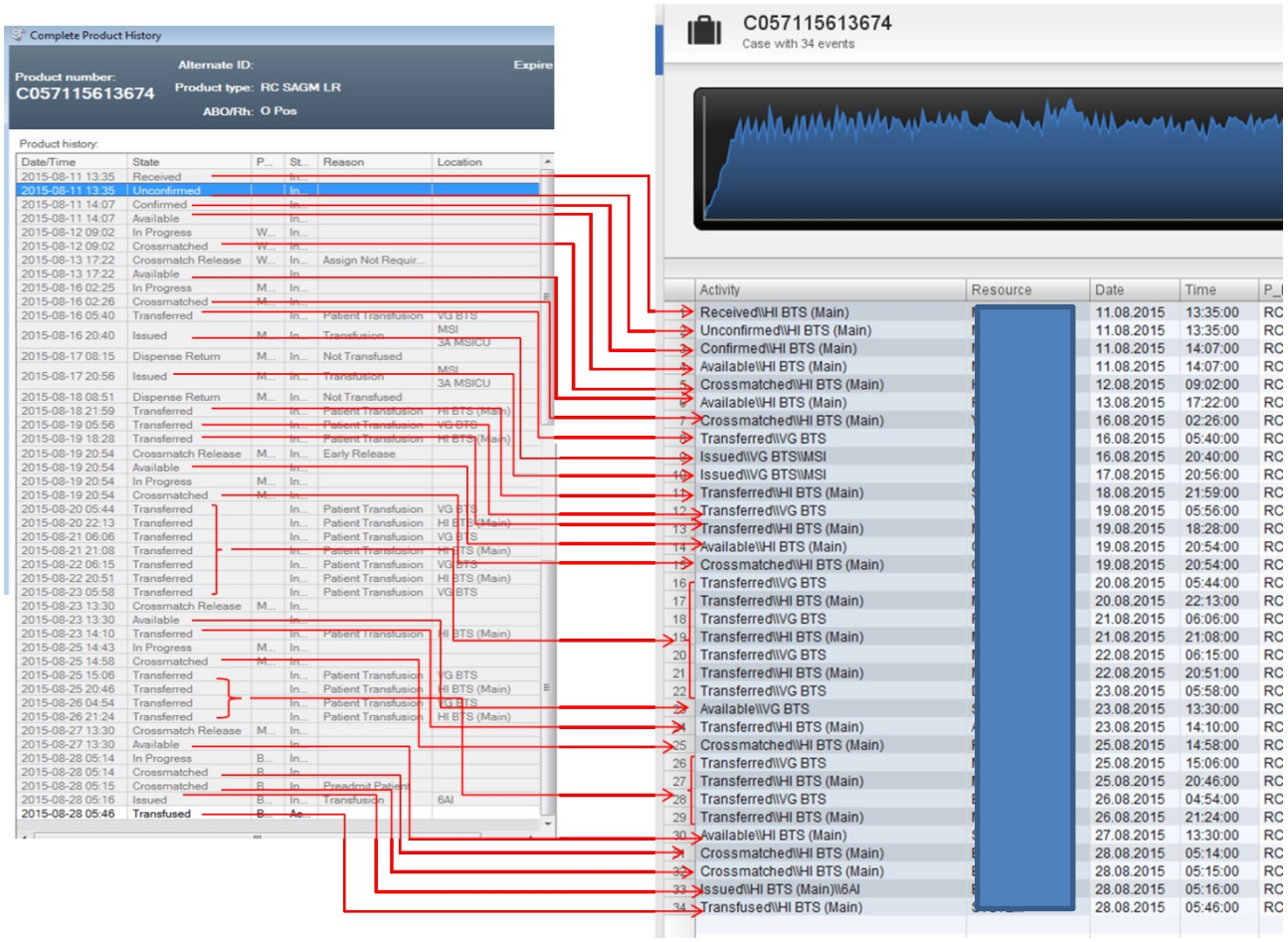


Figure 24: Example of third iteration/exploration cycle of unit C057115613674 with comparison between Complete Product History and Disco 'Cases' tab outputs(red lines indicate concordance).

The fourth iteration involved further refinement of the activity with the addition of CROSSMA_RELEASE_REASON, DISPENSE_RETURN_REASON, and REINSTATE_REASON fields (Table 26), which pertained to reasons which units could be returned or crossmatched. By examining the process maps, it appears that there was separation of the crossmatch by various return and dispense reasons (Figure 25); however, there were artifacts that were introduced when the Complete Product History was attempted to be matched with the output from the cases tab (blue arrows indicate artifacts, Figure 26). The artifacts related to the timestamps whereby dispense return reasons did not have any timestamps associated with them, even if they had a 'reason', whereas Crossmatch Release reason timestamps were offset onto the next

activity. Overall, this highlighted an issue that was not discovered on the data validation step prior to importation into Disco.

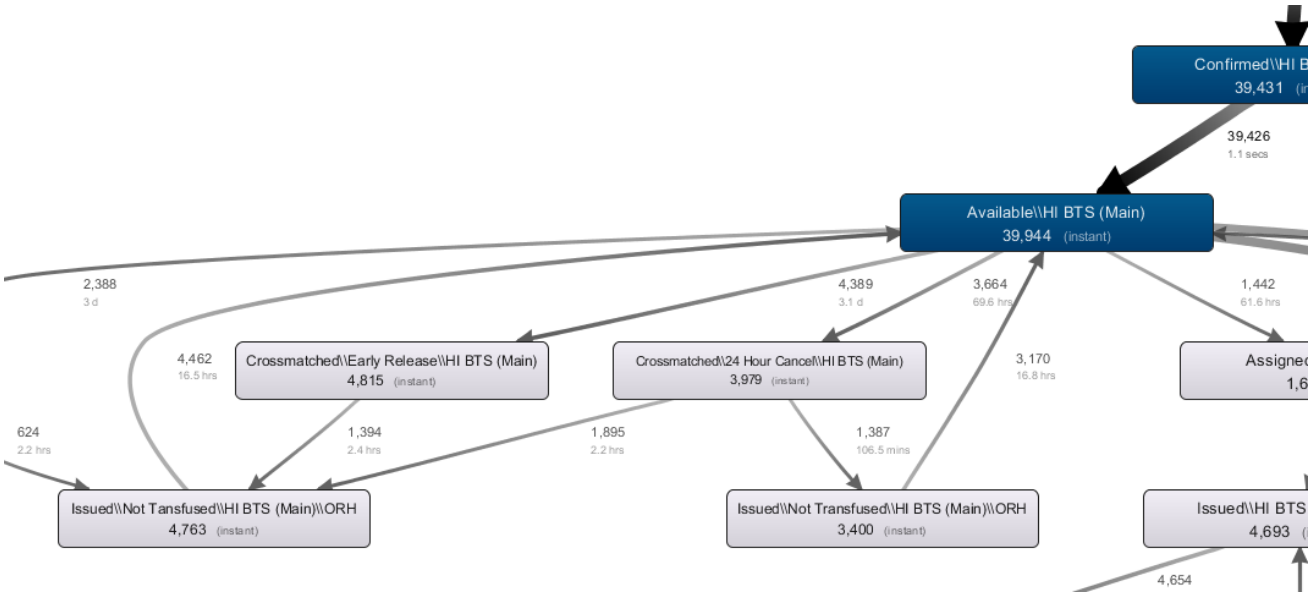


Figure 25: Fourth iteration of data exploration cycle demonstrating separation of crossmatch return reasons

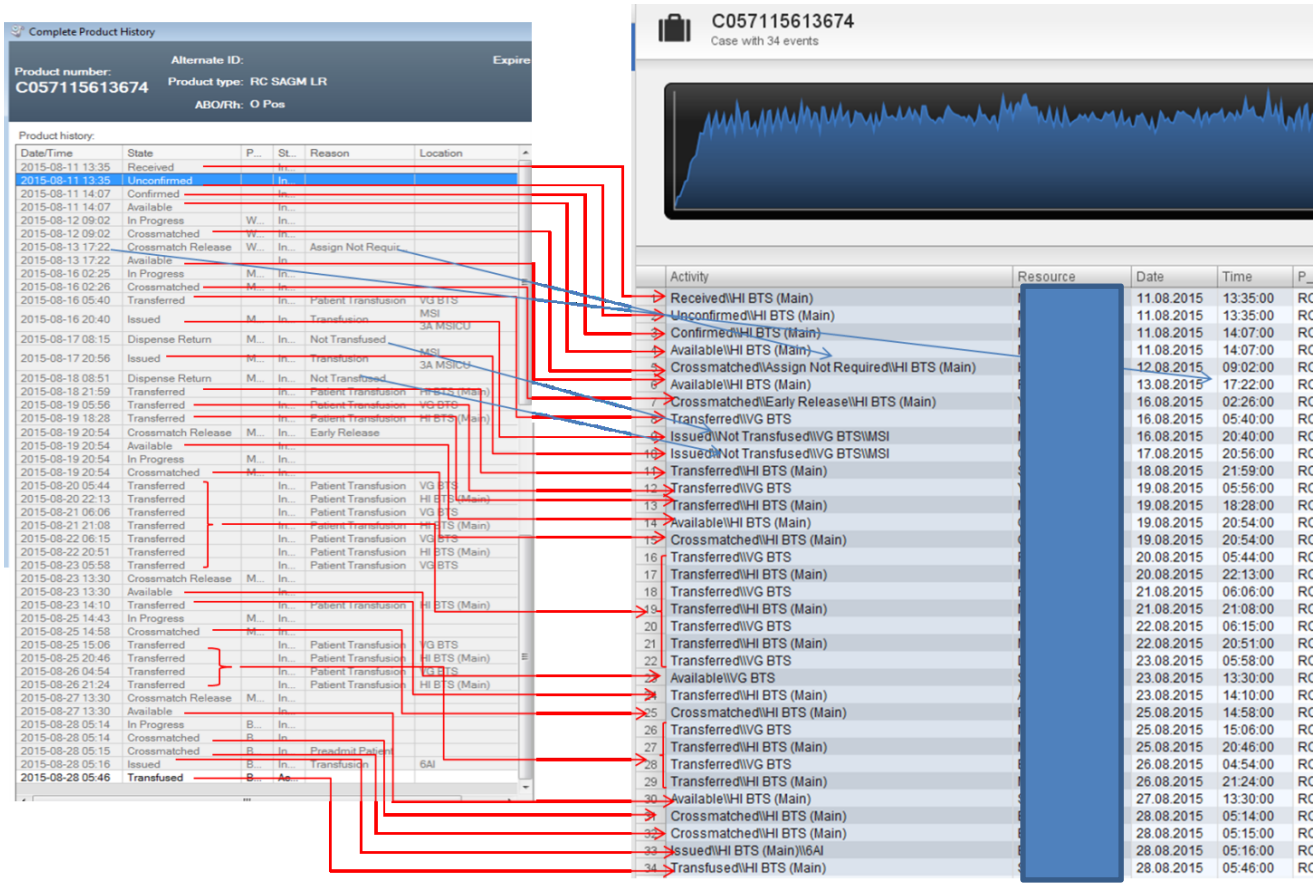


Figure 26: Example of fourth iteration/exploration cycle of unit C057115613674 with comparison between Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance).

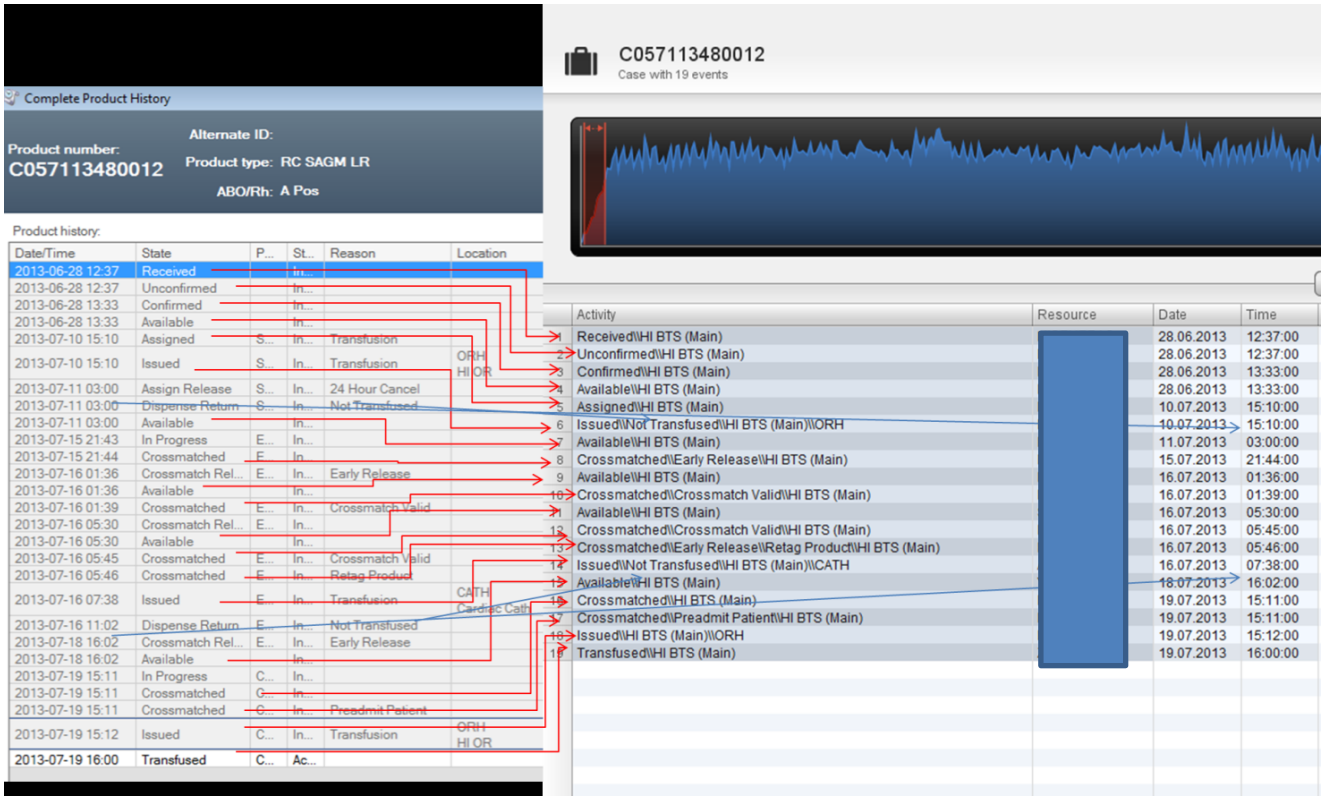


Figure 27: A second example of fourth iteration/exploration cycle of unit C057113480012 with comparison between Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance).

The fifth iteration involved further refinement of the activity with the removal of the DISPENSE_RETURN_REASON field (Table 26), with the rationale that a crossmatch activity (verb) should have been matched by crossmatch associated reasons (object). The dispense reasons (object) should not have been mixed with the crossmatched verb since there were no 'returned dispense' verbs. The output process map, shown in Figure 28 demonstrated separated crossmatch return events; however, the timestamps were still shifted as noted in Figure 28 and Figure 29 and were noted as artifacts. There was a reduction in timestamp and status artifacts.

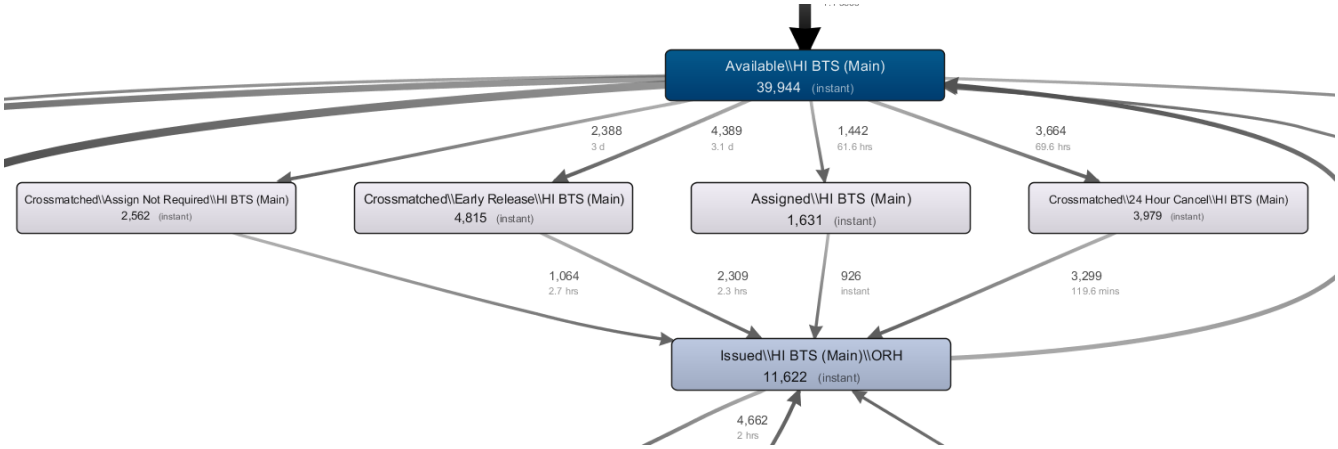


Figure 28: Fifth iteration of data exploration cycle demonstrating separation of crossmatch return reasons

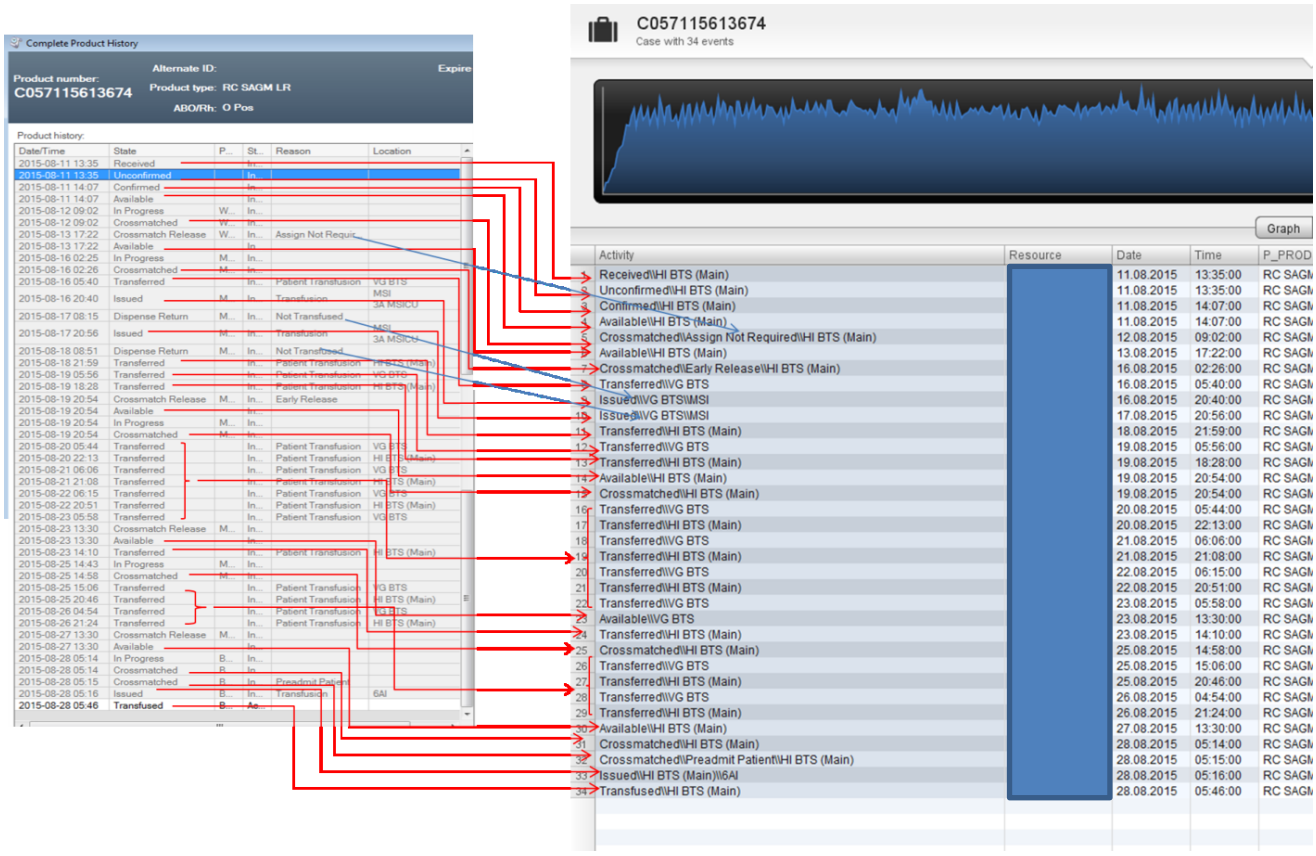


Figure 29: Example of fifth iteration/exploration cycle of unit C057115613674 with comparison between Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance).

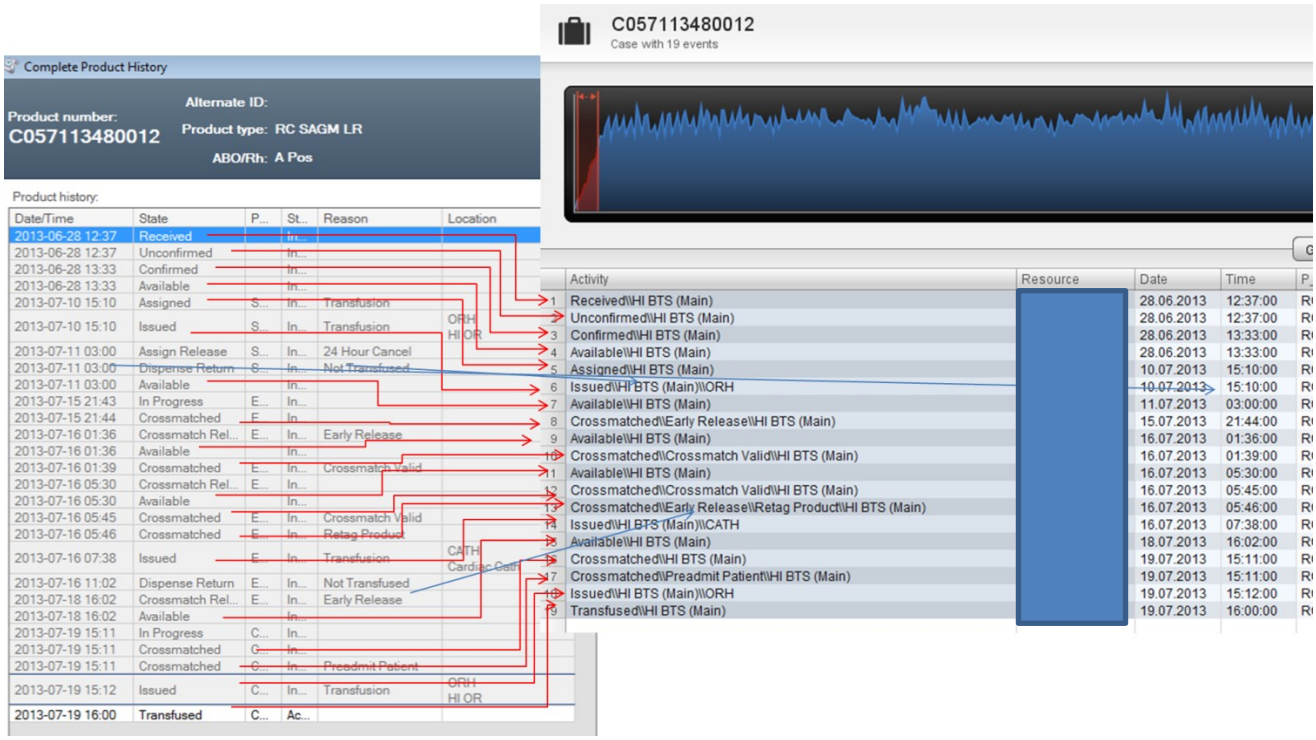


Figure 30: A second example of fifth iteration/exploration cycle of unit C057113480012 with comparison between Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance).

After the 5th iteration, there was an appropriate balance of pathway resolution with explainable artifacts generated in the process mining-generated process maps and data. Overall, as this was the best iteration from a specificity and artifact generation standpoint, this acted as a starting-point for specific red cell and Hemosafe™-related data exploration as data was further optimised for use in the respective focus groups. Also, the artifacts related to the crossmatching returns, and the need for high-resolution accurate data of the crossmatch/crossmatch return process will not be a requirement of any of the focus groups. Nonetheless, this timestamp issue would need to be recognised as a limitation of the dataset.

Table 22: Data Exploration Iterations to Determine Optimal Activity Mapping Parameters for Optimal Concordance between Case Tab, Complete Product History, and Visual Process Map

Iteration of Data Exploration	Parameters	Rationale	Outcome of Case Tab vs. Complete Product History	Comments
1st	Default mapping of core log UNIT_NBR, STATUS, STATUS_DT_TM, and resource	Simplest process map, low resolution	1:1 match when considering case tab activities; previously noted activities missing	Lack of specificity of activities (e.g. 'availability' at the HI BTS is the same as 'availability' at VG BTS) and location
2nd	Additional mapping of STATUS_LOC to 'activity'	Allowed different activities to be separated according to site.	1:1 match when considering case tab activities; previously noted activities missing.	Lack of specificity of activities pertaining to transfusion and crossmatch[e.g. Issued\\HI BTS (Main)].
3rd	Additional mapping of DISPENSE_LOCN to 'activity'	Allowed different activities to be separated according to site, especially at transfusion	1:1 match when considering case tab activities; previously noted activities missing. The location codes for transfusion were correct.	Lack of specificity of activities pertaining to crossmatch[e.g. Issued\\HI BTS (Main)\\6AI].
4th	Additional mapping of CROSSMA_RELEASE_REASON and DISPENSE_RETURN_REASON and	Allowed for the separation of crossmatch and return reasons.	Incorrectly inserted Reason code into the next state (blue arrows) with next timestamp if unit	Despite having increased specificity to separate activities from each other, the insertion of the Reason code occurred at the next or previous timestamp, or with

Iteration of Data Exploration	Parameters	Rationale	Outcome of Case Tab vs. Complete Product History	Comments
	REINSTATE_REASON to 'activity'		was returned for various reasons. Correctly gives crossmatch reason if in the affirmative and not returned.	no timestamp at all (in the case of Dispense Returns), thus introducing an artifact. This only seems to happen for the negative for both crossmatch returns and dispense returns.
5th	Removal of DISPENSE_RETURN_REASON from 'activity'	Removed an improper verb-object reference. Reinstate and crossmatch return reasons (objects) refer to crossmatch (verb). The dispense reasons (object) should not be mixed in there, as there is no verb (dispense returned)	Still persistent errors with timestamps; however, dispense return reason no longer there and process maps are cleaner than 4th iteration. Crossmatch valid reason has correct timestamp.	Problem for crossmatch release is that it appears to happen at the beginning of the crossmatch timeperiod, rather than when it is released at the end. This would need to be recognized as a limitation of the data set. No other data artifacts have been introduced.

3.4.4 Red Cell Wastage Specific Data Exploration

The purpose of this cycle of data exploration was to ensure that the process maps effectively represented all of the data elements for the purposes of the RBC wastage focus group, and ensured that the correct time period was used. The first phase of exploring red cell wastage involved mapping the DISPOSE_REASON to the activity, with the rationale that it could act as an 'object' to specify the verb of 'disposed'. As noted in Figure 31, the reason for disposal was integrated into the visual process map. After mapping, as noted in Figure 32 and Figure 33, there were no significant qualitative differences found between the fifth iteration and the red cell wastage iteration (6^{th} iteration), in that there were no artifacts introduced. The Assign Release state on the product history was discovered to be new, and was likely a state related to dispense return, which was not captured by the dataset (Figure 32).

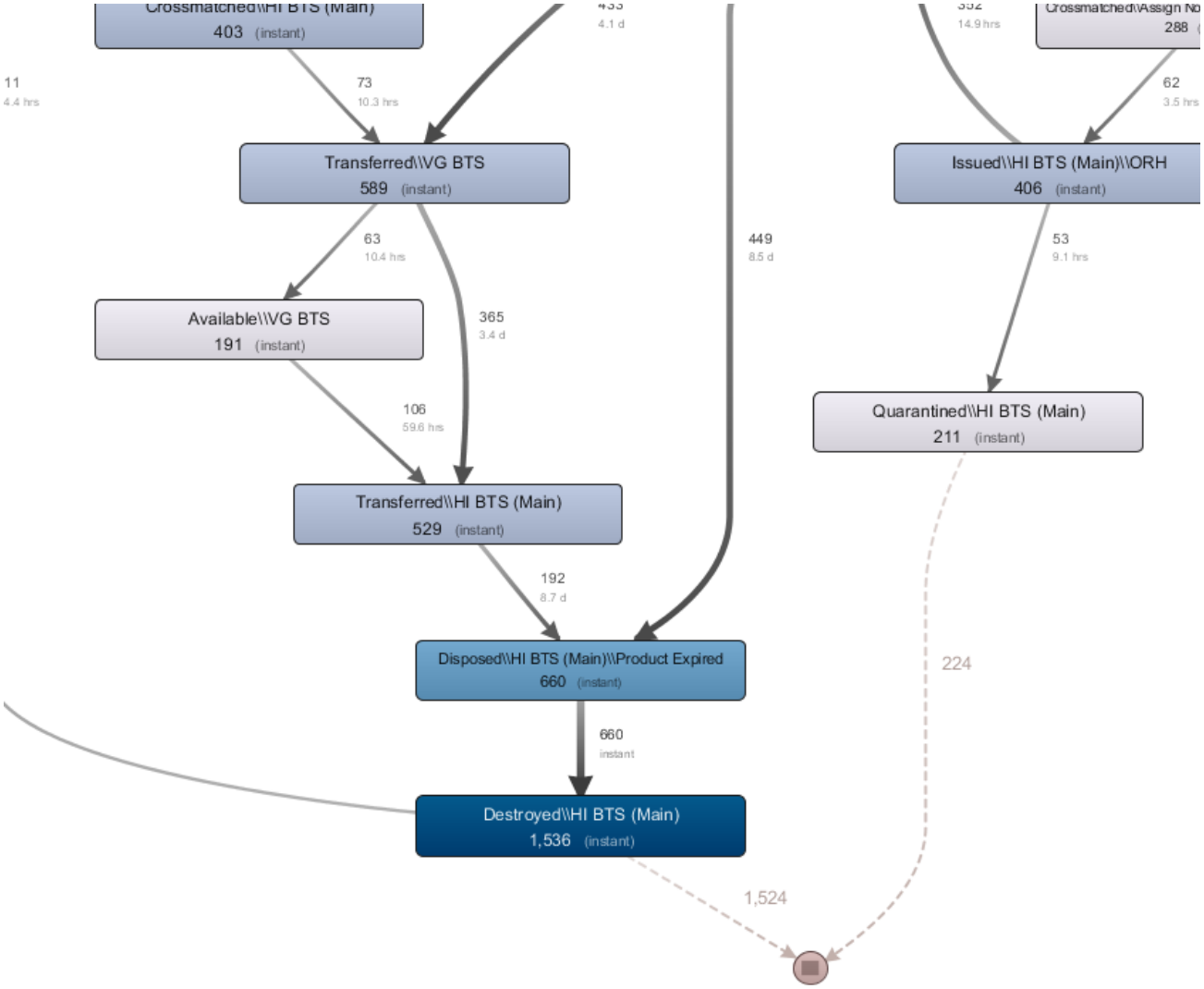


Figure 31: Process map from DISPOSE_REASON mapped to activity field in data exploration cycle

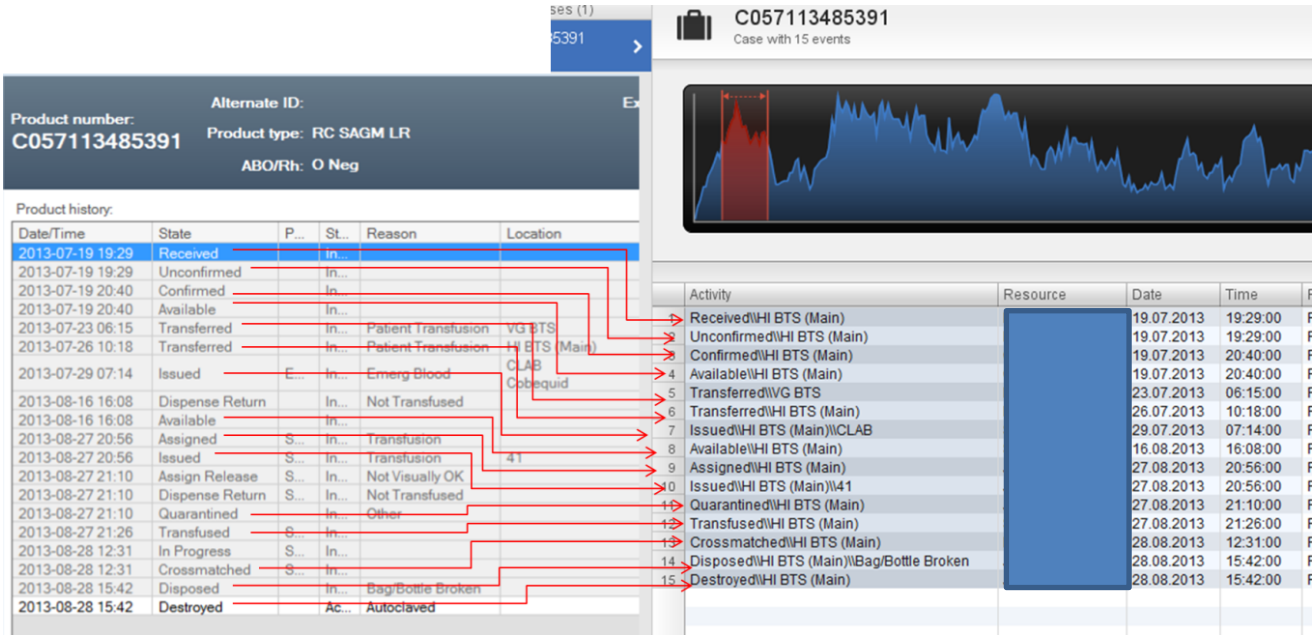


Figure 32: First example of red cell wastage iteration/exploration cycle of unit C057113485391 with DISPOSE_REASON mapped to activity field and comparing Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any).

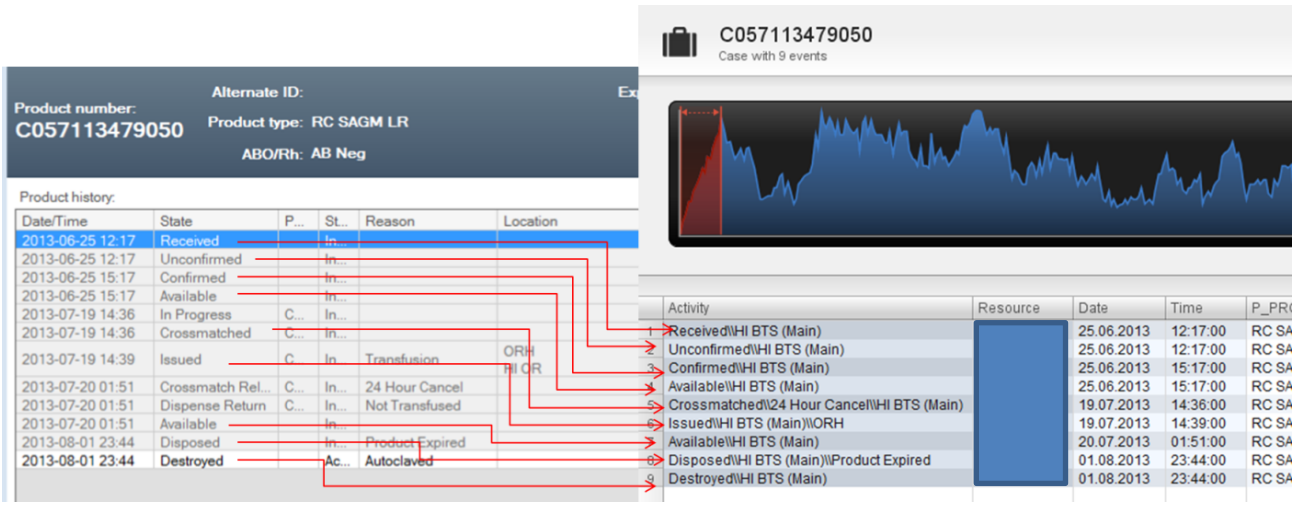


Figure 33: Second example of red cell wastage iteration/exploration cycle of unit C057113479050 with DISPOSE_REASON mapped to activity field and comparing Complete Product History and Disco 'Cases' tab outputs (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any)

The second phase of exploring red cell wastage was done by leaving the DISPOSE_REASON to be mapped to the 'other' category, thus using it only as a filtering variable, rather than being displayed as an activity on the process map. By comparing Figure 31 and Figure 34, the lack of displaying a DISPOSE_REASON would have introduced a second step in the in depth analysis phase of the process mapping focus group, as the results will have to be refiltered and redisplayed. Otherwise, there were no artifacts that were introduced as part of this iteration as noted.

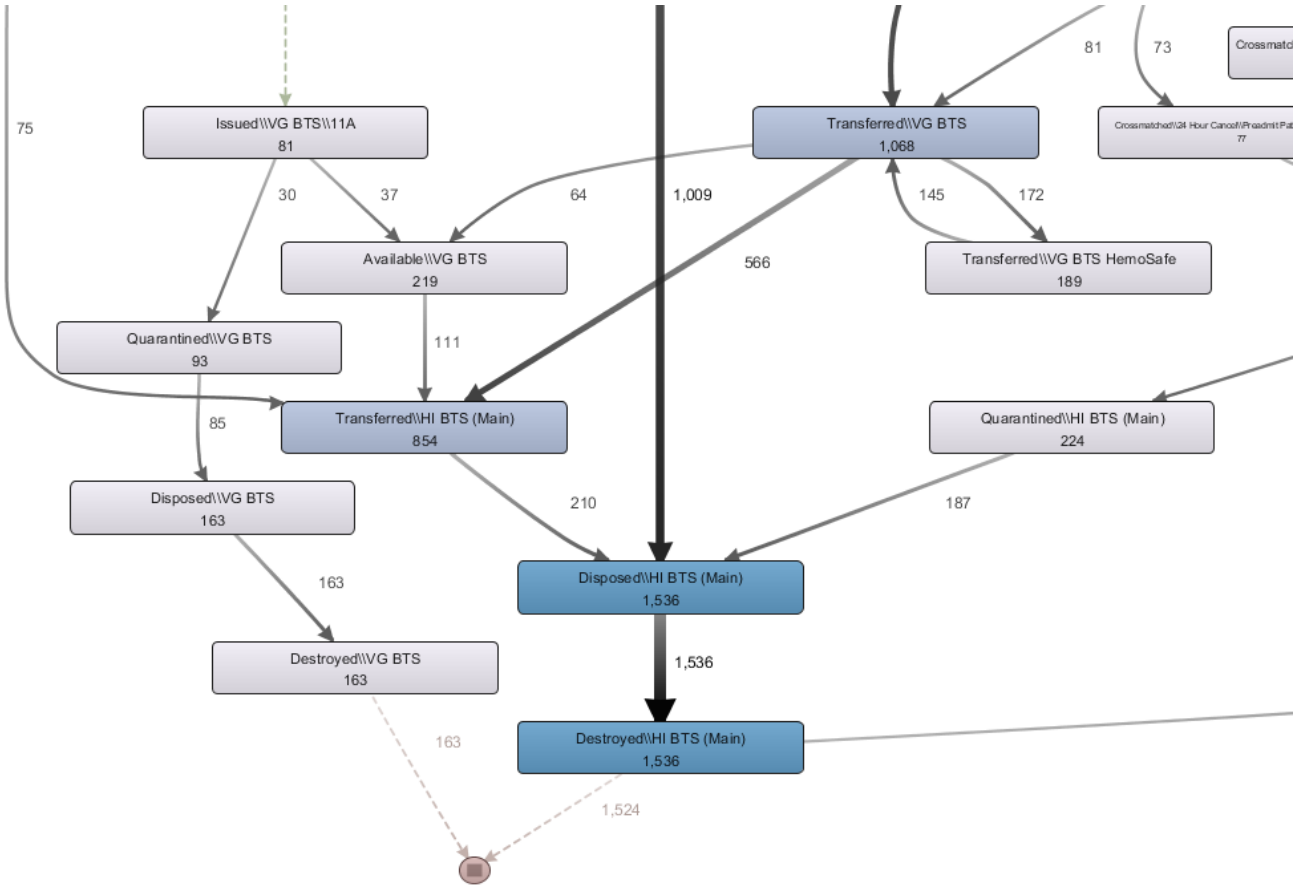


Figure 34: Process map from RBC wastage mapped to 'other' field in data exploration cycle

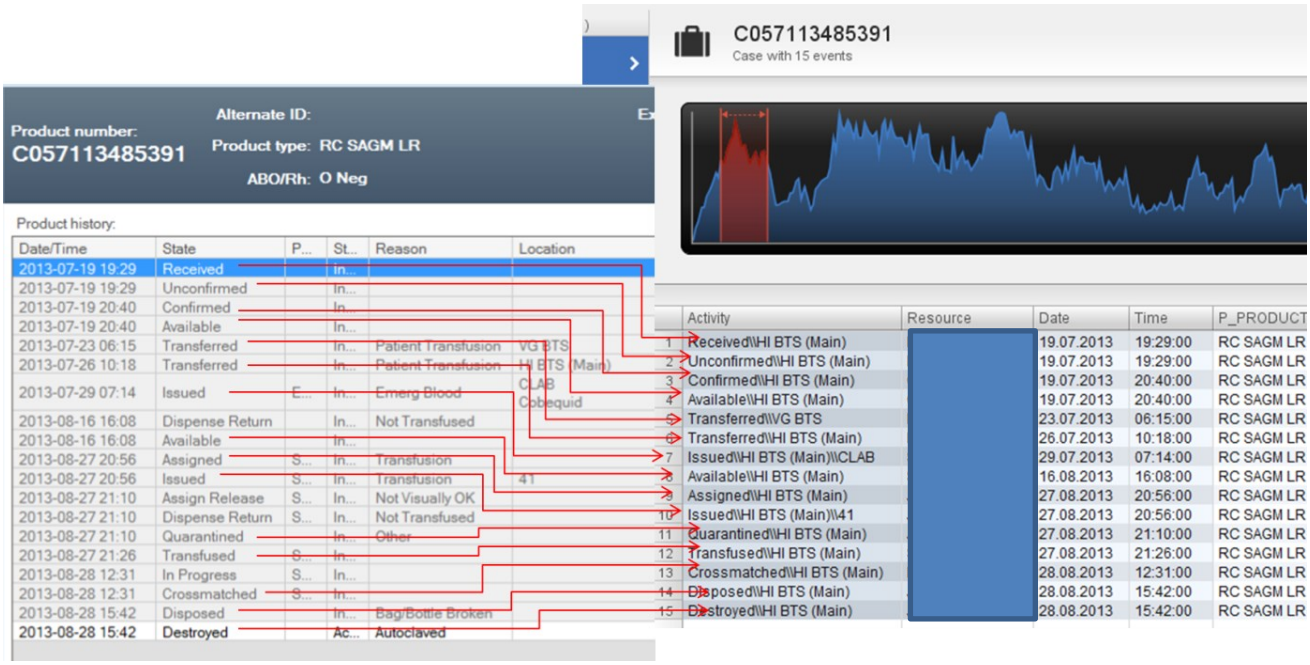


Figure 35: First example comparing Complete Product History and Disco 'Cases' tab outputs for unit C057113485391 when DISPOSE_REASON mapped to 'other' field (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any)

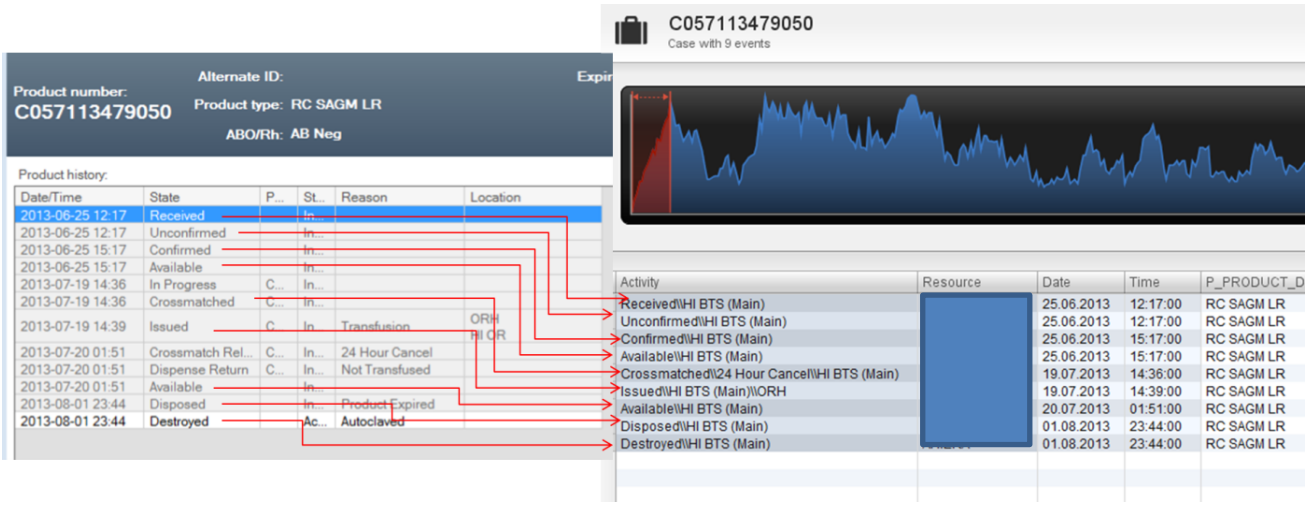


Figure 36: Second example comparing Complete Product History and Disco 'Cases' tab outputs for unit C057113479050 when DISPOSE_REASON mapped to 'other' field (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any)

Table 23: Data Exploration Iterations to Determine Optimal Activity Mapping Parameters for Optimal Concordance between Case Tab, Complete Product History, and Visual Process Map for the RBC Wastage Focus Group

Dispose reason mapping state	Rationale	Outcome of Case Tab vs. Complete Product History	Comments
Mapped as an activity	No need to filter the result; the disposal reason shows up on the process map. Disposed (verb) corresponds to disposal reason (object)	1:1 match when considering case tab activities; previously noted activities missing and previously noted issues in the first five iterations persisting.	No new issues with data inconsistency. Activity nodes within the process map directly represent the dispose reason, so there is no need to filter.
Mapped as 'other' variable	Cleaner process map look, allow values to be filtered instead.	1:1 match when considering case tab activities; previously noted activities missing and previously noted issues in the first five iterations persisting.	Lack of specificity of process map visually. Extra step of filtering needs to occur, so need to go back and forth between the Map and Filter functions on Disco.

The mapping of DISPOSE_REASON to the 'activity' field represented the starting point for the Hemosafe™ fridge specific data exploration section below.

The second objective of data exploration was to understand the scope of the time period but also generate some statistical data related to RBC wastage. During the entire time period, there were 29019 unique units in the dataset, with the largest proportion of final dispositions being transfusions, representing 95.68% of total final dispositions, and the 1255 (4.32%) remaining dispositions consisted of unit transfers to transfusion services outside of the Central Zone or to wastage. Of the 1255 disposed units, there were 967 (77.1%) units disposed due to wastage. The largest wastage reason was product expiry at 55.43% of total wastage reasons. The largest non-expiry wastage reasons were "Exceeds time out of fridge" and "Security Tag Broken or Absent" at 8.48% and 6.00 % respectively. The AB neg, AB pos, B neg, and O negative units accounted for 47.3% of expired product and qualitatively A pos, O pos and B pos units accounted for most of the non-expiry wastage reasons. Table 24 represented the absolute frequency of wastage reasons with heatmapping using conditional formatting (red-yellow-green color scale, with red being the maximum and green being the minimum).

Table 24: Absolute Frequency of Wastage Reasons for RBC Units

Count of ABORH	Column Labels								
Row Labels	**A Neg**	**A Pos**	**AB Neg**	**AB Pos**	**B Neg**	**B Pos**	**O Neg**	**O Pos**	**Grand Total**
No Segments	1								1
Label Problem		1							1
Patient Deceased							1		1
In Lab Temp Unacceptable	1								1
Washed not used		2							2
Improper packing of blood crates							1	2	3
Return to CBS Modify							3	1	4
Storage Equipment Failure		4				1	4		9
Bag/Bottle Broken	1	2					2	7	12
Failed Visual Inspection	1	3		2		2		5	13
Positive DAT	1	5		1	1	1	2	6	17
In Lab Temp/Visual		5	1	2		3	7	3	21
CBS Request	1	8					3	10	22
No Secure Seal	4	3			3		12	3	25
Bag Spiked	1	10		1	1	1	8	6	28
Received from Outside Hospital	2	11					5	14	32
Return Temp Unacceptable	1	14			1	14	4	10	44
Return Temp/Visual	4	16				4	6	25	55
Security Tag Broken or Absent	11	10	1		3		7	26	58
Exceeds time out of fridge	3	17				2	29	31	82
Product Expired	55	15	116	128	102	9	111		536
Grand Total	**87**	**126**	**118**	**134**	**111**	**37**	**205**	**149**	**967**

Since wastage was so infrequent, it was important to balance having enough events to drive creation of the process map and also have enough timeliness so that the participants could provide insight into those process maps. As noted in Table 25, the heatmap (red-yellow-green color scale, with red being the maximum and green being the minimum) demonstrated that expiry in 2014 was high, and that apart from that, wastage reasons were smaller in 2013, likely from the truncated time period. Therefore, it was reasonable to obtain data from the time period of 2014-2015, as the addition of the 2013 period may have introduced recall bias (i.e. focus group participants not remembering details) with little additional gain in the number of cases used to derive the process map. The period of September 1, 2013 to September 30, 2015, was arbitrarily chosen, as it represented the most recent two years' worth of data in the dataset.

Table 25: Absolute Frequency of Wastage on a Yearly Basis of Collected Data.

Count of ABORH	Column Labels			
Row Labels	**2013**	**2014**	**2015**	**Grand Total**
No Segments	1			1
Improper packing of blood crates			3	3
Patient Deceased			1	1
Washed not used			2	2
Label Problem		1		1
In Lab Temp Unacceptable		1		1
Return to CBS Modify		2	2	4
Bag/Bottle Broken	1	3	8	12
Storage Equipment Failure		4	5	9
Failed Visual Inspection		6	7	13
Positive DAT	5	9	3	17
CBS Request	6	11	5	22
No Secure Seal	14	11		25
In Lab Temp/Visual	10	11		21
Bag Spiked	6	15	7	28
Security Tag Broken or Absent	11	17	30	58
Received from Outside Hospital	2	18	12	32
Return Temp Unacceptable		19	25	44
Exceeds time out of fridge		35	47	82
Return Temp/Visual	20	35		55
Product Expired	93	311	132	536
Grand Total	**169**	**509**	**289**	**967**

3.4.5 Hemosafe™ Fridge Specific Data Exploration

The purpose of this cycle of data exploration was to ensure that the process maps are effectively represented all of the data elements for the purposes of the Hemosafe™ fridge focus group. The rationale for doing the RBC wastage data exploration first was that the Hemosafe™ fridge was already inadvertently integrated into the data exploration cycle, as it was an inventory location found in the STATUS_LOC field and has already been introduced in cycle #1. This final exploration cycle was to confirm that no artifacts have been introduced into the Hemosafe location from DISPOSE_REASON mapping, and that the Hemosafe location could be seen in the process maps and the Case tab. As noted in Figure 37, the Hemosafe location was easily identified with inflows and outflows, and in Figure 38 and Figure 39, there were no artifacts introduced, as expected, by comparing the Disco 'cases' tab with Complete Product History.

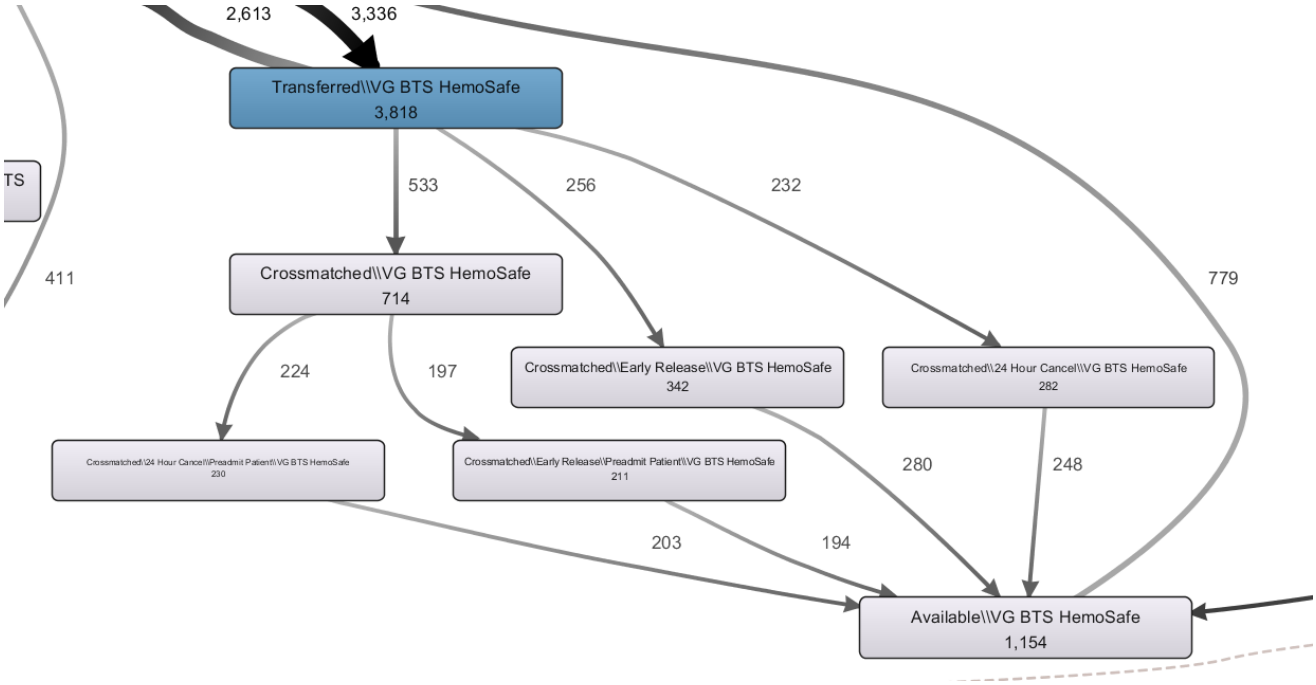


Figure 37: Hemosafe fridge inventory location process map for purposes of data exploration

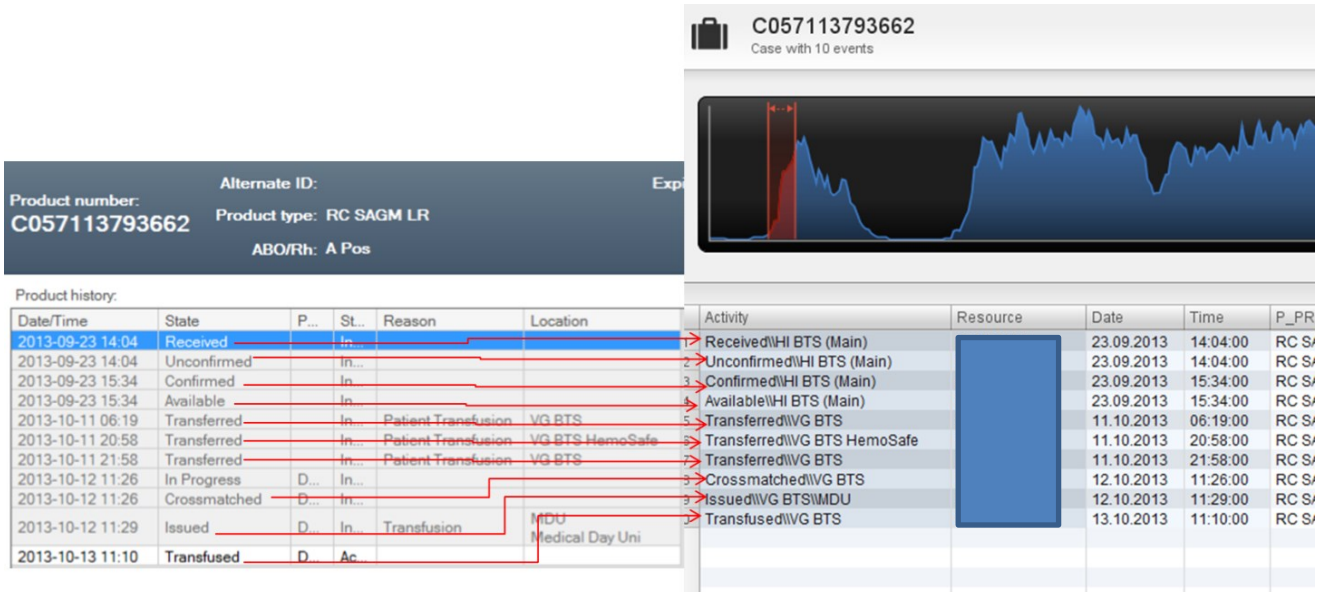


Figure 38: First example comparing Complete Product History and Disco 'Cases' tab outputs for unit C057113793662 for Hemosafe™ data exploration cycle (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any)

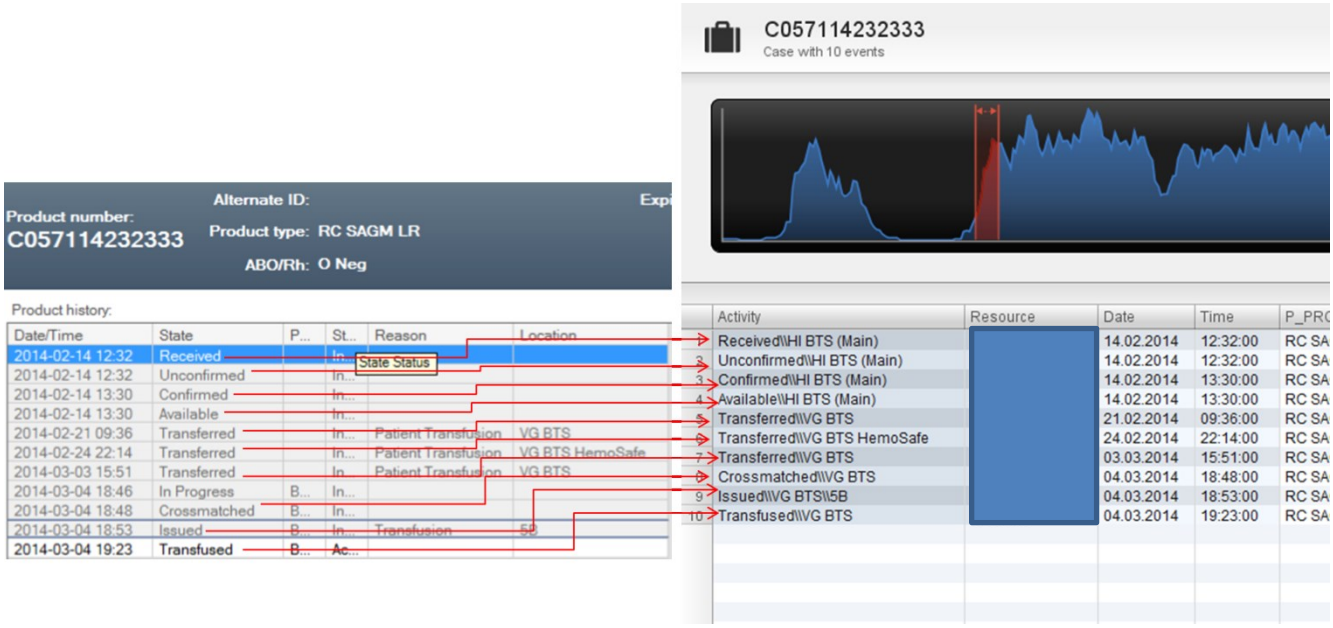


Figure 39: Second example comparing Complete Product History and Disco 'Cases' tab outputs for unit C057114232333 for Hemosafe™ data exploration cycle (red arrows indicate concordance; blue arrows indicate discordance found outside of previously noted discordances, if any)

The overall objective of the data exploration step was accomplished, and it was determined that the data which had been imported into Disco needed increased verb-object specificity to ensure adequate detail in the generated process maps, and that there were predictable artifacts which were generated. The overall verb-object iteration steps were noted on Table 26.

Table 26: Iterative Mapping of Master CSV Data Fields to Disco Data Importation Fields during Data Exploration Step using the TRUISM Framework.

Field Name on final query of table	Disco mapped fields, 1st Iteration of Exploration	Disco mapped fields, 2nd Iteration of Exploration	Disco mapped fields, 3rd Iteration of Exploration	Disco mapped fields, 4th Iteration of Exploration	Disco mapped fields, 5th Iteration of Exploration	Disco mapped fields, RBC Wastage (6th iteration)	Disco mapped fields, Hemosafe Usage (7th iteration)
P_PRODUCT_DISP	Other	Other	Other	Other	Other	Other	Other
UNIT_NBR	Case ID	Case ID	Case ID	Case ID	Case ID	Case ID	Case ID
CHK_DGT	Other	Other	Other	Other	Other	Other	Other
ABORH	Other	Other	Other	Other	Other	Other	Other
SUPPLIER	Other	Other	Other	Other	Other	Other	Other
ORIG_VIS_INSP	Other	Other	Other	Other	Other	Other	Other
SPECIAL_TREATMENT	Other	Other	Other	Other	Other	Other	Other
STATUS	Activity	Activity	Activity	Activity	Activity	Activity	Activity
DURING_DOWNTIME	Other	Other	Other	Other	Other	Other	Other
CROSSMA_RELEASE_REASON	Other	Other	Other	Activity	Activity	Activity	Activity
DISPENSE_RETURN_REASON	Other	Other	Other	Activity	Other	Other	Other
REINSTATE_REASON	Other	Other	Other	Activity	Activity	Activity	Activity
STATUS_LOC	Other	Activity	Activity	Activity	Activity	Activity	Activity
WARD_LOCN	Other	Other	Other	Other	Other	Other	Other
DISPENSE_LOCN	Other	Other	Activity	Other	Other	Other	Other
DISPOSE_REASON	Other	Other	Other	Other	Other	Activity	Activity
STATUS_DT_TM	Timestamp	Timestamp	Timestamp	Timestamp	Timestamp	Timestamp	Timestamp
EXPIRY_DT	Other	Other	Other	Other	Other	Other	Other
USERNAME	Resource	Resource	Resource	Resource	Resource	Resource	Resource

A secondary objective of the Hemosafe™ data exploration was to define the scope of the time period used. Hemosafe™ use appears regular, but very small, representing 0.59% (163/27824) units over the time period. The usage pattern appeared very regular over time even though it was small (Table 27). As the Hemosafe™ implementation was February 1, 2014, the time period of March 1,2014 – September 30, 2015 (after implementation) made intuitive sense, and if

comparison was required, then the corresponding time period prior to implementation could be chosen as September 1, 2013-February 1, 2014. This would allow for the most recent Hemosafe™ usage patterns to be fresh in the participants' memory.

Table 27: Inventory location where a unit of RBCs originates from just prior to transfusion, with heatmapped cells (red-yellow-green color scale, with red being the maximum and green being the minimum)

Count of ABORH	Column																											
	2013					2014													2015									Grand Total
Row Labels	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	
DG BTS	4	90	88	103	118	127	96	136	155	139	160	159	139	110	103	97	113	99	118	128	135	97	94	103	107	20	2838
HC BTS		11	31	21	20	14	26	25	43	45	26	25	25	38	34	33	29	17	27	21	35	20	33	21	15	8	643
HI BTS (Main)	13	288	449	555	549	537	487	429	512	443	448	467	437	488	576	547	525	574	526	583	563	536	531	638	567	217	12485
VG BTS	2	333	460	448	413	481	497	386	474	633	558	541	453	448	510	537	587	489	452	583	402	535	461	439	437	136	11695
VG BTS HemoSafe			1				1	14	12	13	21	9	4	11	4	5	11	22	7	6		3	4	11	3	1	163
Grand Total	**19**	**722**	**1029**	**1127**	**1100**	**1159**	**1107**	**990**	**1196**	**1273**	**1213**	**1201**	**1058**	**1095**	**1227**	**1219**	**1265**	**1201**	**1130**	**1321**	**1135**	**1191**	**1123**	**1212**	**1129**	**382**	**27824**

3.5 Evaluation Framework

The goal of the evaluation framework was to determine if these process mining-generated maps could result in 'guidelines for process improvement', the final result as outlined by the PMMF (Figure 5, Figure 40), a guideline on how to conduct a case study in the process mining environment. This framework was applied to two case studies, one which addressed RBC unit wastage at our institution which is a long-standing unsolved question in the literature regarding avoidable RBC unit wastage [2]. The second case study was also not documented in the literature, and addressed the role of how an automated RBC product dispensing fridge was used in the transfusion service. This case study featured the manual construction of the process map prior to commencing the focus group to examine the differences between a manually constructed process map and one constructed from process mining. Although both case studies were selected due to their lack of resolution in the transfusion inventory literature, these two case studies were also timely topics that needed to be addressed at our blood transfusion service and were not able to be addressed at all due to a significant lack of human resources. Process mining-derived process maps pertaining to these case studies were presented to focus groups consisting of SMEs in transfusion medicine and moderated by a SME (the researcher, myself) in transfusion medicine and process mining, the former element previously described to

be a critical element for success by De Weerdt *et al.* [48]. The focus groups used a Concurrent Think Aloud (CTA) technique adopted from usability testing principles (usability.gov) whereby participants were encouraged to think out loud as they interacted with the process maps, and provided a prospective ”running stream of consciousness” as they reflected on the meaning of the process map (usability.gov) (Table 28). The CTA was chosen as this emulated the communication that occurs during the CZBTS Quality Assurance Committee.

Table 28: Usability Testing Techniques and Their Attributes Adapted from Usability.gov

Technique name	**Attributes and comments**
Concurrent Think Aloud (CTA)	Allows for real-time feedback, and the understanding of participants thoughts while they work through an issue. This can increase the time to perform a task and affect accuracy.
Retrospective Think Aloud (RTA)	Poor data collection, as participants need to delay comments until the end of the session. Does not interfere with task performance time, but difficulty remembering can adversely affect quality of data.
Concurrent probing (CP)	Understands participant thought while they work through a problem, but can interfere with natural thought processes.
Retrospective probing (RP)	Does not interfere with task performance, but participants can have difficulty in remembering, resulting in poor quality data.

The goal of the moderator was to ask open-ended questions and operate the process mining software so that the software interface was not a barrier to the participants while they self-interpreted the maps. This process was designed to mimic the feedback that was typically informally gathered during quality assurance committee meetings when KPI-related data is provided to SMEs for review as a group. For the purposes of this research, the feedback of the focus groups was documented, codified, and thematically analysed using a qualitative evaluation methodology. The main question of the evaluation framework was whether or not any process-improvement related guidelines arose as a result of the two focus group exercises.

3.5.1 The Process Mining Methodology Framework in the Context of the TRUISM Framework

The study will use the Process Mining Methodology Framework as outlined in Figure 40 below [48] and the validated process maps and data from the TRUISM Framework (Figure 15) will be used as inputs. The PMMF would have specifically taken the Case Data perspective (red arrow, Figure 40), and thus its pathway, as the intent of the research was not to examine either control-flow or organisational perspectives. The objective of the research was to demonstrate if any guidelines for process improvement are derived, and qualitatively evaluated the context in which the discussion of those guidelines occurred.

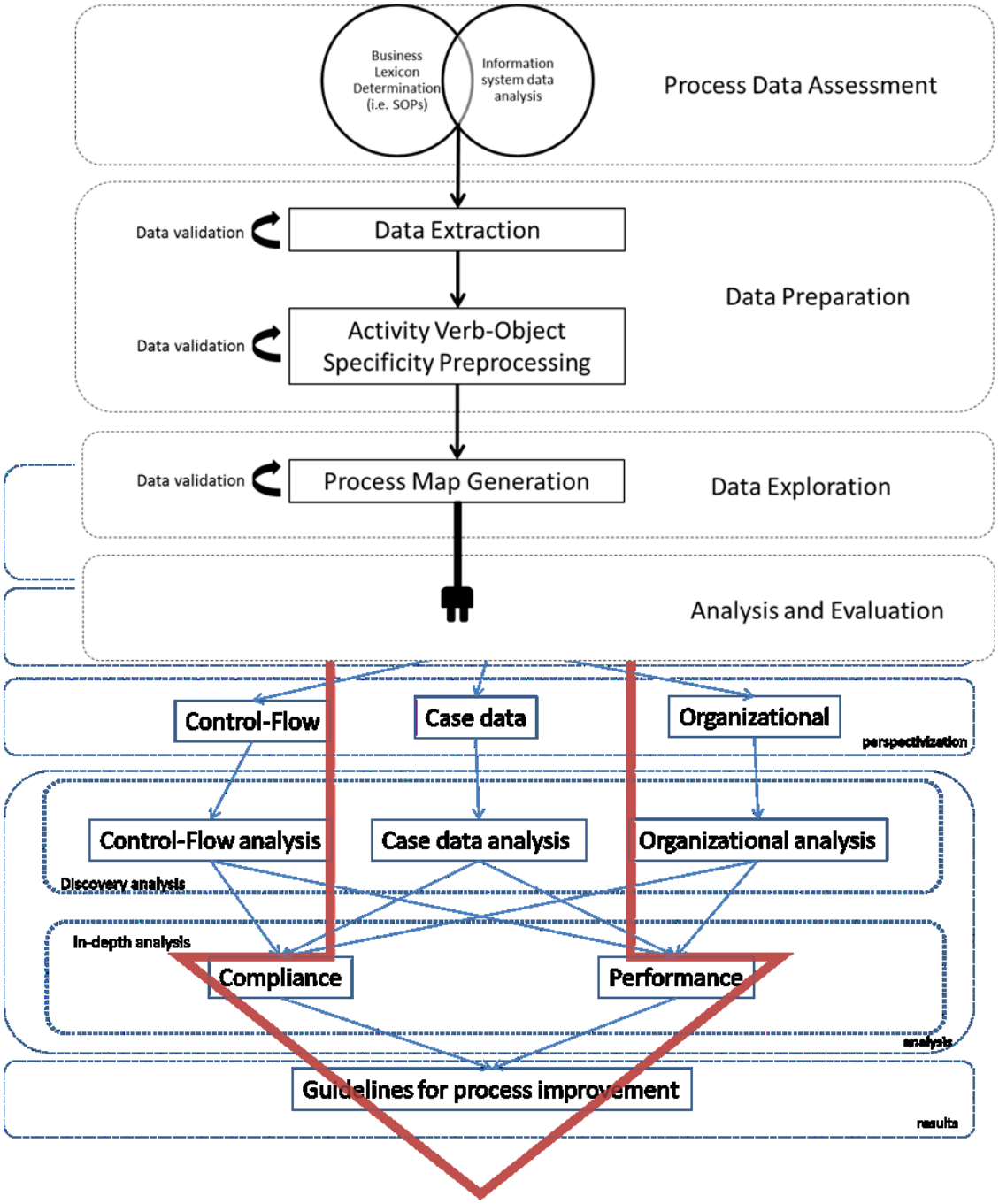


Figure 40: TRUISM Framework plugged into Process mining methodology framework highlighting approach for two case studies (adapted from De Weerdt *et al*.[48])

3.5.2 Interpretation of Qualitative Methodology Results

A qualitative methodology was used to evaluate the comments and questions arising from both case study focus groups. The qualitative evaluation methodology consisted of two separate thematic coding stages and a descriptive stage of the results, the latter according to the PMMF[48].

The first coding stage consisted of thematic codification of comments with two predefined codes-"value based insight" and "process based insight", consistent with that described in other industries [49, 50]. This allowed for the determination of how the SMEs were thinking about the process maps. The rationale is that if the process mining exercise generated more value-based thinking, then it would be less likely that this technique could provide more details than those provided by current value-based techniques. Likewise, if the process mining exercise demonstrated increased process-based insight, then this would have highlighted a complementary nature of this technique to traditional value-based techniques.

The second stage of the qualitative evaluation methodology was to use arbitrarily defined free-codes summarizing the intent of the comments and questions, which allowed for a high-resolution but thematic view into the comments and questions.

The final portion of the analysis included a narrative description of any guidelines for process improvement that may have arisen from the focus group exercises. If the exercise was successful and functional, then there will likely be an element of process insight with resultant formulation of guidelines for process improvement. These results would validate the potential for real-world application and business use of process mining in the understanding of the transfusion inventory.

Within the Hemosafe™ analysis, there was a sub-analysis designed to address the question of the compliance of the real process to the process as anecdotally known by the focus group participants, and encompassed the compliance portion of the in-depth analysis noted in the PMMF [48]. Prior to the focus group starting, participants were given approximately 15 minutes to construct a process map using building blocks which were pre-defined from a subset of activities modified from a representative Hemosafe™ process map in Figure 41 as outlined in Table 29, with some modifications to simplify and group some activity codes

Table 29: Activities from Figure 41 for Use As Building Blocks for Manual Process Creation in Hemosafe™ Focus Group, and Modifications from Original Activity Codes

Activity	Modification
Received HI BTS	Unmodified
Unconfirmed HI BTS	Unmodified
Confirmed Hi BTS	Unmodified
Available HI BTS	Unmodified
Transferred VG BTS	Unmodified
Available VG BTS	Unmodified
Transferred VG BTS Hemosafe	Unmodified
Available VG BTS Hemosafe	Unmodified
Crossmatched VG BTS Hemosafe	Unmodified
Issued VG BTS Hemosafe 11A	Unmodified
Transfused VG BTS Hemosafe	Unmodified
Crossmatched VG BTS	Unmodified
Issued VG BTS (MSI/11A/8A/MDU)	Combined separate MSI, 11A 8A, and MDU codes.
Transfused VG BTS	Unmodified

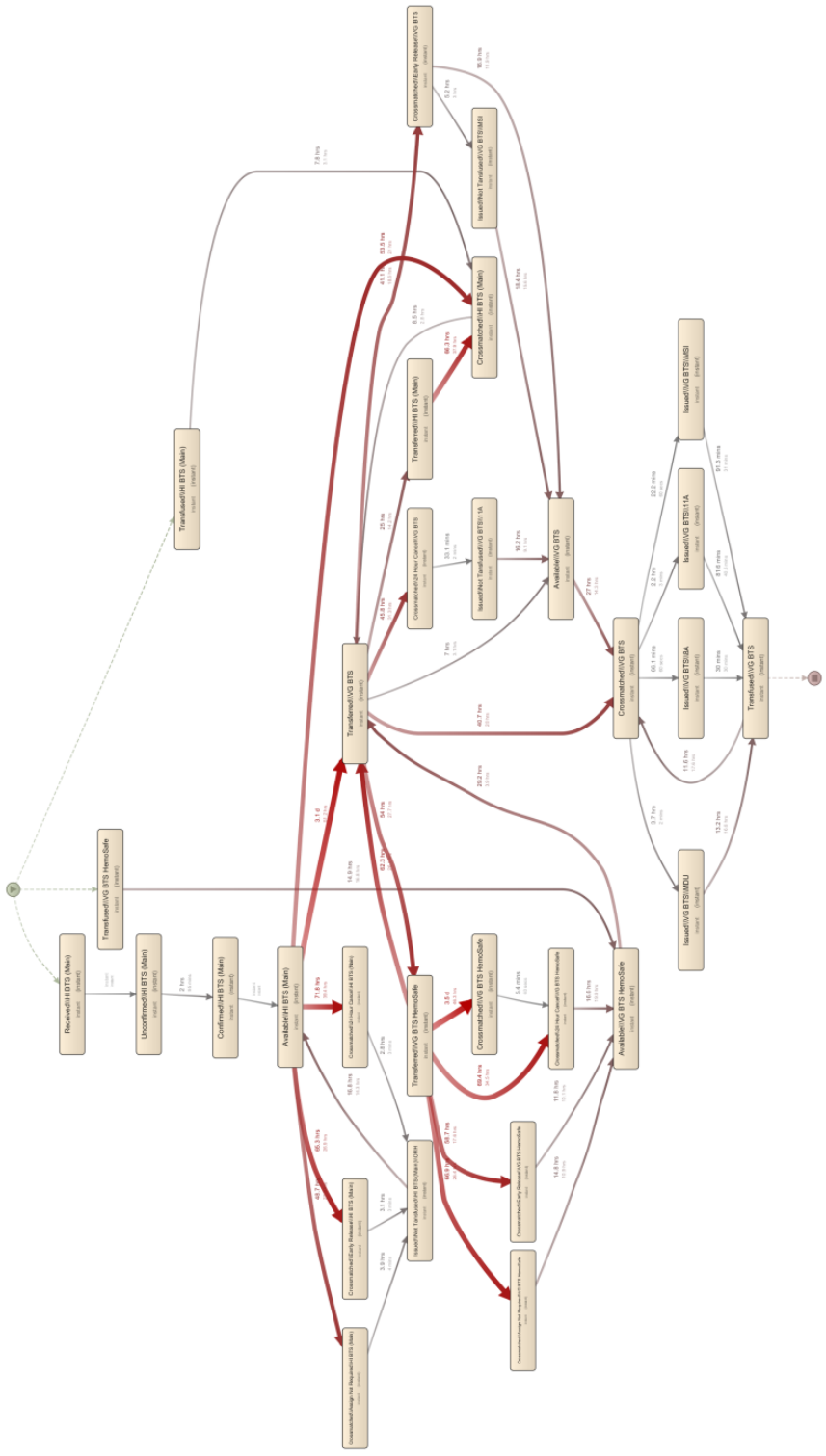


Figure 41: Process map defining the 'universe' of process activities relating to the Hemosafe™ fridge process (7.8%/0% activity/pathways resolution) (note: figure not intended to show detail)

3.5.3 Focus Group 1: Hypothesis Generation for Root Causes Associated with RBC Unit Wastage Events

Root cause analysis (RCA) is directly quoted by Washington State Department of Enterprise Services as "*a systematic process for identifying "root causes" of problems or events and an approach for responding to them. RCA is based on the basic idea that effective management requires more than merely 'putting out fires' for problems that develop, but finding a way to prevent them*" [223]. In BTS, unwanted wastage is a cost to the system as previously discussed, and it has been noted that the lack of data between hospital and supplier information systems was hampering the ability to reduce avoidable discards or to appreciate how inventories are handled on the hospital side [2]. Currently, the CZBTS only documents and trends the avoidable discard rates and does not examine why they happen. The reality of day-to-day practice is that blood products are wasted in various ways, and it is not practical to manually audit all possible antecedent events related to wastage. Therefore, that critical first step towards performing an ultimate in-depth RCA cannot be done. The goal of this focus group was to use a dynamically-generated process map to derive actionable hypotheses related to the root causes of RBC unit wastage, allow for further investigation, and eventual process improvement, in contrast to what is classically done (Table 24 and Table 25).

The steps involved in the focus group were as follows:

1. SMEs from CZBTS at NSHA, (medical director, manager, other management, technologists) were invited to an unpaid 1-1.5 hour focus group. SMEs from all levels of seniority voluntarily participated. They had prior knowledge of the blood transfusion service business processes and could articulate their thoughts in the context of current practice.
2. The process maps were dynamically-generated according to guidance from the participants from the previously validated and explored dataset from September 1, 2013-September 30, 2015.
3. The participants were then asked to select a wastage reason by consensus, to self-interpret the dynamically-generated process map, and speak their thoughts out loud using the CTA methodology.

4. This was all done in the context of a SME moderator (the researcher) in process mining and blood transfusion. Any software navigation was done by the moderator, who also responded to the participants' requests and encouraged discussion when needed.
5. The following series of open-ended questions were designed to encourage conversation in the context of the CTA protocol if there was lack of conversation flow:

    1. *What was highlighted in the wastage process that you previously knew about?*
    2. *What was highlighted in the wastage process that you didn't previously know about?*
    3. *Was there something missing that you were expecting to see? If so, describe it.*
    4. *What was easy or difficult to use or interpret?*
    5. *Why do you see this as an important tool for a transfusion service to use?*
    6. *How will you do things differently in your transfusion practice as a result of this?*
    7. *What other transfusion or laboratory-specific issues could be addressed using this technique?*

3.5.4 Focus Group 2: Policy Conformance Evaluation of Hemosafe™ Fridge Use

The Hemosafe™ fridge (http://angelantonilifescience.com/) is a 'smart' fridge, analogous to a vending machine for red blood cell units; however, these fridges have features which allow for RBC units to be securely allocated. These fridges were initially bought for the NSHA (then Capital District Health Authority) to theoretically improve the utilisation of RBCs by reducing the amount of RBC units returned from the operating rooms. The VGH site was implemented first with the HI site to follow. Since implementation of the Hemosafe™ fridge at the VGH site, an actual audit of RBC unit utilization was not conducted, as it was assumed that the Hemosafe™ fridge was being used. A high-resolution manual audit was unable to be performed due to constrained staffing levels, and even if an audit was done, there would be multiple points for observation at different times during the day and the data would be very complex to analyse. Therefore, the goal of this focus group was to use a dynamically-generated process map to not only evaluate if the fridge was being used, but how the fridge was being used. This is in addition to what is normally done with KPI-based analysis, which could be a variation of the results found

on Table 27. Another goal of this focus group was to qualitatively determine the differences between a human anecdotally-generated process map and a process mining generated process map.

The steps involved in the focus group are as follows:

1. Similar to the RBC wastage focus group, SMEs from CZBTS at the NSHA (medical director, manager, other management, technologists) were invited to an unpaid 1-1.5 hour focus group. SMEs from all levels of seniority voluntarily participated. They had prior knowledge of the blood transfusion service business processes and could articulate their thoughts in the context of current practice. These SMEs may or may not be similar to those in Focus Group 1, as there may be technologists more familiar with the Hemosafe™ fridge process who are not involved in red cell wastage and vice versa.
2. Prior to the exercise featuring process mining generated maps, participants were asked to construct a process map from building blocks given to them (10-15 minutes maximum) on a whiteboard. These building blocks consisted of the activities related to the Hemosafe™ pathway and were previously printed on paper and cut out. They were allowed to embellish the process map in any way, but at minimum draw arrows between the various activities. There was no help given from the moderator of the focus group.
3. As the implementation date for the Hemosafe™ fridge was February 24, 2014, a process map, filtering for the term 'Hemosafe' within the Disco software, was generated for the period before and the period after Hemosafe™ implementation. It was likely that the period before implementation also included some validation activities.
    a. September 1, 2013-February 1, 2014 (before implementation)
    b. March 1, 2014 – September 30, 2015 (after implementation)
4. A SME moderator (the researcher) navigated the process mining application and applied filtering functions while the participants self-interpreted the map data and examined different aspects of it while they spoke their thoughts aloud (CTA methodology).
5. The following series of open-ended questions were designed to encourage conversation in the context of the CTA protocol if there was lack of conversation flow:

1. *What was highlighted in the Hemosafe process that you previously knew about?*

2. *What was highlighted in the Hemosafe process that you didn't previously know about?*
3. *Was there something missing that you were expecting to see? If so, describe it.*
4. *What was easy or difficult to use or interpret?*
5. *What have you learned about the use of the Hemosafe fridge at the VG from this analysis?*
6. *How will you do things differently in your transfusion practice as a result of this?*
7. *Why do you see this as an important tool for a transfusion service to use?*
8. *What other transfusion or laboratory-specific issues could be addressed using this technique?*

3.6 Thematic Analysis Methodology for Qualitative Analysis Component

The goal of the thematic analysis was to categorize the comments arising from both process mining focus groups in order to characterise the thought process and conversation of the group of participants. This thematic analysis was intended to show that process mining generated meaningful process-related discussion, if any, can lead to the generation of guidelines for process improvement.

The first step was to transcribe individual phrases and sentences of each focus group into the qualitative analysis software in their respective analyses. The second step was to qualitatively determine, in broadest terms, the type of discussion arising from the process mining focus group. Comments were classified according to whether they were 'value-based insight', 'process-based insight'. Comments could be also be assigned both categories and those that were neither process nor value-based were not codified and had a null code value (Table 30).

Table 30: Examples of Codification into Value or Process-based Comments.

Examples of comments codified as "value-based insight"	- one percent used - so, the 35 units are much less than true transfusion - OR has a very high number
Examples of comments	- huge problem is the unidirectional interface between Cerner and

codified as "process-based insight"	the blood fridge - is VG site used for after-hours storage? - it is going to be more complex at the HI - more reasons, more errors, and make sure there are no unnecessary ones - OR people confused with stamping and how.
Examples of comments codified as both "process-based insight" and "value-based insight"	- how did those 35 units get transfused?
Examples of comments that were not codified (null)	- if we filter on transfused, you can't figure out the units that aren't transfused - Can you explain the arrows and their thickness? (in the frequency process map) - Can you explain the arrows? - I'm not seeing it

The third step of the thematic analysis was to create codes according to the comments that were captured. This step was done to examine if there are deeper themes which could subgroup comments, and would allow for a deeper qualitative characterisation of the discussion. The assignment of this code was mutually exclusive of the category that was assigned in the prior value-based/process-based coding exercise in the first step. The assignment of this code was arbitrary and the same code could be assigned to different comments. One comment could also have several codes assigned.

The thematic code assignment for both steps of the analysis was done by the unblinded researcher. For both steps, the groundedness of a code referred to the number of times that this code was used in the code assignment exercise.

The qualitative analysis software that was used was the ATLAS.ti™ platform (7.5.15) [http://atlasti.com/].

3.7 Knowledge Translation

The knowledge translation component of the research involves sharing results with the transfusion community and having the results of the exercise influence transfusion practice. This is analogous to the 'guidelines for process improvement' result which is the goal of process mining, as outlined by De Weerdt [48]. The most local definition of the transfusion community where knowledge translation can occur is in the context of the participants of the focus groups and case studies involving relevant business problems or areas of the process which require optimisation. As the research was conducted using business problems relevant to the SMEs of the focus groups, there may be a direct translation of any results into practice. The measurement of the impact of the translation into practice was not evaluated in nor was it the focus of this research. The intentional engagement of SMEs within the focus groups allows the stakeholders to be immediately engaged in a potential transformation of their business process. This consensus-based process discovery cycle based on the visualization of the process map generated using process mining ideally should lead to problem-solving and process optimization if the process is not optimised. Conversely, if the process is optimised, then the exercise would confirm that it is functioning well.

Knowledge translation can be applied to the greater transfusion community but may be more difficult given the already variable business practices highlighted by Devine *et al. [2].* Knowledge translation in the wider transfusion community may involve comparing one's own process map with that of other institutions. As an example, de-identified transfusion product process log data from an institution with a high performing inventory could be provided as a reference for comparison against another transfusion service's process mining-generated map. This could allow SMEs at other institutions to potentially benefit from the discussion that arises from the visualisation of the reference inventory's optimised processes and inventory flows. This type of sharing may have to occur in the context of providing that institution's standard operating procedure, standardized product and activities nomenclature, and may require appropriate data sharing agreements. This concept, however, will not be further explored in the research.

Chapter 4 Results

The results section will be divided into a general description of the focus group participants, a thematic examination of the coded comments, and whether or not tangible guidelines for process improvement were found, the latter outlined as the output as defined in the final phase of the PMMF [48].

4.1 Red Cell Wastage Focus Group Process Mining Analysis

4.1.1 RBC Wastage Focus Group Participant Characteristics

There were 9 participants in this focus group and participant characteristics are shown below in Table 31. Participants were mostly female, having a mean of 25 years' experience in BTS, from both medical and administrative domains. The administrative participants had variable administrative rank and had work experience from the main transfusion sites. Apart from having no participants from the medium or small community hospitals, the participants were representative of the typical clinical/administrative composition of a blood transfusion services quality committee. The focus group took 1:02:37 hours to complete according to the timestamp data.

Table 31: Red Cell Wastage Focus Group Participant Characteristics

Gender	Male Female	1/9 (11.1%) 8/9 (88.9%)
Years of experience in Blood Transfusion	Range Mean Median	7-41 years 25.0 years 23 years
Administrative Rankings	senior clinical informatics analyst hematopathologist, BTS medical director quality coordinator technical specialist lab technologist supervisor technical manager	
Site of Current Work	Mackenzie Building Halifax Infirmary Victoria General Hospital	

4.1.2 RBC Wastage Thematic Analysis

There were 163 separate comments and questions documented from the discussion of the focus group. The thematic analysis demonstrated a 7.8 fold difference in how comments were grounded on process-based vs value-based insight (Figure 42). There were 40 comments that were unable to be classified in either category.

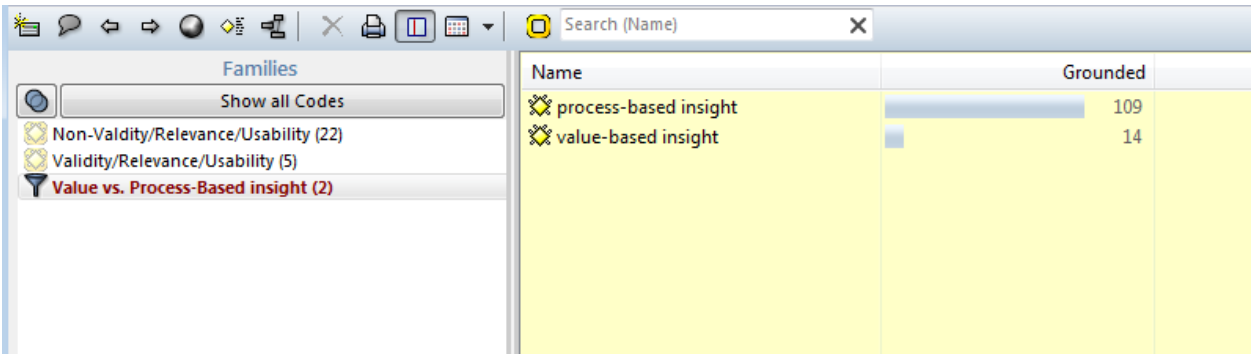


Figure 42: Qualitative analysis of groundedness of 163 comments according to process or value based themes for RBC Wastage focus group (Atlas Ti™code manager screenshot)

The thematic free-coding analysis (Figure 43) demonstrated that the top four most grounded codes (248/344, 72.1%) were related to insight (25/344, 7.3%), process confirmation (34/344, 9.9%), clarification (74/344, 21.5%), and future application (115/344, 33.4%). The rest of the groundedness codes in decreasing frequency pertained to specific discard reasons, but there was also discussion about CSA and AABB standards, and how they applied to the process maps. The free-codes did not add up to the total number of comments as multiple thematic classifications occurred with the same comment.

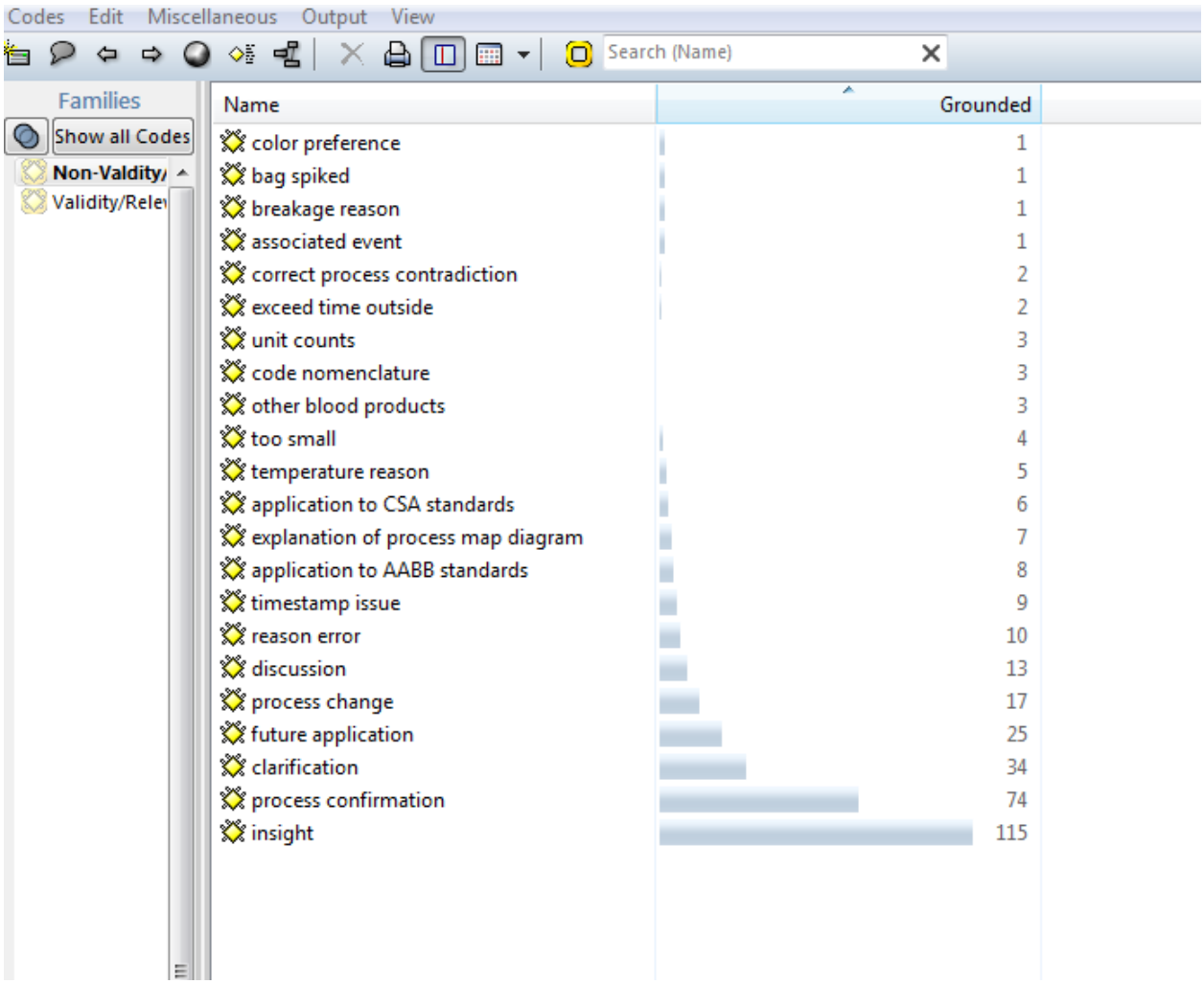


Figure 43: Qualitative analysis of groundedness using free-coding of concepts related to wastage process maps (Atlas Ti™ code manager screenshot)

4.1.3 Process Mining Methodology Framework – Results of RBC Wastage Process Mining Focus Group

There were several guidelines for process improvement resulting from the focus group exploration of the process maps, fulfilling the functional end goal of the PMMF.

The first guideline for process improvement originated from a conversation on the discrepancy between the AABB and CSA standards pertaining to temperature vs. time on returned blood products. This conversation originated from the participants noticing that there was a common location for both of these wastage reasons, which was from the operating rooms. This triggered a secondary conversation on the potential conflict of the two standards, and they also noted that it was difficult to timestamp the units and they would be almost universally discarded. Therefore, the BTS elected to rely on temperature (the AABB standard) rather than timestamp for return products, as it made the process easier for them and eliminated one step for the operating room staff. A deviation form would have to be filled out every time this temperature process was exercised, but that was felt to be an acceptable process, allowing for the reduction of avoidable wastage. Although there was a parallel discussion on potentially having a time stamping cheat sheet posted on the fridge to remind staff how to use it, this train of thought was eventually abandoned.

The second guideline for process improvement involved units being wasted secondary to a security tag being absent or broken, which brought up a secondary discussion about accelerating a blood product transfer agreement with the Prince Edward Island government to avoid wastage of these units during transfer. The status of the transfer agreement was to be followed up after the process mining focus group. Most of the reasons, however, for this wastage reason resulted from the impression that there was increased staff turnover at CBS, and that they had been forgetting to put on security tags prior to sending the units.

The third guideline for process improvement involved technologist reflection on how they input error codes. They indicated that they will need to increase the rigour of their error code entry by a technologist into the LIS, as there were processes that did not make sense. For instance, case variant #1 and #2 (Table 32) illustrated the immediate receipt and immediate process of destruction of RBC units with a broken or absent security tag. Case variant #3 fell outside that process, as it demonstrated a unit going through the confirmation process before being transferred prior to destruction from the same reason. Upon further questioning, this could never physically have happened, as this error code should have always followed the sequence of

activities outlined in case variants 1 and 2 (Table 32). Therefore, the focus group indicated that the third variant represented a coding error.

Table 32: Example of Case Variant Analysis of Units Wasted Due to Security Tag Broken or Absent (17 Total Variants)

Case Variant	Frequency N(%)	Activities
1	11/35 (31.43%)	Received\\HI BTS (Main) →Unconfirmed\\HI BTS (Main) →Disposed\\HI BTS (Main)\\Security Tag Broken or Absent →Destroyed\\HI BTS (Main)
2	4/35 (11.43%)	Received\\HC BTS →Unconfirmed\\HC BTS →Quarantined\\HC BTS →Disposed\\HC BTS\\Security Tag Broken or Absent → Destroyed\\HC BTS
3	2/35 (5.71%)	Received\\HI BTS (Main)→ Unconfirmed\\HI BTS (Main) → Confirmed\\HI BTS (Main)→ Available\\HI BTS (Main)→ Transferred\\VG BTS → Transferred\\HI BTS (Main)→ Quarantined\\HI BTS (Main) →Disposed\\HI BTS (Main)\\Security Tag Broken or Absent → Destroyed\\HI BTS (Main)

4.2 Automated Blood Fridge Analysis Results

4.2.1 Hemosafe™ Focus Group Participant Characteristics

There were 8 participants in this focus group and participant characteristics are demonstrated below in Table 33. All participants were female, having a mean of 20 years of experience in BTS. Participants were solely in the administrative and technical domains, with variable rank. As in the RBC wastage focus group, they were from the main hospital transfusion laboratory, and there were no participants from the medium or small community hospitals. The composition of the focus group was consistent with a typical clinical/administrative committee, such as the quality committee, with the exception of the lack of a physician being present, though it was noted that the moderator (the researcher) would typically fill this role. The focus group took 58:48 minutes to complete the session based on timestamps.

Table 33: Hemosafe™ Utilization Focus Group Participant Characteristics

Gender	Male Female	0/8 (0.0%) 8/8 (100.0%)
Years of experience in Blood Transfusion	Range Mean Median	4-41 years 20.0 years 20.5 years
Grouped Administrative Rankings	lab technologist supervisor technical specialist senior clinical informatics analyst quality coordinator technical manager	
Site of Current Work	Mackenzie Building Halifax Infirmary Victoria General Hospital	

4.2.2 Compliance (Conformance) Analysis of Process Map

The comparison of the manually-generated process map (Figure 44) to Figure 45 and Figure 46, which are low-resolution (<2% activity and pathway resolution) process mining-generated process maps, yielded several key findings. The first difference in the process maps is that the manually-generated process map demonstrated a direct flow from receipt to transfusion. There were no transfer loops between the VG and HI inventories, or between the VG Hemosafe and VG general inventories, as noted in the process mining-generated maps. The manually-generated process map did not contain any details about time or frequency of RBC units between the process steps. There was also no appreciation in the manually-generated process map for the eventual transfusion destination of the medical day unit (MDU), which represented outpatient oncology services. This meant that the Hemosafe™ fridge was not used to supply the operating rooms directly, and was, instead, used to supplement regular non-operating room inventory in the rest of the hospital system. This caused an unnecessary flow-through and looping of RBC inventory units through this fridge location. There were some equivalent process representations between the manually-generated and process mining-generated maps. Firstly,

when there was transfusion from the Hemosafe™ fridge, both types of maps demonstrated transfusion to the 11A location and 10A operating rooms. Also, the receipt process was consistent between the two types of maps, with a one-way flow from receipt to confirmation and the unit becoming available.

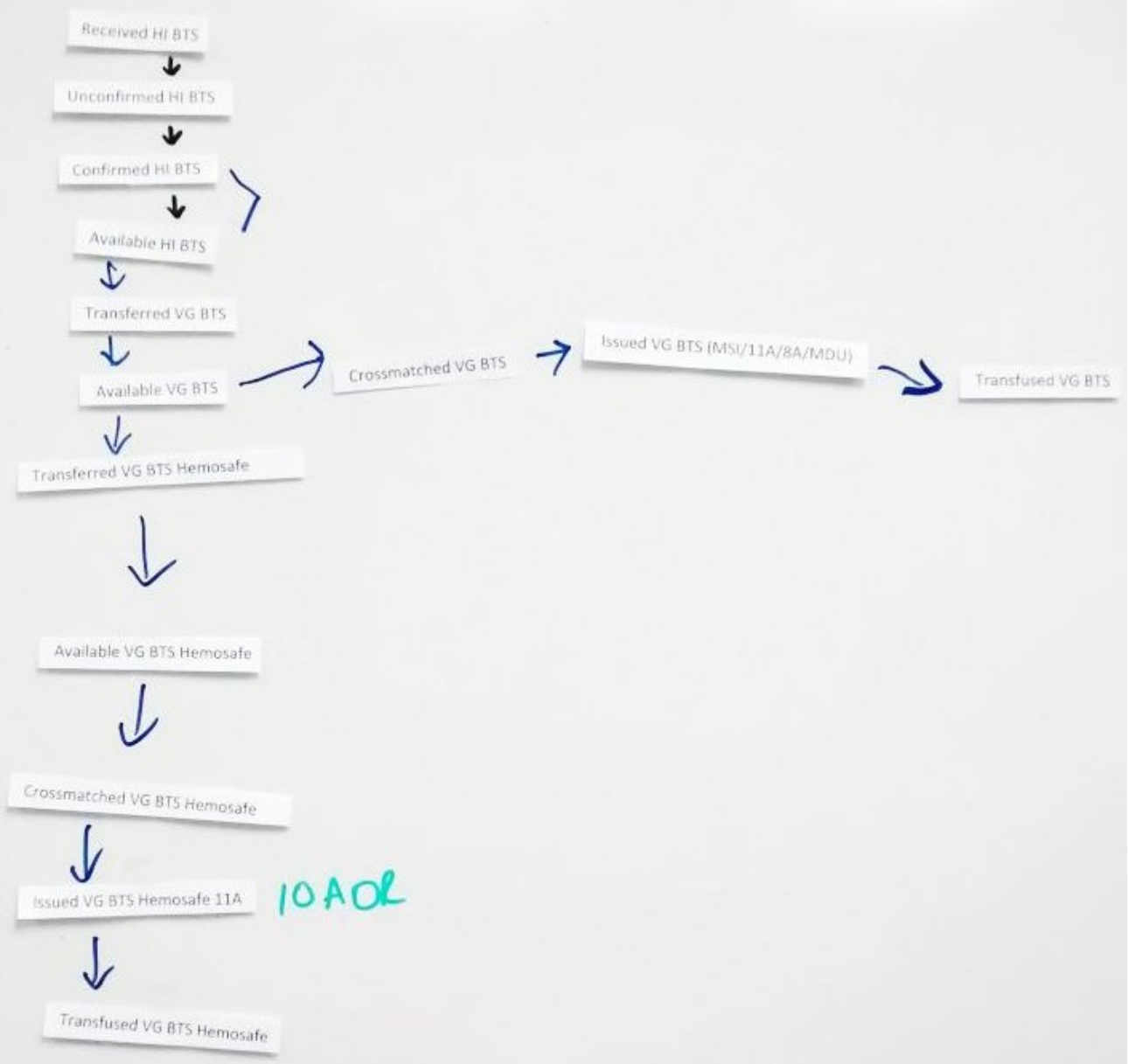


Figure 44: Picture of focus group members' anecdotal process map involving Hemosafe™ fridge prior to formal process mining exercise.

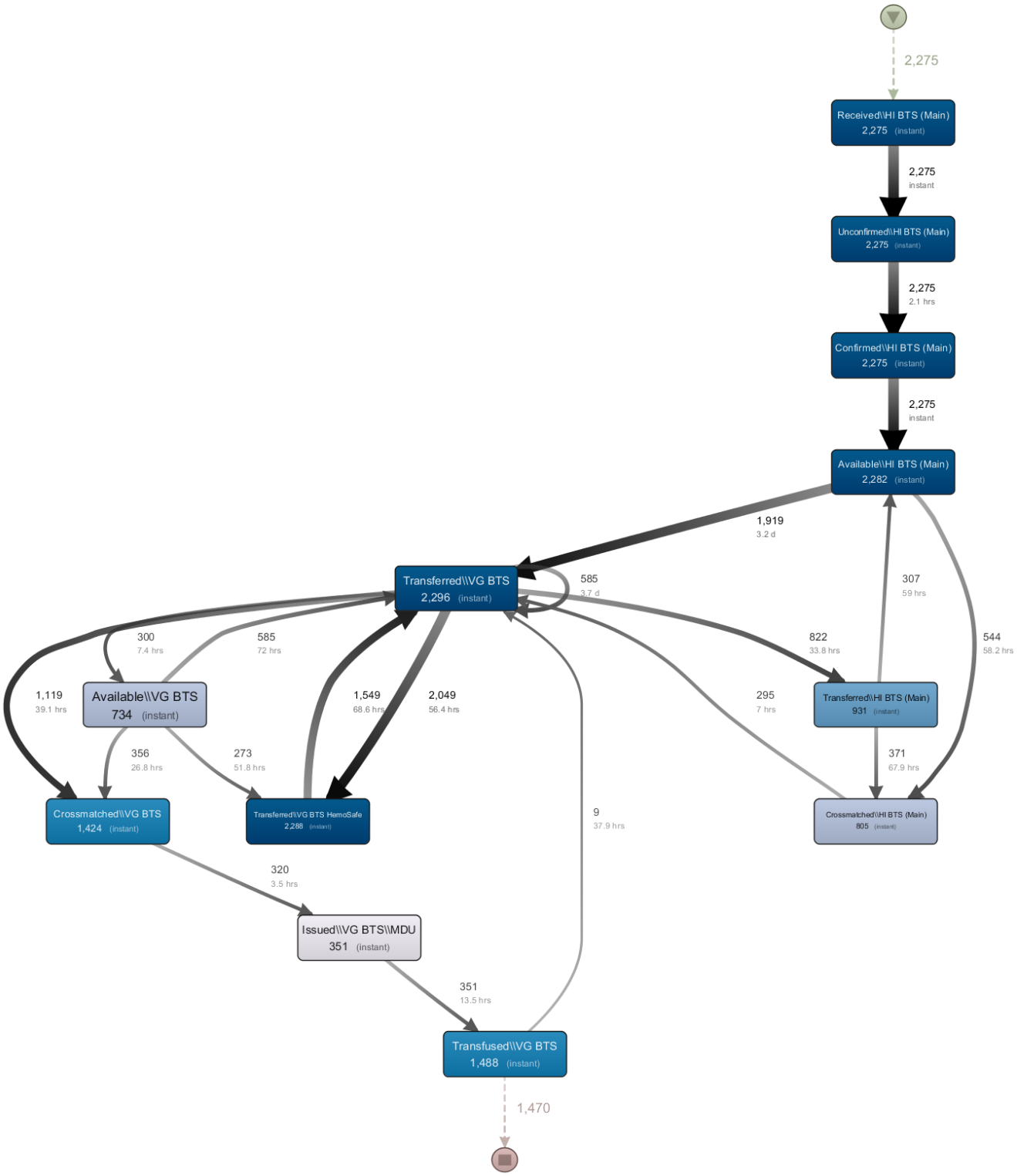


Figure 45:Low Resolution Case Frequency (bold) and mean duration (small print) Process Map Generated at 1.5% and 1.7% Activities/Paths resolution

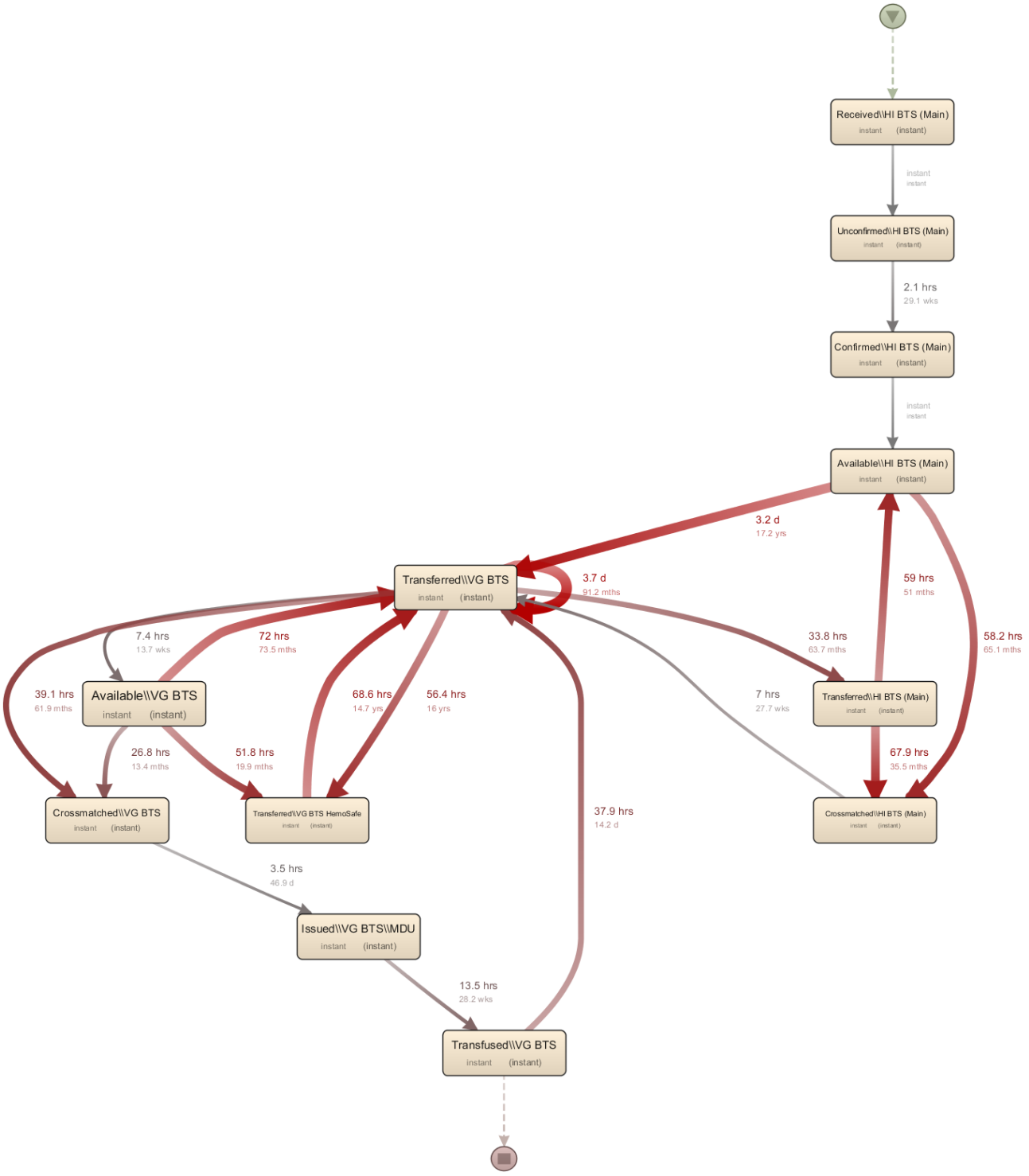


Figure 46: Low Resolution Mean Duration (bold) and Total duration (small print) Process Map Generated at 1.5% and 1.7% Activities/Paths resolution

There were 141 separate comments and questions documented from the focus group discussion. The first thematic analysis demonstrated a 4.7 fold difference in how comments were grounded on process-based versus value-based insight (Figure 47). There were five comments that were unable to be classified in either category.

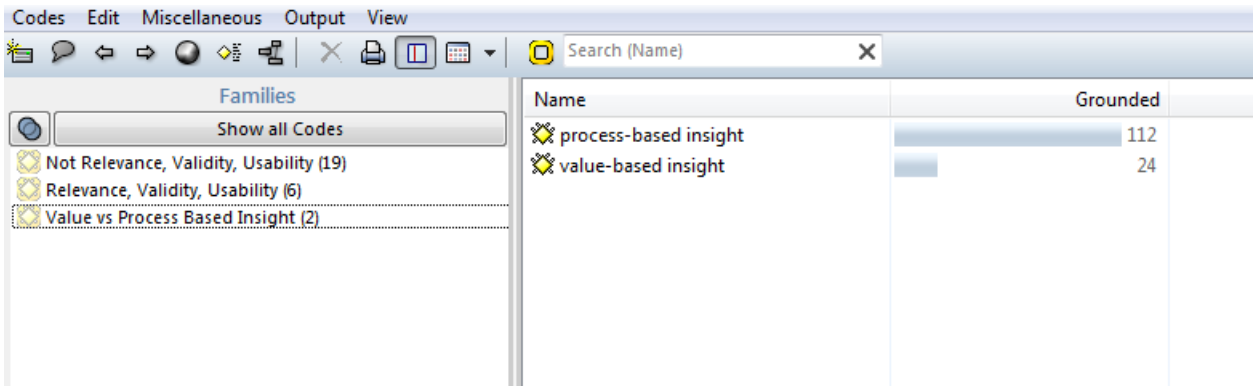


Figure 47: Qualitative analysis of groundedness of 141 comments according to process or value based themes for Hemosafe™ focus group (Atlas Ti™ code manager screenshot)

A thematic free-coding analysis of the comments demonstrated that the majority of comments were grounded on process confirmation and insight (54.6%, 143/262), but mostly to insight (38.2%, 100/262) (Table 34).

Table 34: Qualitative Analysis of Groundedness using Free-coding of Concepts related to Hemosafe™ Usage Process Maps

Name	Grounded
text size	1
questioned data	1
usability	1
process comparison	2
training issue	2
unidirectional interface	2
data limitation	3
clarification	3
questioned process map	3
filtering question	3
unit counting	7
process change	9
process deviation	9
error in status	12
error detection	13
future application	15
reason code	33
process confirmation	43
insight	100

4.2.3 Process Mining Methodology Framework – Results of Hemosafe™ Process Mining Focus Group

There were several guidelines for process improvement resulting from the focus group exploration of the process maps, fulfilling the functional intent of the PMMF.

The first guideline for process improvement identified various weaknesses with the Hemosafe™ workflow. Firstly, it was noted that a bidirectional interface should be explored with the Hemosafe™ company, or that process improvements with the unidirectional interface should be done in the meantime to improve usability of the fridge. The interface contributed to a complex workflow from a technologist and nurse standpoint, which may have deterred the use of the

fridge. It was also noted that this inventory location was used as a flow-through inventory location for RBC units on their way to being transfused at non-operating room destinations as discussed previously. This flow-through nature was not previously known nor intended, as the manually created process map demonstrated the unidirectional flow to 11A and 10A operating rooms (Figure 44, vertical flow). The BTS needed to internally examine how they could improve that process and simplify it. It was also determined that the internal workflow of the fridge was not available. This missing information could have been helpful to the process mining exercise, as the information resided on the BloodTrak™ middleware, rather than the LIS. Access to this middleware was not provided by the Hemosafe™ despite many prior requests. It was suggested that Hemosafe™ was to be engaged, and at the time of the writing of this manuscript, they sent a consultant to the CZBTS for a site visit, consultation, and generation of a report with some recommendations. The recommendations included the purchase of new fridges with bidirectional interfaces and relocation of them in the operating rooms.

The second guideline for process improvement involved a similar issue as previously discussed in the RBC wastage focus group. The process maps highlighted the wrong error or disposition codes being manually entered into the LIS. This resulted in increased vigilance by way of a technologist-generated e-mail reminder to the staff of the importance of proper code assignment in the LIS.

The third guideline for process improvement involved the use of live codes for training nursing and technologist staff on how to use the Hemosafe™ fridge. This was picked up by noting that there were some RBC units which had abnormal repetitive process pathways (data not shown), which skewed the process data in the LIS. Once this was identified, a coding exercise was performed during a transfusion quality assurance committee meeting, and new training codes were implemented to allow for training-related processes to be filtered out in future process mining analysis.

The fourth guideline for process improvement was an acknowledgement that the complex operational nature of the VG Hemosafe™ implementation described above needed to be avoided in the future plans for a second fridge at the HI site. It was noted that the first fridge was implemented without significant PInG input. Therefore, the guideline for process improvement was to involve PInG more formally in the earlier stages of design and examine if there were any workflow changes that could be done prior to the HI site Hemosafe™ going live. This could have included various interface workarounds or custom coding in the LIS to emulate some functionality of a bidirectional interface, potentially improving usability and adoption of its use.

4.3 Summary of Results

Overall, the process mining focus groups for both RBC wastage and Hemosafe™ fridge usage demonstrated primarily process-based insight versus value-based insight. There were various guidelines for process improvement that were generated from those exercises.

Chapter 5 Discussion and Conclusion

This chapter will highlight the conclusion of the research and discuss aspects of the research, including the validity of the research design, results interpretation, insight that was gained, limitations of the research, and recommendations for future research.

5.1 Conclusion

The research demonstrates the utility of using a process-mining based transfusion inventory framework (TRUISM Framework) to understand the RBC inventory as a process map, resulting in SMEs making meaningful observations and ultimately derive guidelines for process improvement. It highlights the potential of process mining using data from a process-unaware system to generate process maps leading to process improvement guidelines in a real-life setting.

The first case study demonstrates how SMEs in transfusion medicine, composed of participants, who typically would be found in transfusion quality assurance or process committees, used process mining to yield process-based insight from LIS data on RBC wastage reasons, in addition to some value-based insight. Free-coding analysis demonstrates that process mining in this context also brought process-based clarification and confirmation in the participants. The process mining exercise also precipitated guidelines for process improvement, including the need to have blood returns done by temperature alone rather than time, the need to follow-up on a transfer agreement to avoid discarding units that came from other provinces, and that the technologists needed to be reminded to properly enter error codes into the system.

The second case study demonstrates the utility of process mining in two different ways by examining the role of the Hemosafe^{TM} fridge in the transfusion inventory using LIS data. The first way is that it highlights the limitations of human anecdotal experience on processes, by way of

manually-generated process maps, and highlights that process mining-derived maps are much more reflective of the actual process than manually-created maps. The second way that process mining demonstrates utility is that the insight from the discussion was nearly all process-based and very little was value-based. In addition, a similar group of SMEs demonstrate that process mining brought process insight and confirmation of the existing process. The guidelines for process improvement resulting from the focus group included the re-engagement of the manufacturer, the use of internal resources to further improve the functioning of the workflow, accurate entry of error or disposition codes into the LIS, and a larger dictionary of codes being required to resolve the training environment from the live environment.

5.2 Discussion

5.2.1 Validity of Research Design

The validity of the research design can be rationalized from different perspectives. Firstly, the overall goal of the research was to build on the emerging understanding of the blood transfusion community, through the literature, that the RBC inventory is a process [36, 42]within a larger blood product supply chain [2, 5, 22, 36-46], and therefore could be understood with process related methodologies in addition to the traditional value based ones. As current frameworks were not directly applicable to the transfusion inventory, a generalizable framework which applies to transfusion inventory data in any LIS in any transfusion service was created. The TRUISM Framework is intended to increase the specificity of process map activities and yield a highly valid process map using iterative validation steps during data preparation and exploration in the right process documentation environment. This technique needed to be applied in an environment which could be immediately translated into day-to-day transfusion community practice with very low complexity and cost, as suggested by Devine *et al.* [2]. For this process mining research to be generalizable, it needed to occur in a day-to-day practice environment, using data that was easily obtainable, and with people who are directly involved with transfusion inventory management. To most emulate day-to-day practice, this is why the

research was done in focus groups that were similar in composition and format to a transfusion quality assurance committee, and using data that was readily available in the LIS.

The case study approach was used, as it closely followed a typical audit of a blood transfusion process. At CZBTS, usually, a problem will be identified by the quality assurance (QA) committee, the data surrounding a process will be gathered, validated, and analysed with a conclusion and next steps generated at the end of the exercise. The research was done using case studies in two areas pertaining to RBC inventory management techniques: RBC wastage and fridge inventory use. These have not been assessed in the literature but the RBC wastage question directly pertains to the avoidable wastage discussed in Devine *et al.* [2], and involves a deeper and alternative analysis of that question. Both case studies highlighted difficult and unsolved day-to-day transfusion inventory management questions in the value-based paradigm and in the context of limited financial and human resources.

The core of the research design leveraged and augmented a previously established methodology in the process-mining literature. The Process Mining Methodology Framework (PMMF) by De Weerdt *et al.* [48] illustrated how a case study should be conducted in the context of process mining with the end goal of guidelines for process improvement. This research developed the TRUISM Framework, adapted from elements of the PMMF, and was used to transform the process-unaware LIS data to be used as a reliable input for process mining software. This ensures that the guidelines for process improvement and case data analysis were grounded on valid process log data and a valid process maps. Incidentally, the comment coding component (free-coding) of the analysis practically augments the results component of PMMF. If no guidelines for process-related improvements can be generated, then a free-coding analysis of comments from SME examination of a process mining-generated map, is still helpful, in that the exercise at minimum would have yielded some process-based insight. These comments could include process confirmation insight, especially if a process is shown to be already optimised. Comment generation was done using the Concurrent Think Aloud (CTA) technique adopted from already-established usability testing principles whereby participants were encouraged to think out loud (usability.gov) as they interacted with the process mining in real time. There were

other feedback gathering methodologies including Retrospective Think Aloud (RTA), Concurrent Probing (CP) and Retrospective Probing (RP) techniques which could have been used, but none of them could have emulated what happens in real-time discussion at committee meetings. In addition, the real-time feedback allowed the moderator to manipulate the process mining tool and adapt to the questions of the participants. This could not have been done using other techniques, as they were either retrospective or required the in-depth probing of a participant. Although the feedback could have been gathered on questionnaires or surveys, none of these techniques would have approximated the real-life ebb and flow of discussion that occurs when data is presented in front of a transfusion quality committee. A thematic analysis was performed on the comments and questions associated with the focus groups, and this allowed the group's thought processes to be simplified, and allowed for the classification of those thoughts and comments to either value based or process based insights. The Hemosafe™ group, was asked to manually generate a process map according to anecdotal knowledge of the process. This allowed for a comparison between a process mining-generated map and a manually-generated process map according to blood transfusion SMEs. This process is valid, given the time constraints of the focus group, and is an efficient way of gathering data from a diverse group of users by having them work together. Another possibility is to gather individual manually-generated process maps from each participant, but this is inefficient, removes the communication between participants with possible loss of the richness of the discussion. The entire context of this was moderated by a combined transfusion SME and process mining expert (the researcher), which, in conjunction with transfusion SMEs, was suggested in the literature for effectiveness [48]. The moderator (the researcher) also serves as an expert in the functioning of the process mining tool to allow the SMEs to concentrate on the interpretation of the process maps without the requirement to learn the tool.

The choice to use the data that was already resident in the LIS, without patient identifiers or significant preprocessing was designed to emulate the data that was normally available during a quality assurance process, and was designed so that ethics approval was not necessary. RBC unit data was chosen, as RBCs are the most frequently used cellular blood products in the transfusion inventory [2], are stored and transported to multiple fridges and inventory locations at the CZBTS, and are important in the reimbursement for blood utilisation. There was also

justification for the study of the RBC unit inventory, as opposed to other processes or other products, as this is the only area in the literature whereby the inventory is described as a life cycle [37]. The goal of this study was to introduce the process mining methodology to other transfusion suppliers and hospital sites to complement their value-based methods in their day-to-day quality improvement workflow.

The data validation process was done continuously and iteratively according to the TRUISM Framework, and though there are some artifacts generated, they are not material to the objectives of the focus groups. The data validation was a key focus of the research, as the data was originating from a process unaware system [47] and was constructed using data from numerous tables of the LIS.

The process mining software, Disco, was chosen as it had a user-friendly interface, data importation capabilities and qualitative descriptive tools. It is also based on the fuzzy mining algorithm, which had previously described desirable characteristics. This was the only software tool available (at the time of writing this manuscript) that had these qualities, and represented a nimble tool which could be used in focus groups.

5.2.2 Results Interpretation and Insight Gained

The main accomplishment of the research was to develop a generalizable framework by which inventory data from process unaware information systems can be turned into valid process maps. In this case, it was developed in a transfusion inventory environment. This TRUISM Framework (TRansfUsion Inventory proceSs Mining Framework) allows one to assess the environment in which the transfusion process is taking place to determine if a lexicon is being used and if that data is being captured in the LIS. This Framework also provides guidelines on how to structure queries, improve the specificity of verb-object activities, and ensure that there is validation at every iterative step of data manipulation and handling, with the intent to

generate a valid process map based on valid process data. The intent of this Framework is to allow transfusion inventory data to be handled and plugged in as an input in any subsequent analysis. In this case, the PMMF [48] was chosen as the subsequent analysis framework and ultimately resulted in derivation of guidelines for process improvement. In this research, the TRUISM and PMMF framework combination was applied to two case studies – one in RBC wastage, and the other regarding the Hemosafe^{TM} fridge.

An incidental accomplishment of the research was the application of the free-coding analysis technique used to analyse focus group comments and augment the PMMF framework. This quantifies process-related insight and discussion prior to deriving guidelines for process improvement, if any. The application of this would be in a theoretical case where the process is already optimised, and there are no process improvements required. Through the comment gathering steps and subsequent free-coding analysis technique, confirmation of the optimised process can happen as the endpoint. Another theoretical application is when a SME group cannot come to a consensus on guidelines for process improvement in an evaluation of a suboptimal process, it is still possible to derive meaningful process-based insight. The PMMF does not appear to address the scenarios whereby no guidelines for process improvement occur, and this is a weakness of this framework.

The results from the RBC wastage focus group demonstrate the effectiveness of using a process-mining approach to a traditionally value-based, KPI-driven approach to the understanding of transfusion inventory wastage. As the traditional KPI-based approach has been the default paradigm for understanding the inventory it was surprising, that within the 1 hour-long focus group, the goal of the Process Mining Methodology Framework [48] was achieved, with three separate process improvement guidelines. The first guideline regarding the consolidation of the AABB and CSA guidelines effectively reduces the process complexity by streamlining the process by which RBC units are returned from the operating rooms. Despite the various initiatives which our institution instituted since 2006, it wasn't until the RBC return process was visualized as a process that the similarities between two different RBC wastage reasons could be discovered. As the process maps were being dynamically generated, various members were able to observe

the commonalities, and resulted in this guideline being tabled. The second guideline resulting from the process mining exercise was that some avoidable RBC wastage could have been prevented by an RBC transfer agreement between provinces, even though the main reason was thought to be due to high staff turnover at CBS. The exploration of this peripherally-related concept suggests that the process mining exercise may have resulted in an environment where process exploration was at the forefront of the participants' minds, and the exploration of this issue simply brought forward a related thought. The third guideline was a data integrity guideline which, in the value-based paradigm, would not have been present in a tangible way. Previously, in all circumstances examined, it was not possible to discover if the incorrect RBC wastage codes were being used, and it was generally accepted that the correct code was being used. In this instance, process mining uncovered the incorrect usage of a code, as the context that the code was used in was in an incorrect process. This is an important finding that highlights the potential role for process mining to error check a subset of RBC wastage codes which have a predefined process. It is likely that this error occurs in other institutions, which may result in incorrect interpretation or calculation of RBC wastage KPIs.

The key to the success of this process mining focus group can be also attributed to the diverse mix of expertise and experience found in the participants. Moreover, due to the CTA structure of the focus group, participants were able to reinforce and shape each other's thinking, especially since not all participants had similar insight into the overall process. For example, if the focus group was primarily composed of physicians, they may not have had much insight into the day-to-day working of the inventory, but may have insights into policies. Likewise, if the group was composed of primarily bench technologists, then they may not have had appreciation of overarching policies, though they may be expert in the day-to-day coding of wastage reasons. Therefore, it was serendipitous that there was a broad mix of experience and expertise, mimicking what is typically found in committees, such as the Quality Assurance Committee at CZBTS. It is likely that the results may not have been as successful or in-depth if there was no SME interpretation of the process maps in conjunction with a process mining expert, as indicated in the literature [48].

In addition to the broad guidelines, some finer academic details were also elucidated from the results. Most of the comments from the free-coding analysis were based around insight and confirmation of the process, as well as potential future applications and process changes. When

grouped into higher-level themes, such as value-based or process-based thinking, it was found that there were predominantly process-based comments originating from the exercise, and four times less value-based ones. There are several potential explanations for this. The first and most likely explanation is that participants were biased to know that this was a process mining focus group, and thus came prepared with a process mindset. An alternative explanation could be that the members of the group had already in-depth knowledge of various wastage rates from their extensive history working in blood transfusion services, and that there was no point to revisiting that knowledge, and thus bringing up those value-based concepts would not have helped the focus group significantly. This potentially parallels the thinking of Devine *et al.* [2] whereby "*the presence of gaps in the information required to really have an optimal inventory management system… [and] this lack of information hampers the ability to drive down the rate of avoidable discards"*, suggesting that what is known about avoidable wastage has already been maximized at the individual ends of the 'rope', and effort must be focused on the minimizing this perceived and functional information gap to optimise the blood supply chain.

It is possible that the lack of value-based responses in the focus group could simply reflect that the maximum utility of KPI descriptors have already been reached, and that the focus group is ready and willing to explore alternative analytical techniques such as process mining. While the statement concerning the 'information gap' by Devine *et al.* [2] may be correct, this research suggests that it may only be correct in the current value-based paradigm in how the inventory is understood. This research further suggests that process mining-based techniques applied to the same sources of information can further deepen the understanding of the transfusion product inventory in both the supplier and the hospital transfusion service. Perhaps the ends of the 'rope' should be thought of as a mine as shown in Figure 48 which depicts the value-based paradigm in context with the process-based one. Within the process-based paradigm, the dotted line reflects the current shallow value-based mining of the inventory data happening at both blood supplier and hospital BTS, and the solid red line depicts a greater mining depth with process-based analytical techniques. With a process-based paradigm, the existing minesites can be explored and exploited in greater depth and breadth than currently done with solely value-based techniques.

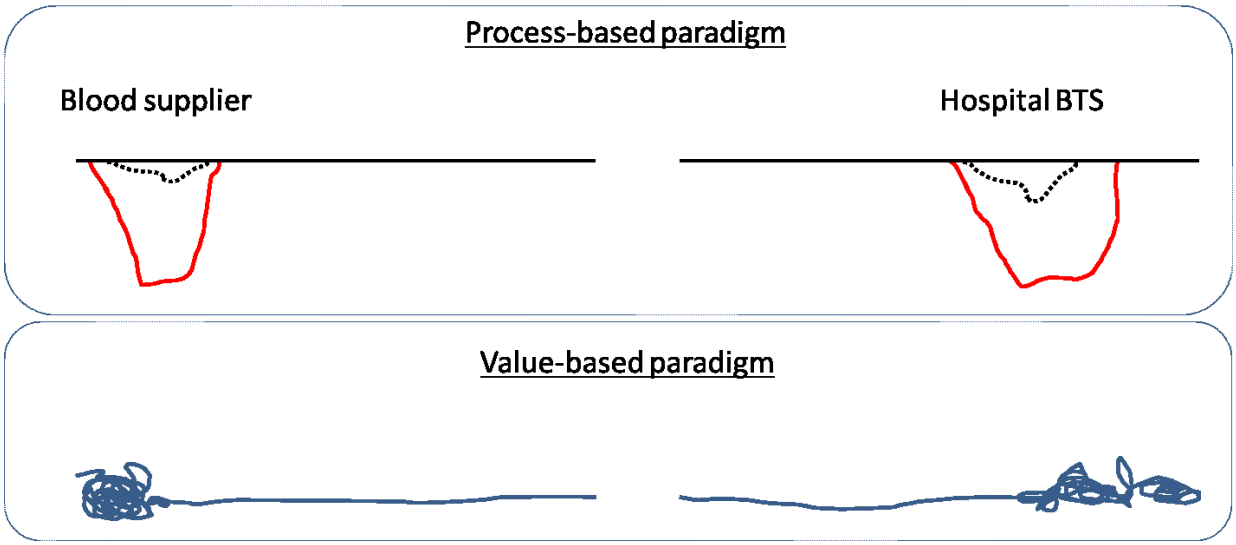


Figure 48: Process-based paradigm mine vs. value-based paradigm rope of transfusion inventory data (dotted line depicts current depth of value-based analysis; solid red line depicts depth of process-based analysis)

This may lead to further improvement of wastage metrics or transfusion efficiency without even needing to address the issue of the gap. For instance, in the case of avoidable discards, this research demonstrated further guidelines for process improvement where previous value-based analysis was unable to provide. The next frontier would be to re-evaluate global institutional transfusion data based on this paradigm prior to embarking on linking supplier and hospital information systems. In addition, as previously mentioned, there were data integrity issues uncovered by the process mining methodology. It is likely that this phenomenon of incorrect coding also occurs in transfusion services globally, and this would be a potential blind spot to correct prior to the linking of information across the gap as Devine *et al.* suggest should happen, since incorrect conclusions could easily unknowingly arise. Both of these illustrations suggest a complementary role for process mining in the current value-based paradigm in which RBC wastage is analysed and understood.

The results from the Hemosafe^{TM} usage focus group also demonstrate the effectiveness of the process mining technique to understand how, not only how much, an inventory location is being used. The traditional value-based paradigm would have only allowed the transfusion service to

quantify that less than a fraction of a percent of all units that were transfused came from the VG Hemosafe inventory location, but an appreciation of how that location was being used in the greater inventory would be difficult to know. Similar to the RBC wastage focus group, it was surprising that within one hour the Hemosafe inventory focus group was able to achieve multiple process improvement guidelines, which fulfilled the goal of the PMMF [48]. The participants in this group were essentially the same as in the RBC wastage focus group with the exception that there was no other physician participant apart from the moderator (the researcher), which may have influenced the clinical breadth of guideline development, as there would have been limited medical perspective. Nonetheless, the first guideline for process improvement was not directly visualized on the process map, but may have been prompted by the overall process-based intent of the exercise as noted in the RBC wastage focus group. This directly related to the need to simplify the process of accessing blood by technologists and nursing staff. As the process information was unavailable since it was internal to the Hemosafe™ rather than in the LIS, there was a significant effort devoted to working around this shortcoming by engaging PInG resources to improve the unidirectional interface and to engage the manufacturer regarding a consultant assessment and the potential for a bidirectional interface. Another important process guideline that required improvement related to data integrity whereby a limited dictionary of codes was being used for training and live purposes. Due to the lack of training codes, existing live codes were being used improperly. These coding irregularities in both cases would likely have not been discovered using value-based methods, but were only discovered when the coding was deemed to be out of place by the SME participants when they realised that the codes followed an incorrect pathway in the RBC inventory process. The majority of the comments were overwhelmingly process-based, rather than value based because the value-based analysis was limited to very small numbers and there was already an existing understanding that the fridge was not being used. Any value-based discussion be fruitless in solving the questions of 'why' and 'in what way' that fridge was being used. Thematic free-coding also demonstrated that the utility of the process mining exercise allowed the participants to confirm the processes that were already known to them. The manually-generated process map, using the building blocks, was qualitatively simplistic without any inventory looping arising from the Hemosafe inventory location. This highlights the potential difficulty of even performing a manual audit, as it would be difficult to know which parts of the process to observe in order to do the audit correctly, especially if those parts of the

process are not apparent to the auditor. Process mining therefore could be used to inform a manual auditing process, suggesting places for observation, as it would highlight complexities of the process map that would not be apparent, even to a group of SMEs cooperating to manually generate a process map. While there is no literature on how to perform process compliance studies in the process mining or transfusion literature, this qualitative approach could represent one way of doing that by describing differences in manually- and process mining-generated maps.

In both focus groups, the commonality of incorrect coding is highlighted as a guideline for process improvement. While it may seem trivial, the proper coding of processes in the information system affects the downstream interpretation of KPIs, the very measures that global transfusion services rely upon to evaluate their inventory performance. This raises a greater philosophical question of how well we understand transfusion inventories. The phenomenon of 'garbage in, garbage out' is commonly understood. This research proposes that there is a lower-cost way of identifying the existence and amount of some incorrect coding within in supplier and hospital information systems. In the day-to-day practice of transfusion services globally, there is no assurance that completely correct codes are used as inputs for calculating KPIs. Therefore, it may be difficult to perform accurate benchmarking between institutions and define avoidable wastage events as noted in Devine *et al*. [2]. This is even more important when dealing with very small absolute and relative numbers, as is the case for wastage events in transfusion inventories, and in the case of inventory that is rarely used from a certain location. By not using process mining to correctly identify incorrect processes within a subset of these wastage reasons, benchmarking and linking of information systems may not have as large of a benefit as anticipated and/or may lead to improper conclusions. The research also demonstrates the ability of process mining to augment the auditing process. This may allow the BTS to monitor and ultimately troubleshoot problems in areas where rare events can occur at any time of the day, spread over any location in the blood transfusion inventory geographically, without the need for significant human resources to observe these deviations when they happen. Human resources would be required at different phases of the auditing process, such as the definition of attributes, data extraction, preprocessing, and exploration

phases, and interpretation of process mining-generated maps in both process unaware or aware systems.

There were several assumptions which were foundational for this research to occur. The first assumption has been validated, in that inventories can be thought of as processes, such as initially described by Katsaliaki and Brailsford [36] or more recently described as a 'lifecycle' Cheng *et al.* [37]. Of those two papers, the only paper that described a real inventory (and not a simulation) was the latter paper. The research extends the concept of the lifecycle mentioned in Cheng *et al.*, whereby RBCs are described with KPIs, such as age at receipt and discard rates, but were also acknowledged to have process-like characteristic, such as time spent outside the BTS, or number of times returned to inventory. The authors of this paper may have been unaware of the documentation of these process activities in the laboratory information system at the time. Given the articulation of the process-based paradigm applied to transfusion inventory management, the description of the RBC inventory can now be studied as a process. Moreover, this research has demonstrated that the activities associated with these processes are captured by the LIS, and can be retrieved and visualized/analysed using process mining techniques. The second assumption that has been validated is that this data can be retrieved from a system that is not process-aware [47], and can be done using, at minimum, level 3 quality data as the process log of the RBC unit could be pieced together from the data located in different tables. By examining the table structure of the LIS from which the data was derived, these data elements are the most granular elements describing an RBC unit, some of which is represented in Appendix B: Laboratory Information Systems, in Table 36 and Figure 55. This research is the second paper that references process mining in the context of an inventory. There are some key differences of this research versus the research by Paszkiewicz [218], in that their data originates from a WFM without modification and conformance testing is done by comparing a process map of a mattress life-cycle generated through process mining with a model *du jure*. This research used data originating from a process unaware system and used process mining for process discovery. The conformance phase of the research examined how a manually-generated process map differed from the process mining-generated map. Moreover, the research highlights that by using process mining to augment the traditional value-based (KPI-based) view on inventories, it is possible to further obtain a deeper understanding of them,

though there is a requirement to ensure that the data already exists in the LIS, and that there are minimal artifacts introduced if the data is manipulated from a non-process aware system. As an extension of this research, it is likely that this process information is captured in high granularity in other transfusion information systems globally.

One area where this research can be extended is in the area of benchmarking, as benchmarking represents a functional use and application of KPI-based methodologies. It has already been highlighted that benchmarking relies on accurate KPIs, and this research has demonstrated that not all codes entered in the LIS are entirely correct. As noted, if a KPI is based on an improperly coded wastage reason, then any benchmarking using that KPI is invalid. Moreover, in very uncommon wastage events, a very small shift can yield very large percentage changes. Process mining can also possibly change the philosophy of benchmarking. Currently, benchmarking is done using summary variables such as outdate percentages and other variables that describe hospital distance from blood supplier [20, 23]. One of the goals of benchmarking is to identify and communicate best practices for others to potentially emulate [20, 22, 23]. Devine *et al.* also indicate that "there is a need for constructive dialogue amongst blood suppliers and users of blood products to identify a set of best/better practices that can be implemented" [2], implying that benchmarking will be an important tool in that process. One assumption with regards to identification and application of 'best/better practices' is that the adoption of those practices is uniform and complete in those best practice sites, and that there are no other confounding reasons as to why sites have high performance, but this cannot be guaranteed. In addition, there is no guarantee that compliance to policies will be uniform in those centres adopting those identified best/better practices, as there may also be confounding factors preventing adoption or possibly enhancing adoption. Therefore, the concept of benchmarking potentially could be a non-equivalent comparison. The potential use for process mining in this instance would be to examine high-performing institutions and decompose their inventory handling into sub-processes that could be further benchmarked or studied for compliance. On the application side, process mining could be used to determine whether or not there are changes before and after the implementation of a policy on inventory sub-processes. The use of process mining could be used to complement conventional KPI-based benchmarking to determine whether or not institutions are truly equivalent. The potential use for process mining in this instance would

be to examine the effects of a policy change on decomposed sub-processes of an institution, and determine an institution's adherence to the policy. Although it was not studied in this research, process mining has the ability to decompose the inventory process into multiple sub-processes which can each be analysed. There is no transfusion literature to benchmark sub-processes, either inter- or intra-institutionally, such as product receipt time, crossmatching times, and recycling times from smaller peripheral sites.

One area in which this sub-process scale benchmarking can occur, and may be theoretically helpful, is in identifying areas where blood products age unnecessarily, as identified by Cheng *et al.* [37]. In the literature, increasing crossmatch times have been modelled and simulated to increase outdate rates [18, 30]. The mechanism of this is unknown, but it is likely that increased crossmatch times cause a RBC unit to unnecessarily age, by the amount of the crossmatch expiry time, prior to being returned to the inventory. This pushes the RBC unit much closer to its expiry date. Therefore, it may be that any delay or shunting of a RBC unit's duration of availability could potentially contribute to increased outdating. These latencies have not been studied in the literature, because until now, they have largely been an unknown and unquantifiable variable, but they can be mitigated by internal procedures and policies of the BTS once the root cause is known. One other issue with KPI-based measures, that were previously discussed (Table 6), is the variability of mathematical equations by which inventory wastage and size can be expressed. Although this is speculative, if inventory can be understood as a process, and that process can be made most efficient, it should follow that a possible benchmarking variable could be the time that a unit spends in the transfusion service. This is not currently used as a KPI. Time stamps are inherent to the process log and they are invisibly and constantly collected with negligible overhead, therefore their visualization through process mining could represent another KPI by which transfusion services are compared with each other. Although time does not have any mathematical manipulation associated with it, similar to an absolute inventory size, the standardized definition of the various signposts (activities) may need to be discussed.

Another aspect of benchmarking that could benefit from process mining is in the area of inventory size, and thus turnover. Currently, transfusion services globally use various inventory indices to describe the amount of inventory available (i.e. days on hand, ISI). It is known that an inventory that is too large will have a high ISI and increased number of days on hand.

Conceptually, units sitting in those large inventories will not 'circulate' much and this could be measured by latency between a unit's availability and its final disposition. Cheng *et al.* [37] attempted to measure this latency but it was not thought of in the context of benchmarking. Could one use a simpler time variable to describe the relative flow (i.e. flux) of RBC units through an inventory? If generalised globally, one could benchmark the flux of a high-performing inventory versus that of a marginally-performing inventory, with the goal of mitigating wastage. Perhaps high-flux inventories (i.e. ones with units that have low latency times between events) are the ones that have better performing outdate rates. Conversely, there may be a minimum floor that cannot be penetrated, otherwise there would be shortages. There is no literature in this area, but this latency concept could be an interesting phenomenon to explore and potentially benchmark institutions with each other. Ideally, if transfusion services globally collect this type of timestamp data to begin with, then this should be a concept that is easily generalizable. By using process-based techniques on a traditionally value-based benchmarking paradigm, it may be possible to achieve different ways of expressing inventory performance, determining the equivalency of transfusion inventories, and decomposing complicated processes into simpler ones.

There are other areas where the results of this research could be applied. In the area of Lean Sigma process improvement techniques applied to blood transfusion, where the goal is to '*define, measure, analyse, improve, and control*' as described by Heitmiller *et al.* [38], a potential explanation of their reduction in RBC wastage is the Hawthorne effect whereby process performance metrics are improved by the mere observation of the subjects in question. Although there is not enough detail on what is captured in their information systems, there may have been a possibility to use process mining in the measurement and control phases of their study. The major benefit of process mining is that no direct observation is required.

With increased cost pressures and diversity of products, the need to maximise performance from transfusion inventories globally is a trend that will persist [2]. While the transfusion inventory management literature is fairly mature, and best practices have already been defined for the past half-century, Devine *et al.* highlight a paradox whereby that is not the case:

'while there is clearly no singular right way to handle inventory management and the choices involved therein, there is a need for constructive dialogue amongst blood suppliers and users of blood products to identify a set of best/better practices that can be implemented without prohibitory levels of cost or complexity to increase the availability, safety and quality of blood products for the patients who need them while at the same time increasing the efficiency of this important aspect of blood transfusion management. [2]'.

This quote highlights the need for communication and identification of best practices to improve efficiency, without significantly increase cost or jeopardizing the integrity of the system. Despite the matured body of literature since the 1960s, the only practice that seems to be generalized and widely adopted according to Devine *et al.*, is the concept of FIFO [2]. One can speculate that there may be potential reasons for this paradox, including the expository rather than practical application of benchmarking techniques, the unavailability of methods to directly assess inventory performance, including the information gap noted previously.

This research demonstrates it is possible to reconstruct the 'process log' of RBC inventory transactions from a process-unaware laboratory information system and that further insight on inventories that previously characterised using KPIs is possible without prohibitory complexity or costs. This research suggests that process mining is an ideal tool for application to transfusion inventory management data. It is a technique that describes day-to-day real inventories and the process map that is generated is easily understandable by end-users and drives meaningful discussion and inventory introspection. Process mining has the potential to define new KPIs, and could be used to accurately compare processes and sub-processes between transfusion inventories. The application of process mining could significantly shift the paradigm of transfusion inventory management, similar to the financial auditing environment, and could provide the ability for best practices and better practices to be discovered, communicated, and disseminated, thereby maximizing the efficiency of inventories globally.

5.2.3 Limitations of Research

The process map is dependent on appropriate and accurate coding by the technologists who enter in error codes and reasons. There were codes which were found to be incorrect due to the process mining pathway not making sense, but there is a possibility for silent errors to occur where the code is incorrectly inputted but does not generate a significantly incorrect process pathway. In this research, it would limit the ability to generate valid and additional process improvement guidelines. It is noted that this limitation would also occur if one was to analyze the KPIs associated with that process.

As noted in the literature, if the data does not exist within the dataset, such as in the case of manual data capture or if data is captured elsewhere and not incorporated into the dataset, then it is not possible to analyze portions of the process. The dataset was unable to capture details associated with the transfusion (i.e. nurse, perfusionist), porter, and courier resources. Data from CBS was unavailable. Data from the Hemosafe™ middleware (BloodTrax™) was not available, and may have been helpful in generating additional process improvement guidelines.

A limitation to this research is that the data originated from a process unaware system. There was missing data with regards to STATUS attributes, which required the native STATUS attributes to be multiply mapped to the 'activity' attribute within Disco. As a result, there was generation of artifacts, but these artifacts could be explained. Ideally, those missing status attributes should be found and integrated into the data set. Also, by multiple mapping of the attributes, excess process granularity may have been generated in the dataset versus the number of actual process steps. For instance, if five RBC units went to a particular location, but were returned, the presence of two crossmatch reasons would split those five RBC units into two groups, rather than one, thus generating ten variants.

A limitation of the focus group was that it was not possible to filter out and attribute contributions of people within that group. The results of the comments, and therefore the thematic coding may have been influenced by a minority of participants in the focus group who

were outspoken. Conversely, this may have not been the setting for less outspoken members to normally express their thoughts and comments, and this would have also biased the results. It is also unknown if the process map information was differentially interpreted by members having different roles or experience levels in blood transfusion services. The number of guidelines for process improvement may have been affected by this.

A limitation of the research may have been the time difference between the time of focus group and the date period for which the query was run. It was attempted to make this as short as possible so that there was little difficulty recalling any inventory processes that may have occurred, as the data would have been very recent in nature. This would have affected the comments and thus the thematic coding, as well as the number of guidelines for process improvement.

Another limitation of the conformance testing aspect of the research is that participants of the focus group should have been allowed to meet and construct a process map *de jure* of the Hemosafe™ fridge process as it is supposed to happen, and how they think it is happening without any assistance. As it stood in the research, participants were given already constructed building blocks. This would have allowed for a true comparison between the various perspectives on the process map, but unfortunately, the question was not asked in this study.

Another limitation of the study involves the Atlas Ti™ coding coming from an unblinded SME. It is unknown what contribution this may have in the results of the thematic coding; however, the effect of this bias this is counterbalanced by the fact that overall concrete process improvement were identified as a result of the process mining exercise, which justified the overall methodology. Any bias in the thematic coding would only affect the academic portion of the research.

Another possible limitation of the study could be that the moderator (the researcher) is a SME in blood transfusion services, particularly the inventory, which may increase the frequency of leading questions and which may unknowingly have biased conversation, which may also have caused some areas to be explored either too deeply or superficially. This may also result in leading questions or unknown biases which may have affected the qualitative outcomes of the research.

One potential limitation of the research is in its generalizability, in that not all LISs are configured or potentially may capture this data in high granularity. The research uses data from Cerner PathNet Millennium™, which is a modern LIS. To answer the question of whether data of sufficient resolution might exist in an older information system, a process map was generated using Meditech data from an implementation at another hospital (Appendix B: Laboratory Information Systems). Meditech Magic™ has been implemented at that hospital since 1989, and it is likely that most transfusion services will use systems or newer ones to capture transfusion inventory process data. A process map mined from Meditech Magic™ is shown (Figure 54,Appendix B: Laboratory Information Systems) to demonstrate that there is potential for generalisability, though it is not specifically proven in this research. It is also noted that these LISs are likely process-unaware, and thus it behooves the implementer of the information system to deliberately codify the blood transfusion inventory process within the system in order that process maps can be generated from this data. Codification would also include using human readable variables, an emphasis on discrete coding rather than free textual fields, and appropriate activity-level details with the goal of constructing a verb-object [48]. This codification would have to be matched by the physical processes, such as scanning and documentation into the LIS at different checkpoints along the process.

5.2.4 Recommendations for Further Research

In the area of general transfusion inventory, further research would validate the use of this methodology at another institution to visualize their total inventory flow. In particular, it would

validate that this methodology works in an environment where the moderator is a process expert but not necessarily the SME of that environment with the ability to comment on the institution's own process map. As an extension of this, further research would be to compare the process maps of different institutions that feature the same topology (i.e. hub and spoke) for benchmarking purposes. Another area of research is to decompose the practices of transfusion services having the best inventory management KPIs, and determine how similar they are in their sub-processes. The goal of this comparison study would be to determine if there are any inventory practices that can be translated to other transfusion services. As an extension of that, in the transfusion services which implement those practices, are there meaningful changes in their sub-processes, either in performance or frequency, that can be observed?

In the area of benchmarking, further research could determine if it is possible to group institutions according to how sub-processes in their inventory are functioning, in the context of other variables, such as transfusion rates. This would require multicentre participation. Further research could examine if new variables can be introduced into the benchmarking lexicon. The ideal performance times of sub-processes such as RBC unit confirmation, RBC unit return, and RBC unit latency can be defined. Perhaps a new flux variable could be defined to measure the extent of RBC flow through an inventory. This benchmarking could be done within an institution and between institutions globally.

In the area of policy conformance, further research could determine the extent of policy adherence of transfusion services with emphasis on policies that could directly affect the transfusion inventory, such as recycling, crossmatching, ordering and receipt policies using process mining. Further research could involve determining how various storage environments impact the RBC flow pathway. In particular, if a policy change intervention has occurred, does the process map display that after a return to steady state? Further research could also include examining using process mining on different inventory products, such as platelets, plasma, and plasma protein products.

Further research can occur with process mining in other areas of transfusion services outside of the inventory. In the area of the document management system (i.e. Paradigm) or an enterprise wide system (i.e. SAP), there is potential research to examine the performance of the document approval process. Further research to look for areas of potential efficiency gains using process mining can occur in the laboratory testing processes, specimen receipt processes, and other transfusion laboratory processes.

In the area of donor services, it is known that CBS is currently implementing a paperless front-end for donors. It would be informative to see how well the process is performing, and whether or not there are any process deviations from what they are expecting. It is the understanding that this paperless workflow system has a high likelihood to be process aware, and would represent a rich data source for process mining.

In the area of the healthcare system, it may be possible to use process mining to determine patient flow through different services of the hospital, to potentially determine bottlenecks in the system. Likewise, the performance of laboratory testing, transportation processes, troubleshooting of automated laboratory instruments, and other processes that are documented within the LIS can also be process mined.

Although this is not specifically a research need, with the increased rationalization of healthcare information systems, with initiatives such as One Person, One Record (Nova Scotia), further research could involve the approach to codifying variables and fields in the system so that process mining can potentially occur and meaningfully improve health system performance.

Appendix A: Literature Search Strategies

For the transfusion medicine inventory papers:

The initial literature search involved using the PubMed database (www.pubmed.org) with the following search terms: ("Inventories, Hospital"[Mesh] AND ("blood transfusion"[MeSH Terms] OR ("blood"[tiab] AND "transfusion"[tiab]) OR "blood transfusion"[tiab] OR "blood products"[tiab] OR "transfusion products"[tiab] OR blood[tiab])) OR (("inventory"[All Fields] OR "equipment and supplies"[MeSH Terms] OR ("equipment"[All Fields] AND "supplies"[All Fields]) OR "equipment and supplies"[All Fields]) AND ("Blood Banks/economics"[Mesh] OR "Blood Banks/instrumentation"[Mesh] OR "Blood Banks/methods"[Mesh] OR "Blood Banks/organization and administration"[Mesh] OR "Blood Banks/standards"[Mesh] OR "Blood Banks/supply and distribution"[Mesh] OR "Blood Banks/trends"[Mesh] OR "Blood Banks/utilization"[Mesh])). English language articles were only considered, and searching was done during the writing of the manuscript.

A comprehensive recent review paper on transfusion inventory management methodologies was done by Belien *et al*[25], and this was used as the input paper for the Web of Science (WOS) database. The articles this paper cited were then exported into Endnote X4, and the duplicate citations were removed. The articles that remained were added to the articles obtained from the initial literature search. There were also papers that could not be obtained due to lack of access.

For the process mining papers, this was the search strategy:

The literature search involved a combination of these various strategies, as process mining is less so in the literature realm, but more so in the practical realm. A topic search for ‘process mining’ and ‘process mining applications’ was done using Web of Science. With WOS, process mining articles were found and the articles and their citations were used to find other articles. Also, popups of ‘other users also viewed these articles’ or recommended articles. The ‘cited’ and ‘cited by’ fields on WOS were examined for articles on any article which was located, and as new articles were discovered, the same process would repeat.

Appendix B: Laboratory Information Systems

Meditech Magic™

Meditech Magic™ (Westwood, Massachusetts; meditech.com) has been operating at the Izaak Walton Killam Hospital since 1989 and is still currently in operation as the institution-wide hospital information system and the laboratory information system. Data is structured in a branching tree hierarchy, as highlighted in Figure 49.

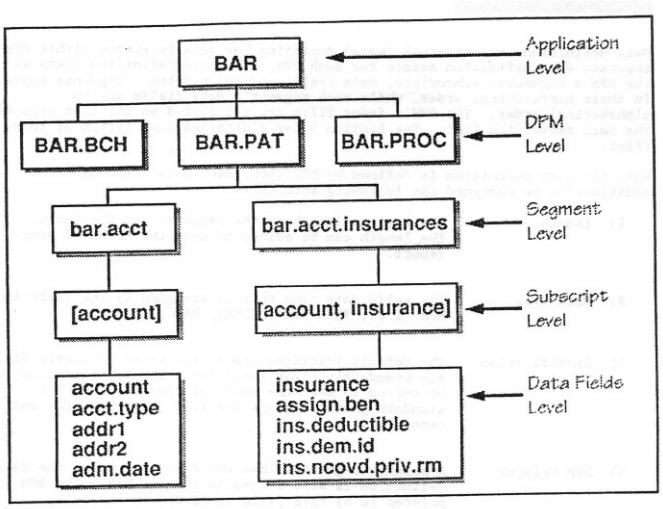


Figure 49: Data structure of the Meditech Magic™ laboratory information system

The red cell unit data is within 'lab.b.unit.main' (third line from the top in Figure 50; Data Definition List queried on March 10, 2012) and this is the highest granularity data that can be derived from the laboratory information system.

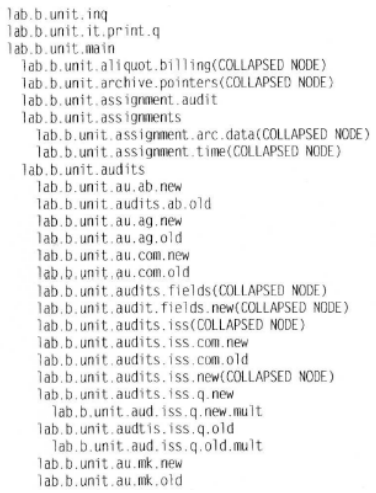
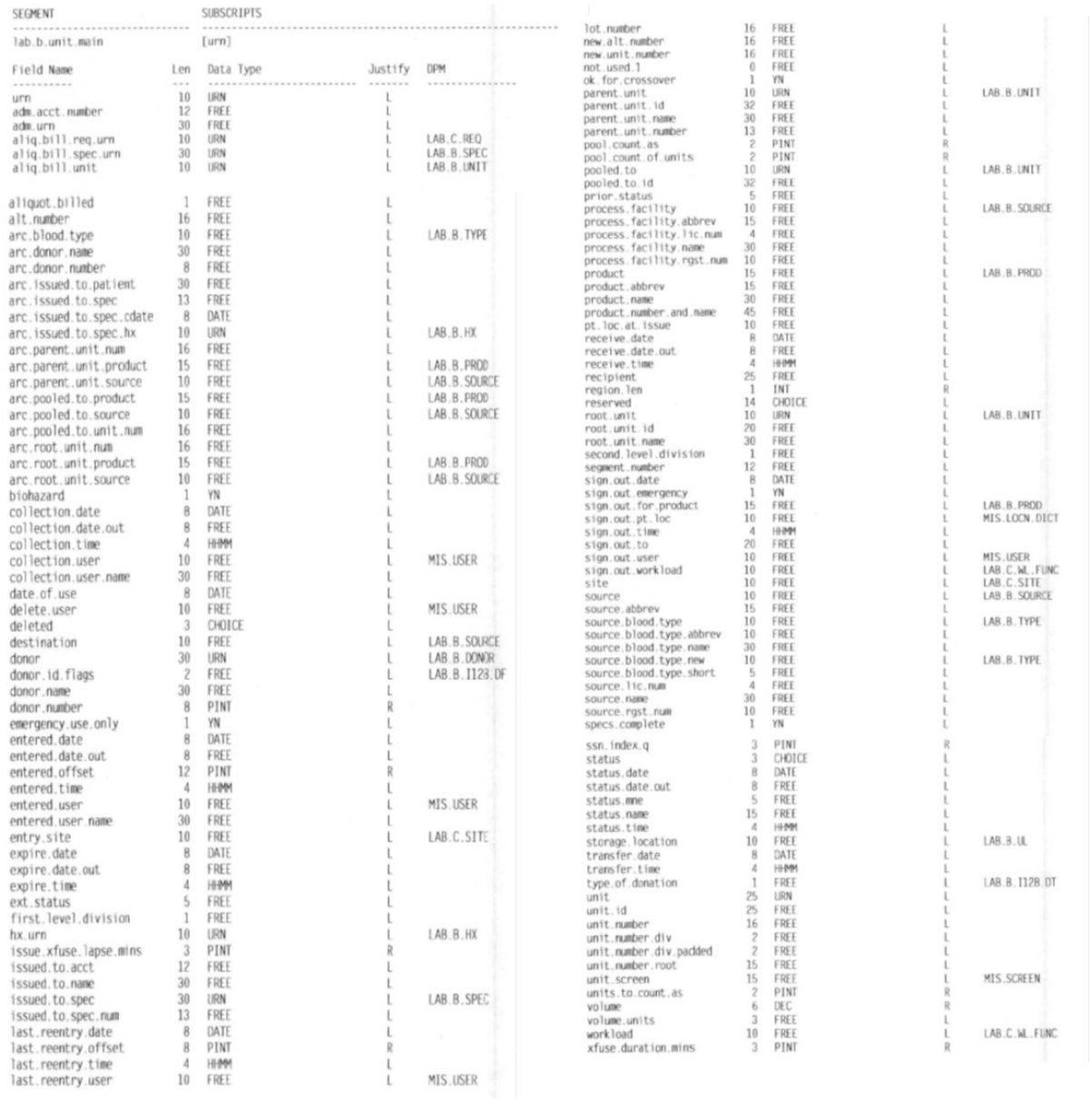


Figure 50: screenshot of segment architecture of blood bank unit data (Meditech Magic™)

Within the *unit.main* branch are the data fields and data types that pertain to all possible blood products, with red cell units being a subset of the bigger laboratory dataset (Figure 51). Within Figure 51, a product is described with multiple general attributes (i.e. location, blood type, expiry dates, issuance and pooling activities), timestamps (collection times, expiry, status, transfer, signout), and there are also fields for users. Specific unit-related information, including barcode and final disposition are located in Figure 52 and Figure 53. A data type table is shown in Table 35.



SEGMENT	SUBSCRIPTS
lab.b.unit.main	[urn]

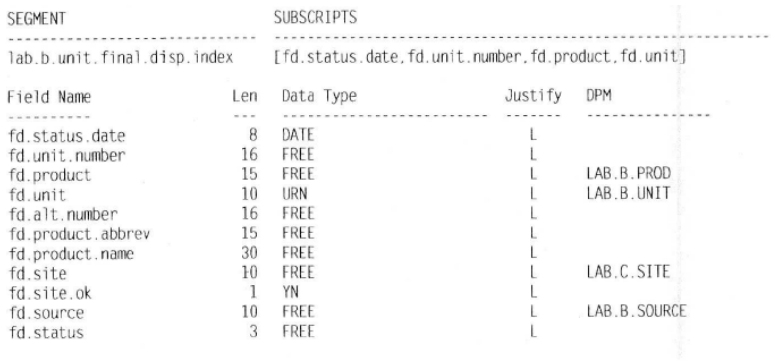
Field Name	Len	Data Type	Justify	DPM
urn	10	URN	L	
adm.acct.number	12	FREE	L	
adm.urn	30	FREE	L	
aliq.bill.req.urn	10	URN	L	LAB.C.REQ
aliq.bill.spec.urn	30	URN	L	LAB.B.SPEC
aliq.bill.unit	10	URN	L	LAB.B.UNIT
aliquot.billed	1	FREE	L	
alt.number	16	FREE	L	
arc.blood.type	10	FREE	L	LAB.B.TYPE
arc.donor.name	30	FREE	L	
arc.donor.number	8	FREE	L	
arc.issued.to.patient	30	FREE	L	
arc.issued.to.spec	13	FREE	L	
arc.issued.to.spec.cdate	8	DATE	L	
arc.issued.to.spec.hx	10	URN	L	LAB.B.HX
arc.parent.unit.num	16	FREE	L	
arc.parent.unit.product	15	FREE	L	LAB.B.PROD
arc.parent.unit.source	10	FREE	L	LAB.B.SOURCE
arc.pooled.to.product	15	FREE	L	LAB.B.PROD
arc.pooled.to.source	10	FREE	L	LAB.B.SOURCE
arc.pooled.to.unit.num	16	FREE	L	
arc.root.unit.num	16	FREE	L	
arc.root.unit.product	15	FREE	L	LAB.B.PROD
arc.root.unit.source	10	FREE	L	LAB.B.SOURCE
biohazard	1	YN	L	
collection.date	8	DATE	L	
collection.date.out	8	FREE	L	
collection.time	4	HHMM	L	
collection.user	10	FREE	L	MIS.USER
collection.user.name	30	FREE	L	
date.of.use	8	DATE	L	
delete.user	10	FREE	L	MIS.USER
deleted	3	CHOICE	L	
destination	10	FREE	L	LAB.B.SOURCE
donor	30	URN	L	LAB.B.DONOR
donor.id.flags	2	FREE	L	LAB.B.I123.DF
donor.name	30	FREE	L	
donor.number	8	PINT	R	
emergency.use.only	1	YN	L	
entered.date	8	DATE	L	
entered.date.out	8	FREE	L	
entered.offset	12	PINT	R	
entered.time	4	HHMM	L	
entered.user	10	FREE	L	MIS.USER
entered.user.name	30	FREE	L	
entry.site	10	FREE	L	LAB.C.SITE
expire.date	8	DATE	L	
expire.date.out	8	FREE	L	
expire.time	4	HHMM	L	
ext.status	5	FREE	L	
first.level.division	1	FREE	L	
hx.urn	10	URN	L	LAB.B.HX
issue.xfuse.lapse.mins	3	PINT	R	
issued.to.acct	12	FREE	L	
issued.to.name	30	FREE	L	
issued.to.spec	30	URN	L	LAB.B.SPEC
issued.to.spec.num	13	FREE	L	
last.reentry.date	8	DATE	L	
last.reentry.offset	8	PINT	R	
last.reentry.time	4	HHMM	L	
last.reentry.user	10	FREE	L	MIS.USER
lot.number	16	FREE	L	
new.alt.number	16	FREE	L	
new.unit.number	16	FREE	L	
not.used.1	0	FREE	L	
ok.for.crossover	1	YN	L	
parent.unit	10	URN	L	LAB.B.UNIT
parent.unit.id	32	FREE	L	
parent.unit.name	30	FREE	L	
parent.unit.number	13	FREE	L	
pool.count.as	2	PINT	R	
pool.count.of.units	2	PINT	R	
pooled.to	10	URN	L	LAB.B.UNIT
pooled.to.id	32	FREE	L	
prior.status	5	FREE	L	
process.facility	10	FREE	L	LAB.B.SOURCE
process.facility.abbrev	15	FREE	L	
process.facility.lic.num	4	FREE	L	
process.facility.name	30	FREE	L	
process.facility.rgst.num	10	FREE	L	
product	15	FREE	L	LAB.B.PROD
product.abbrev	15	FREE	L	
product.name	30	FREE	L	
product.number.and.name	45	FREE	L	
pt.loc.at.issue	10	FREE	L	
receive.date	8	DATE	L	
receive.date.out	8	FREE	L	
receive.time	4	HHMM	L	
recipient	25	FREE	L	
region.len	1	INT	R	
reserved	14	CHOICE	L	
root.unit	10	URN	L	LAB.B.UNIT
root.unit.id	20	FREE	L	
root.unit.name	30	FREE	L	
second.level.division	1	FREE	L	
segment.number	12	FREE	L	
sign.out.date	8	DATE	L	
sign.out.emergency	1	YN	L	
sign.out.for.product	15	FREE	L	LAB.B.PROD
sign.out.pt.loc	10	FREE	L	MIS.LOCN.DICT
sign.out.time	4	HHMM	L	
sign.out.to	20	FREE	L	
sign.out.user	10	FREE	L	MIS.USER
sign.out.workload	10	FREE	L	LAB.C.WL.FUNC
site	10	FREE	L	LAB.C.SITE
source	10	FREE	L	LAB.B.SOURCE
source.abbrev	15	FREE	L	
source.blood.type	10	FREE	L	LAB.B.TYPE
source.blood.type.abbrev	10	FREE	L	
source.blood.type.name	30	FREE	L	
source.blood.type.new	10	FREE	L	LAB.B.TYPE
source.blood.type.short	5	FREE	L	
source.lic.num	4	FREE	L	
source.name	30	FREE	L	
source.rgst.num	10	FREE	L	
specs.complete	1	YN	L	
ssn.index.q	3	PINT	R	
status	3	CHOICE	L	
status.date	8	DATE	L	
status.date.out	8	FREE	L	
status.mne	5	FREE	L	
status.name	15	FREE	L	
status.time	4	HHMM	L	
storage.location	10	FREE	L	LAB.B.UL
transfer.date	8	DATE	L	
transfer.time	4	HHMM	L	
type.of.donation	1	FREE	L	LAB.B.I128.DT
unit	25	URN	L	
unit.id	25	FREE	L	
unit.number	16	FREE	L	
unit.number.div	2	FREE	L	
unit.number.div.padded	2	FREE	L	
unit.number.root	15	FREE	L	
unit.screen	15	FREE	L	MIS.SCREEN
units.to.count.as	2	PINT	R	
volume	6	DEC	R	
volume.units	3	FREE	L	
workload	10	FREE	L	LAB.C.WL.FUNC
xfuse.duration.mins	3	PINT	R	

Figure 51: lab.b.unit.main (main table of red cell unit) fields in Meditech Magic™

SEGMENT	SUBSCRIPTS
lab.b.unit.ee.barcode.info	[ee.bc.info.q]

Field Name	Len	Data Type	Justify	DPM
ee.bc.info.q	2	PINT	R	
ee.bc.info.line	80	FREE	L	

Figure 52: lab.b.unit.ee.barcode.info [edit and entry of unit barcode] fields in Meditech Magic™



SEGMENT: lab.b.unit.final.disp.index
SUBSCRIPTS: [fd.status.date,fd.unit.number,fd.product,fd.unit]

Field Name	Len	Data Type	Justify	DPM
fd.status.date	8	DATE	L	
fd.unit.number	16	FREE	L	
fd.product	15	FREE	L	LAB.B.PROD
fd.unit	10	URN	L	LAB.B.UNIT
fd.alt.number	16	FREE	L	
fd.product.abbrev	15	FREE	L	
fd.product.name	30	FREE	L	
fd.site	10	FREE	L	LAB.C.SITE
fd.site.ok	1	YN	L	
fd.source	10	FREE	L	LAB.B.SOURCE
fd.status	3	FREE	L	

Figure 53: lab.b.unit.final.disp.index [Final Disposition of Red Cell Unit] Fields in Meditech Magic™

Table 35: Table of Data Type and Description for Data in Meditech Magic™

Data Type	Data Type Description
FREE	Freetext
URN	Unique reference number
DATE	Date
YN	Yes/No
CHOICE	A list of choices is displayed, with selection of one from the list.
INT	Integer
PINT	Positive integer
HHMM	Hour:Minute (2 digit hour: 2 digit minute)
DEC	Decimal places

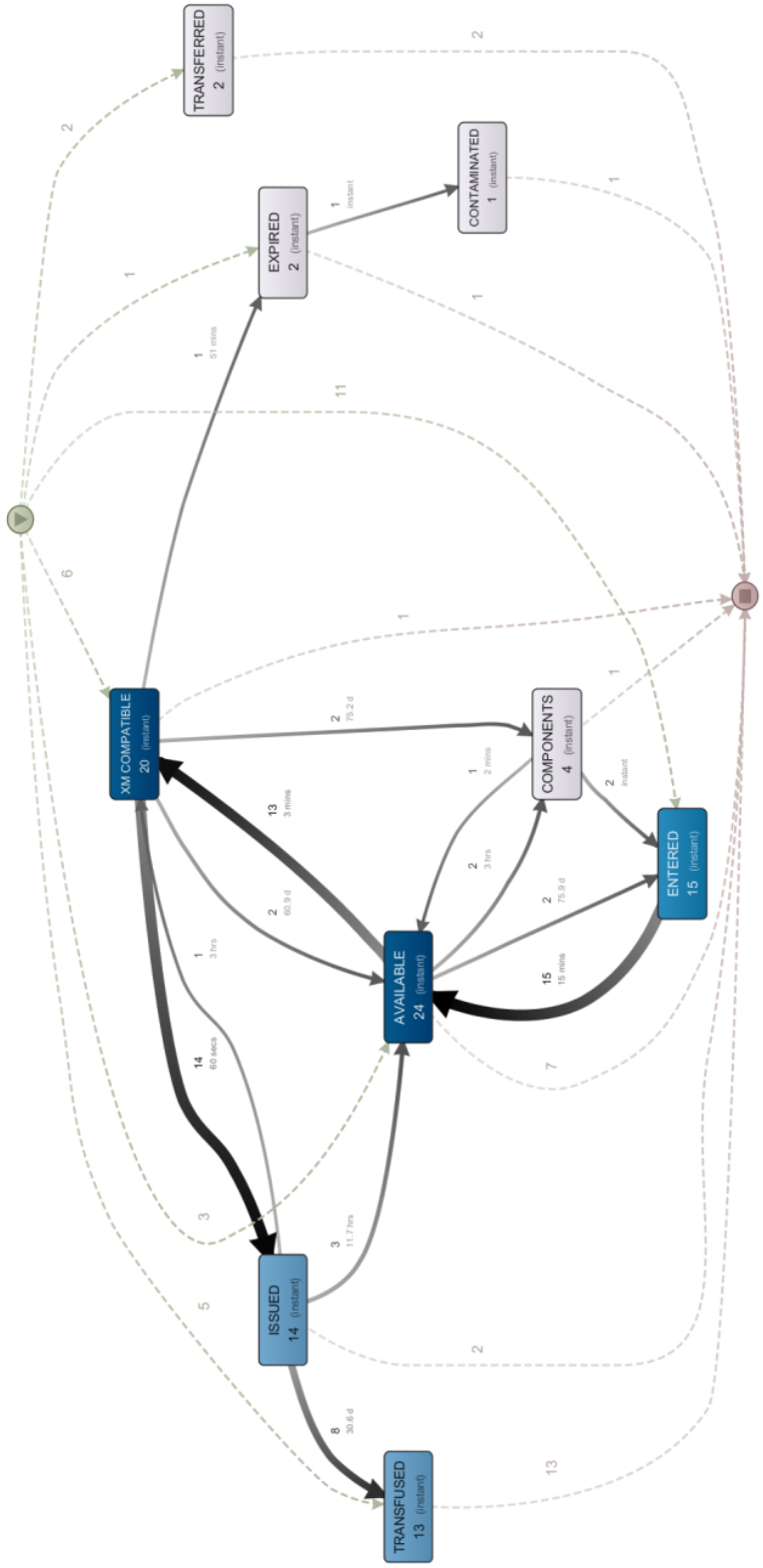


Figure 54: Process Map generated from test data from Meditech Magic™ (June 2015)

Cerner Millennium™

Cerner Millennium™ (Cerner Corporation; Kansas City, USA) has been operational since May 2009 at the Blood Transfusion Services at the Nova Scotia Health Authority, Central Zone. Data is stored in modern database architecture within Oracle tables, and is queried using Cerner Command Language, a Cerner Corporation variation of commonly used SQL query language.

A screenshot of the 'Blood Product' Table is highlighted in Figure 55, with data types shown in Table 36. This table is an example of many tables within the laboratory information system, with selected tables being depicted on the left hand column. Figure 55 shows different attributes of a blood product which are captured, including the product itself, the events that pertain to it (event, disposition events, return events, transfers), personnel, and institutional location attributes also. Status date and time of status, other attributes (i.e. unit volumes, autologous states, supplier codes) are also recorded and tracked (data table not shown). The data is located within each field and corresponds to a consistent data type. For instance, ACTIVE_STATUS_DT_TM has a data type of DQ8 associated with it, which the data type description in Table 36 indicates that it is a date and time in raw format.

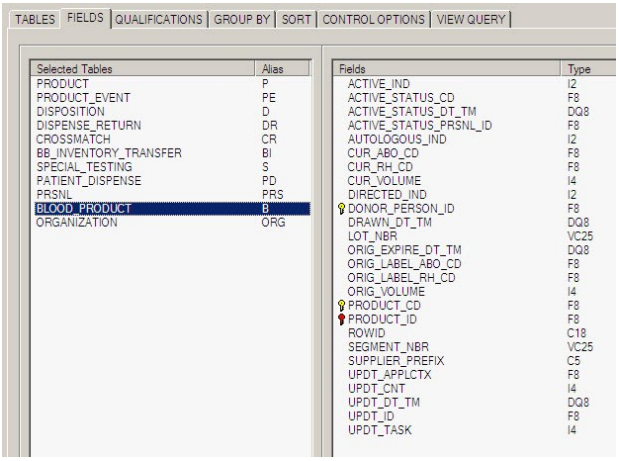


Figure 55:Blood Product Table Fields and Data Types (Cerner Millennium™)

Table 36: Table of Data Type and Data Type Descriptions for Data in Cerner Millennium™ Tables

Data Type	Data type Description
C[#]	Fixed length character string, specified by the #
DM12	12 byte used for timestamp
DQ8	Quadword used for datetime: series of raw numbers used to store dates, must be translated for format as a readable date
F4	Single precision float: positive or negative numbers with range of 0.29*10-38 to 1.7*10+38 and 7 digit precision
F8	Double precision float: positive or negative numbers with range of 0.29*10-38 to 1.7*10+38 and 16 digit precision
[U]I1	Unsigned/signed 1-byte integer: whole numbers, specified by prefixing with "U" for unsigned or leaving it as "I" for signed signed value limits: -128 to +127 unsigned value limits: 0 to 255
[U]I2	Unsigned/signed 2-byte integer: whole numbers, specified by prefixing with "U" for unsigned or leaving it as "I" for signed signed value limits: -32,768 to +32,767 unsigned value limits: 0 to 65,535
[U]I4	Unsigned/signed 4-byte integer: whole numbers, specified by prefixing with "U" for unsigned or leaving it as "I" for signed signed value limits: -2,147,483,648 to +2,147,483,647 unsigned value limits: 0 to 4294967295
VC	Variable length character string. The default maximum is 524,288 characters.

Bibliography

1. Blake, J.T., *On the use of Operational Research for managing platelet inventory and ordering.* Transfusion, 2009. **49**(3): p. 396-401.
2. Devine, D.V., et al., *Inventory management.* Vox sanguinis, 2010. **98**(3 Pt 1): p. e295-363.
3. Shander, A., et al., *Estimating the cost of blood: past, present, and future directions.* Best practice & research. Clinical anaesthesiology, 2007. **21**(2): p. 271-89.
4. Shander, A., et al., *Activity-based costs of blood transfusions in surgical patients at four hospitals.* Transfusion, 2010. **50**(4): p. 753-765.
5. Aledort, L.M., et al., *The cost of blood: Multidisciplinary consensus conference for a standard methodology.* Transfusion medicine reviews, 2005. **19**(1): p. 66-78.
6. Callum, J.L., O.R.B.C. Network, and O.R.B.C.N. Staff, *Bloody Easy 3: Blood Transfusions, Blood Alternatives and Transfusion Reactions: A Guide to Transfusion Medicine*2011: Ontario Regional Blood Coordinating Network.
7. Abraham, I. and D. Sun, *The cost of blood transfusion in Western Europe as estimated from six studies.* Transfusion, 2012. **52**(9): p. 1983-8.
8. Agrawal, S., et al., *Assessing the total costs of blood delivery to hospital oncology and haematology patients.* Current Medical Research and Opinion, 2006. **22**(10): p. 1903-1909.
9. Van Bellinghen, L., et al., *Comparing the provider time and costs for red blood cell transfusions in anaemia management of cancer patients using the activity-based costing (ABC) method in a French and Austrian setting.* Value in Health, 2003. **6**(6): p. 735-735.
10. Glenngard, A.H., U. Persson, and C. Soderman, *Costs associated with blood transfusions in Sweden - the societal cost of autologous, allogeneic and perioperative RBC transfusion.* Transfusion medicine, 2005. **15**(4): p. 295-306.
11. Hadjianastassiou, V.G., G. Virich, and I.A. Lennox, *Use of the blood transfusion service in total knee replacement arthroplasty. The cost implications.* Knee, 2002. **9**(2): p. 145-148.
12. Varney, S.J. and J.F. Guest, *The annual cost of blood transfusions in the UK.* Transfusion medicine, 2003. **13**(4): p. 205-218.
13. Goodnough, L.T., et al., *Medical progress - Transfusion medicine (First of two parts) - Blood transfusion.* New England Journal of Medicine, 1999. **340**(6): p. 438-447.
14. Vamvakas, E.C., *Epidemiology of red blood cell utilization.* Transfusion medicine reviews, 1996. **10**(1): p. 44-61.
15. Pealer, L.N., et al., *Transmission of West Nile Virus through Blood Transfusion in the United States in 2002.* New England Journal of Medicine, 2003. **349**(13): p. 1236-1245.
16. Saberton, P.J., et al., *Geographical variations in the correlates of blood donor turnout rates: an investigation of Canadian metropolitan areas.* International journal of health geographics, 2009. **8**: p. 56.
17. Rock, G., et al., *The supply of blood products in 10 different systems or countries.* Transfusion Science, 2000. **22**(3): p. 171-182.
18. Perera, G., et al., *Hospital Blood Inventory Practice: the factors affecting stock level and wastage.* Transfusion medicine, 2009. **19**(2): p. 99-104.
19. Novis, D.A., et al., *Quality indicators of blood utilization - Three college of American Pathologists Q-Probes studies of 12 288 404 red blood cell units in 1639 hospitals.* Archives of pathology & laboratory medicine, 2002. **126**(2): p. 150-156.

20. Heddle, N.M., et al., *Factors affecting the frequency of red blood cell outdates: an approach to establish benchmarking targets.* Transfusion, 2009. **49**(2): p. 219-226.
21. Collins, R.A., et al., *Effectiveness of Multiple Initiatives to Reduce Blood Component Wastage.* American journal of clinical pathology, 2015. **143**(3): p. 329-335.
22. Apelseth, T.O., et al., *Benchmarking: Applications to Transfusion Medicine.* Transfusion medicine reviews, 2012. **26**(4): p. 321-332.
23. Barty, R.L., et al., *A benchmarking program to reduce red blood cell outdating: implementation, evaluation, and a conceptual framework.* Transfusion, 2015. **55**(7): p. 1621-1627.
24. Chapman, J.F. and R. Cook, *The Blood Stocks Management Scheme, a partnership venture between the National Blood Service of England and North Wales and participating hospitals for maximizing blood supply chain management.* Vox sanguinis, 2002. **83**(3): p. 239-246.
25. Beliën, J. and H. Forcé, *Supply chain management of blood products: A literature review.* European Journal of Operational Research, 2012. **217**(1): p. 1-16.
26. Prastacos, G.P., *BLOOD INVENTORY MANAGEMENT - AN OVERVIEW OF THEORY AND PRACTICE.* Management Science, 1984. **30**(7): p. 777-800.
27. Pierskella WP, R.C., *Optimal issues policies for perishable inventory.* Manage Sci, 1972(18): p. 608-14.
28. Brodheim, E., R. Hirsch, and G. Prastacos, *SETTING INVENTORY LEVELS FOR HOSPITAL BLOOD BANKS.* Transfusion, 1976. **16**(1): p. 63-70.
29. Pereira, A., *Blood inventory management in the type and screen era.* Vox sanguinis, 2005. **89**(4): p. 245-250.
30. Cohen, M.A. and W.P. Pierskalla, *Management Policies for a Regional Blood Bank.* Transfusion, 1975. **15**(1): p. 58-67.
31. Britten, A.F.H. and D.G. Geurtze, *WEEKLY ROTATION OF BLOOD INVENTORY - SYSTEM FOR SUPPLYING SMALL HOSPITALS.* Transfusion, 1979. **19**(6): p. 738-741.
32. Sharpe, C., et al., *Novel web-based real-time dashboard to optimize recycling and use of red cell units at a large multi-site transfusion service.* Journal of Pathology Informatics, 2014. **5**(1): p. 35-35.
33. Chapman, J., *Unlocking the essentials of effective blood inventory management.* Transfusion, 2007. **47**: p. 190S-196S.
34. Chapman, J.F., C. Hyam, and R. Hick, *Blood inventory management.* Vox sanguinis, 2004. **87 Suppl 2**: p. 143-5.
35. Gomez, A.T., et al., *Implementation of a novel real-time platelet inventory management system at a multi-site transfusion service.* Transfusion, 2015. **55**(9): p. 2070-2075.
36. Katsaliaki, K. and S.C. Brailsford, *Using simulation to improve the blood supply chain.* Journal of the Operational Research Society, 2007. **58**(2): p. 219-227.
37. Cheng, C.K., D. Trethewey, and I. Sadek, *Comprehensive survey of red blood cell unit life cycle at a large teaching institution in eastern Canada.* Transfusion, 2010. **50**(1): p. 160-5.
38. Heitmiller, E.S., et al., *Blood wastage reduction using Lean Sigma methodology.* Transfusion, 2010. **50**(9): p. 1887-96.
39. Bellabdaoui, A., et al., *[Optimization of inventory levels in the blood transfusion centers in Morocco].* Transfusion clinique et biologique : journal de la Societe francaise de transfusion sanguine, 2014. **21**(1): p. 15-22.
40. Delen, D., et al., *Better management of blood supply-chain with GIS-based analytics.* Annals of Operations Research, 2011. **185**(1): p. 181-193.

41. Jackups, R., Jr. and S. Kymes, *Comparison of two platelet transfusion strategies to minimize ABO-nonidentical transfusion, outdating, and shortages using a computer-simulated "virtual blood bank".* Transfusion, 2015. **55**(2): p. 348-56.
42. Katsaliaki, K., N. Mustafee, and S. Kumar, *A game-based approach towards facilitating decision making for perishable products: An example of blood supply chain.* Expert Systems with Applications, 2014. **41**(9): p. 4043-4059.
43. Ghandforoush, P. and T.K. Sen, *A DSS to manage platelet production supply chain for regional blood centers.* Decision Support Systems, 2010. **50**(1): p. 32-42.
44. Stanger, S.H., et al., *Blood inventory management: hospital best practice.* Transfusion medicine reviews, 2012. **26**(2): p. 153-63.
45. Vaquero, M.Á.P., et al., *Optimization of the management of platelet concentrate stocks in the Basque Country using mathematical simulation.* Vox sanguinis, 2016. **110**(4): p. 369-375.
46. Zoric, L., et al., *Blood wastage reduction A 10-year observational evaluation in a large teaching institution in France.* European Journal of Anaesthesiology, 2013. **30**(5): p. 250-255.
47. van der Aalst, W., *Process Mining: Discovery, Conformance and Enhancement of Business Processes*2011: Springer Verlag.
48. De Weerdt, J., et al., *Process Mining for the multi-faceted analysis of business processes-A case study in a financial services organization.* Computers in Industry, 2013. **64**(1): p. 57-67.
49. Jans, M., M.G. Alles, and M.A. Vasarhelyi, *A Field Study on the Use of Process Mining of Event Logs as an Analytical Procedure in Auditing.* Accounting Review, 2014. **89**(5): p. 1751-1773.
50. van der Aalst, W.M.P., et al., *Auditing 2.0: Using Process Mining to Support Tomorrow's Auditor.* Computer, 2010. **43**(3): p. 90-93.
51. Bozkaya, M., J. Gabriels, and J.M.v.d. Werf. *Process Diagnostics: A Method Based on Process Mining*. in *Information, Process, and Knowledge Management, 2009. eKNOW '09. International Conference on*. 2009.
52. Rebuge, Á. and D.R. Ferreira, *Business process analysis in healthcare environments: A methodology based on process mining.* Information Systems, 2012. **37**(2): p. 99-116.
53. Mustafee, N., et al., *Facilitating the Analysis of a UK National Blood Service Supply Chain Using Distributed Simulation.* Simulation-Transactions of the Society for Modeling and Simulation International, 2009. **85**(2): p. 113-128.
54. Fontaine, M.J., et al., *Improving platelet supply chains through collaborations between blood centers and transfusion services.* Transfusion, 2009. **49**(10): p. 2040-2047.
55. Hess, J.R., *Red cell freezing and its impact on the supply chain.* Transfusion medicine, 2004. **14**(1): p. 1-8.
56. Rytila, J. and K. Spens, *Using simulation to increase efficiency in blood supply chains.* Management research news, 2006. **29**(12): p. 801-819.
57. Williamson, L.M. and D.V. Devine, *Challenges in the management of the blood supply.* Lancet, 2013. **381**(9880): p. 1866-75.
58. Fan, W., et al., *The impact of increasing the upper age limit of donation on the eligible blood donor population in Canada.* Transfusion medicine, 2012. **22**(6): p. 395-403.
59. O'Brien, S.F., et al., *West Nile virus in 2006 and 2007: the Canadian Blood Services' experience.* Transfusion, 2010. **50**(5): p. 1118-1125.
60. Petersen, L.R. and J.S. Epstein, *Chikungunya virus: new risk to transfusion safety in the Americas.* Transfusion, 2014. **54**(8): p. 1911-1915.

61. Glynn, S.A., et al., *Effect of a national disaster on blood supply and safety - The September 11 experience.* Jama-Journal of the American Medical Association, 2003. **289**(17): p. 2246-2253.
62. Stainsby, D., et al., *Safety of blood donation from individuals with treated hypertension or non-insulin dependent type 2 diabetes - a systematic review.* Vox sanguinis, 2010. **98**(3): p. 431-440.
63. Fontaine, M.J., et al., *Age of blood as a limitation for transfusion: potential impact on blood inventory and availability.* Transfusion, 2010. **50**(10): p. 2233-2239.
64. Stramer, S.L., R.Y. Dodd, and A.T.-T.D. E, *Transfusion-transmitted emerging infectious diseases: 30 years of challenges and progress.* Transfusion, 2013. **53**(10): p. 2375-2383.
65. Drackley, A., et al., *Forecasting Ontario's blood supply and demand.* Transfusion, 2012. **52**(2): p. 366-74.
66. Stramer, S.L., et al., *Emerging infectious disease agents and their potential threat to transfusion safety.* Transfusion, 2009. **49**: p. 1S-235S.
67. Shi, L., et al., *Blood safety and availability: continuing challenges in China's blood banking system.* Transfusion, 2014. **54**(2): p. 471-482.
68. Cable, R.G., et al., *Iron deficiency in blood donors: the REDS-II Donor Iron Status Evaluation (RISE) study.* Transfusion, 2012. **52**(4): p. 702-711.
69. Reynolds, E., C. Wickenden, and A. Oliver, *The impact of improved safety on maintaining a sufficient blood supply.* Transfusion Clinique Et Biologique, 2001. **8**(3): p. 235-239.
70. Goodell, A.J., et al., *Costs, consequences, and cost-effectiveness of strategies for Babesia microti donor screening of the US blood supply.* Transfusion, 2014. **54**(9): p. 2245-2257.
71. Bagozzi, D. *World Blood Donor Day 2006*. 2006; Available from: http://www.who.int/mediacentre/news/releases/2006/pr33/en/.
72. *Canadian Blood Services Inventory Disposition Report*. Available from: https://blood.ca/en/hospitals/blood-component-and-product-disposition-system.
73. Nahmias, S., *PERISHABLE INVENTORY-THEORY - A REVIEW.* Operations Research, 1982. **30**(4): p. 680-708.
74. Dumas, M.B. and M. Rabinowitz, *POLICIES FOR REDUCING BLOOD WASTAGE IN HOSPITAL BLOOD BANKS.* Management Science, 1977. **23**(10): p. 1124-1132.
75. Hanne Jensen, C.M., Rosa Perry, Ralph Almazan. *UC Davis Transfusion Medicine*. Available from: https://www.ucdmc.ucdavis.edu/pathology/services/clinical/clinical_pathology/transfusion_medicine.html.
76. Hirsch, R.L. and E. Brodheim, *BLOOD DISTRIBUTION-SYSTEMS AND THE EXCHANGE OF INFORMATION BETWEEN HOSPITAL BLOOD-BANKS AND REGIONAL BLOOD CENTERS.* Vox sanguinis, 1981. **40**(3): p. 239-244.
77. Greinacher, A., et al., *Impact of demographic changes on the blood supply: Mecklenburg-West Pomerania as a model region for Europe.* Transfusion, 2007. **47**(3): p. 395-401.
78. Greinacher, A., et al., *Implications of demographics on future blood supply: a population-based cross-sectional study.* Transfusion, 2011. **51**(4): p. 702-709.
79. Lattimore, S., C. Wickenden, and S.R. Brailsford, *Blood donors in England and North Wales: demography and patterns of donation.* Transfusion, 2015. **55**(1): p. 91-99.
80. Greinacher, A., K. Fendrich, and W. Hoffmann, *Demographic Changes: The Impact for Safe Blood Supply.* Transfusion Medicine and Hemotherapy, 2010. **37**(3): p. 141-148.
81. Johannsdottir, V., et al., *Blood donors in Iceland: a nationwide population-based study from 2005 to 2013.* Transfusion, 2016. **56**(6): p. 1654-1661.

82. Borkent-Raven, B.A., M.P. Janssen, and C.L. van der Poel, *Demographic changes and predicting blood supply and demand in the Netherlands.* Transfusion, 2010. **50**(11): p. 2455-2460.
83. Ali, A., M.-K. Auvinen, and J. Rautonen, *The aging population poses a global challenge for blood services.* Transfusion, 2010. **50**(3): p. 584-588.
84. Rosenberg, M.W. and E.G. Moore, *The health of Canada's elderly population: current status and future implications.* Canadian Medical Association journal, 1997. **157**(8): p. 1025-1032.
85. Barjas-Castro, M.L., et al., *Probable transfusion-transmitted Zika virus in Brazil.* Transfusion, 2016. **56**(7): p. 1684-1688.
86. Motta, I.J.F., B.R. Spencer, and O.C. Ferreira, *Evidence for Transmission of Zika Virus by Platelet Transfusion.* New England Journal of Medicine, 2016. **375**(11): p. 1101-1103.
87. Musso, D., et al., *Zika virus: a new challenge for blood transfusion.* Lancet, 2016. **387**(10032): p. 1993-1994.
88. Riley, W., M. Schwei, and J. McCullough, *The United States' potential blood donor pool: estimating the prevalence of donor-exclusion factors on the pool of potential donors.* Transfusion, 2007. **47**(7): p. 1180-1188.
89. Custer, B., et al., *Quantifying losses to the donated blood supply due to donor deferral and miscollection.* Transfusion, 2004. **44**(10): p. 1417-1426.
90. Walsh, G.M., et al., *Blood-Borne Pathogens: A Canadian Blood Services Centre for Innovation Symposium.* Transfusion medicine reviews, 2016. **30**(2): p. 53-68.
91. Dodd, R.Y., *Emerging pathogens and their implications for the blood supply and transfusion transmitted infections.* British Journal of Haematology, 2012. **159**(2): p. 135-142.
92. Shehata, N., et al., *Changing trends in blood transfusion: an analysis of 244,013 hospitalizations.* Transfusion, 2014. **54**(10): p. 2631-2639.
93. Greinacher, A. and W. Hoffmann, *Why has the demand for platelet components increased? A commentary.* Transfusion medicine, 2014. **24**(5): p. 257-259.
94. Estcourt, L.J., *Why has demand for platelet components increased? A review.* Transfusion medicine, 2014. **24**(5): p. 260-268.
95. Beckman, N., et al., *Blood banks meet the paradox of Gabriel's Horn: what are the options to maintain supply as demand decreases?* Transfusion medicine, 2016. **26**(3): p. 170-176.
96. Kleineruschkamp, A.G., et al., *Cost analysis of patient blood management.* Anaesthesist, 2016. **65**(6): p. 438-447.
97. Bell, C.F., et al., *Cost-effectiveness of transfusion of platelet components prepared with pathogen inactivation treatment in the United States.* Clinical Therapeutics, 2003. **25**(9): p. 2464-2486.
98. Girona-Llobera, E., et al., *Reducing the financial impact of pathogen inactivation technology for platelet components: our experience.* Transfusion, 2014. **54**(1): p. 158-168.
99. Fragoulakis, V., et al., *The Cost of Blood Collection in Greece: An Economic Analysis.* Clinical Therapeutics, 2014. **36**(7): p. 1028-1036.
100. Jefferies, L.C., B.S. Sachais, and D.S. Young, *Blood transfusion costs by diagnosis-related groups in 60 university hospitals in 1995.* Transfusion, 2001. **41**(4): p. 522-529.
101. AuBuchon, J.P., et al., *Evolution in a centralized transfusion service.* Transfusion, 2011. **51**(12): p. 2750-2757.

102. Carden, R. and J.L. DelliFraine, *An examination of blood center structure and hospital customer satisfaction: what can centralized and decentralized blood centers learn from each other?* Health marketing quarterly, 2005. **22**(3): p. 21-42.
103. Delaney, M., et al., *The immunohematologic and patient safety benefits of a centralized transfusion database.* Transfusion, 2013. **53**(4): p. 771-776.
104. MacIvor, D., D.J. Triulzi, and M.H. Yazer, *Enhanced detection of blood bank sample collection errors with a centralized patient database.* Transfusion, 2009. **49**(1): p. 40-43.
105. Pagliaro, P., R. Turdo, and E. Capuzzo, *Patients' positive identification systems.* Blood Transfusion, 2009. **7**(4): p. 313-318.
106. Lippi, G. and M. Plebani, *Identification errors in the blood transfusion laboratory: A still relevant issue for patient safety.* Transfusion and Apheresis Science, 2011. **44**(2): p. 231-233.
107. Wittmann, G., et al., *Automation and data processing with the immucor Galileo (R) system in a university blood bank.* Transfusion Medicine and Hemotherapy, 2007. **34**(5): p. 347-352.
108. Chang, C., et al., *Evaluation of Erytra (R) fully automated analyser for Routine Use in Transfusion Laboratory.* Transfusion medicine, 2014. **24**(1): p. 33-38.
109. Shin, S.Y., et al., *Evaluation of two automated instruments for pre-transfusion testing: AutoVue Innova and Techno TwinStation.* Korean Journal of Laboratory Medicine, 2008. **28**(3): p. 214-220.
110. South, S.F., T.S. Casina, and L. Li, *Exponential error reduction in pretransfusion testing with automation.* Transfusion, 2012. **52**(8): p. 81S-87S.
111. Pendry, K., *The use of big data in transfusion medicine.* Transfusion medicine, 2015. **25**(3): p. 129-137.
112. *Richard Gillespie. Manufacturing knowledge: A history of the hawthorne experiments. Cambridge: Cambridge University Press, 1991. 304 pp. $44.95 (cloth) (Reviewed by Richard C. S. trahair).* Journal of the History of the Behavioral Sciences, 1994. **30**(3): p. 247-252.
113. Dumas, M., W.M.v.d. Aalst, and A.H.t. Hofstede, *Process-aware information systems: bridging people and software through process technology*2005: John Wiley \\& Sons, Inc. 409.
114. Van der Aalst, W., *Process design by discovery: Harvesting workflow knowledge from ad-hoc executions.* Knowledge Management: An Interdisciplinary Approach, Dagstuhl Seminar Report, 2000(281).
115. van der Aalst, W.M.P., et al., *Process mining manifesto*, in *BPM 2011 Workshops Proceedings*2012, Springer-Verlag. p. 169-194.
116. Golfarelli, M., S. Rizzi, and I. Cella. *Beyond data warehousing: what's next in business intelligence?* in *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP*. 2004. ACM.
117. Grigori, D., et al., *Business process intelligence.* Computers in Industry, 2004. **53**(3): p. 321-343.
118. Weske, M., W.M. van der Aalst, and H. Verbeek, *Advances in business process management.* Data & Knowledge Engineering, 2004. **50**(1): p. 1-8.
119. Aalst, W.v.d., *Process Mining: Overview and Opportunities.* ACM Trans. Manage. Inf. Syst., 2012. **3**(2): p. 1-17.
120. Agrawal, R., D. Gunopulos, and F. Leymann, *Mining Process Models from Workflow Logs*, in *Proceedings of the 6th International Conference on Extending Database Technology: Advances in Database Technology*1998, Springer-Verlag. p. 469-483.

121. Cook, J.E. and A.L. Wolf, *Discovering models of software processes from event-based data.* ACM Trans. Softw. Eng. Methodol., 1998. **7**(3): p. 215-249.
122. Lyytinen, K., et al., *Automating the Discovery of As-Is Business Process Models: Probabilistic and Algorithmic Approaches.* Info. Sys. Research, 1998. **9**(3): p. 275-301.
123. Schimm, G., *Process Miner - A Tool for Mining Process Schemes from Event-Based Data*, in *Proceedings of the European Conference on Logics in Artificial Intelligence*2002, Springer-Verlag. p. 525-528.
124. Schimm, G., *Mining exact models of concurrent workflows.* Comput. Ind., 2004. **53**(3): p. 265-281.
125. Van der Aalst, W., T. Weijters, and L. Maruster, *Workflow mining: Discovering process models from event logs.* Knowledge and Data Engineering, IEEE Transactions on, 2004. **16**(9): p. 1128-1142.
126. Weijters, A. and J. Ribeiro. *Flexible heuristics miner (FHM)*. in *Computational Intelligence and Data Mining (CIDM), 2011 IEEE Symposium on*. 2011. IEEE.
127. Günther, C. and W.P. Aalst, *Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics*, in *Business Process Management*, G. Alonso, P. Dadam, and M. Rosemann, Editors. 2007, Springer Berlin Heidelberg. p. 328-343.
128. Medeiros, A.K.A., A.J.M.M. Weijters, and W.M.P. Aalst, *Genetic process mining: an experimental evaluation.* Data Mining and Knowledge Discovery, 2007. **14**(2): p. 245-304.
129. De Weerdt, J., et al., *A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs.* Information Systems, 2012.
130. Agrawal, R., D. Gunopulos, and F. Leymann, *Mining process models from workflow logs*, in *Advances in Database Technology - Edbt'98*, H.J. Schek, et al., Editors. 1998. p. 469-483.
131. Mannila, H. and C. Meek, *Global partial orders from sequential data*, in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*2000, ACM: Boston, Massachusetts, USA. p. 161-168.
132. Herbst, J. and D. Karagiannis, *Workflow mining with InWoLvE.* Computers in Industry, 2004. **53**(3): p. 245-264.
133. B.F. van Dongen, W.M.P.v.d.A. *Multi-phase process mining: aggregating instance graphs into EPCs and Petri nets*. in *Proceedings of the 2nd International Workshop on Applications of Petri Nets to Coordination, Workflow and Business Process Management (PNCWB).*
134. Gaaloul, W., K. Baina, and C. Godart, *Towards mining structural workflow patterns*, in *Database and Expert Systems Applications, Proceedings*, K.V. Andersen, J. Debenham, and R. Wagner, Editors. 2005. p. 24-33.
135. Weijters, A. and W.M.P. van der Aalst, *Rediscovering workflow models from event-based data using little thumb.* Integrated Computer-Aided Engineering, 2003. **10**(2): p. 151-162.
136. Weijters, A., W.M.P. van der Aalst, and A.K.A. de Medeiros, *Process mining with the heuristics miner-algorithm.* Technische Universiteit Eindhoven, Tech. Rep. WP, 2006. **166**.
137. Greco, G., et al., *Discovering expressive process models by clustering log traces.* Ieee Transactions on Knowledge and Data Engineering, 2006. **18**(8): p. 1010-1027.
138. Maruster, L., et al., *A rule-based approach for process discovery: Dealing with noise and imbalance in process logs.* Data Mining and Knowledge Discovery, 2006. **13**(1): p. 67-87.

139. Ferreira, H.M. and D.R. Ferreira, *An integrated life cycle for workflow management based on learning and planning.* International Journal of Cooperative Information Systems, 2006. **15**(4): p. 485-505.
140. Medeiros, A.K.A.d., *Genetic Process Mining*, 2006, TU Eindhoven: Eindhoven.
141. Gunther, C.W. and W.M.R. van der Aalst, *Fuzzy mining - Adaptive process simplification based on multi-perspective metrics*, in *Business Process Management, Proceedings*, G. Alonso, P. Dadam, and M. Rosemann, Editors. 2007. p. 328-343.
142. Wen, L., et al., *Mining process models with non-free-choice constructs.* Data Min. Knowl. Discov., 2007. **15**(2): p. 145-180.
143. Lamma, E., et al., *Inducing declarative logic-based models from labeled traces*, in *Business Process Management, Proceedings*, G. Alonso, P. Dadam, and M. Rosemann, Editors. 2007. p. 344-359.
144. Greco, G., A. Guzzo, and L. Pontieri, *Mining taxonomies of process models.* Data & Knowledge Engineering, 2008. **67**(1): p. 74-102.
145. Goedertier, S., et al., *Robust Process Discovery with Artificial Negative Events.* Journal of Machine Learning Research, 2009. **10**: p. 1305-1340.
146. Wen, L., et al., *A novel approach for process mining based on event types.* J. Intell. Inf. Syst., 2009. **32**(2): p. 163-190.
147. Folino, F., et al., *Discovering expressive process models from noised log data*, in *Proceedings of the 2009 International Database Engineering & Applications Symposium*2009, ACM: Cetraro - Calabria, Italy. p. 162-172.
148. Ferreira, D.R. and D. Gillblad, *Discovering Process Models from Unlabelled Event Logs*, in *Business Process Management, Proceedings*, U. Dayal, et al., Editors. 2009. p. 143-158.
149. van der Werf, J.M.E.M., et al., *Process Discovery using Integer Linear Programming.* Fundamenta Informaticae, 2009. **94**(3-4): p. 387-412.
150. van der Aalst, W.M.P., et al., *Process mining: a two-step approach to balance between underfitting and overfitting.* Software and Systems Modeling, 2010. **9**(1): p. 87-111.
151. Carmona, J., J. Cortadella, and M. Kishinevsky, *New Region-Based Algorithms for Deriving Bounded Petri Nets.* Ieee Transactions on Computers, 2010. **59**(3): p. 371-384.
152. Werner, M. and N. Gehrke, *Multilevel Process Mining for Financial Audits.* Ieee Transactions on Services Computing, 2015. **8**(6): p. 820-832.
153. Aalst, W.M.P.v.d. and B.F.v. Dongen, *Discovering Workflow Performance Models from Timed Logs*, in *Proceedings of the First International Conference on Engineering and Deployment of Cooperative Information Systems*2002, Springer-Verlag. p. 45-63.
154. van der Aalst, W.M., et al., *Business process mining: An industrial application.* Information Systems, 2007. **32**(5): p. 713-732.
155. Pink, J., A. Thomson, and B. Wylie, *Inventory management in Sydney Public Hospital Blood Banks.* Transfusion medicine, 1994. **4**(7820233): p. 237-242.
156. Dodds, A.J., et al., *A rational approach to crossmatching blood for elective surgery.* Anaesth Intensive Care, 1983. **11**(6407355): p. 16-19.
157. Napier, J.A. and A. Biffin, *The importance of cross-match/transfusion ratio and intervals between successive cross-matches in determining red cell outdate rates.* Clin Lab Haematol, 1983. **5**(6667602): p. 379-385.
158. Axelrod, F.B., A.J. Grindon, and D.H. Vroon, *Disproportionate outdating of group A blood.* Transfusion, 1987. **27**(3590282): p. 219-221.
159. Hirsch, R.L., E. Brodheim, and F.E. Ginsberg, *A COMPUTER-BASED BLOOD INVENTORY AND INFORMATION SYSTEM FOR HOSPITAL BLOOD BANKS AS PART OF A REGIONAL BLOOD-MANAGEMENT PROGRAM.* Transfusion, 1970. **10**(4): p. 194-&.

160. Catassi, C.A. and E.L. Peterson, *The blood inventory control system--helping blood bank management through computerized inventory control.* Transfusion, 1967. **7**(6017396): p. 60-69.
161. Georgsen, J. and T. Kristensen, *From serological to computer crossmatching in nine hospitals.* Vox sanguinis, 1998. **74**: p. 419-425.
162. Davis, R., et al., *Tracking blood products in blood centres using radio frequency identification: a comprehensive assessment.* Vox sanguinis, 2009. **97**(1): p. 50-60.
163. Ledman, R.E. and N. Groh, *Platelet production planning to ensure availability while minimizing outdating.* Transfusion, 1984. **24**(6506187): p. 532-533.
164. Galloway, M.J., et al., *A tabletop exercise to assess a hospital emergency blood management contingency plan in a simulated acute blood shortage.* Transfusion medicine, 2008. **18**(18937738): p. 302-307.
165. Denesiuk, L., et al., *Implementation of a redistribution system for near-outdate red blood cell units.* Archives of pathology & laboratory medicine, 2006. **130**(8): p. 1178-83.
166. Kendall, K.E., *PLANNING FOR INNOVATION IN HEALTH-CARE - A BREAKTHROUGH IN BLOOD BANKING.* Long Range Planning, 1984. **17**(6): p. 127-131.
167. Custer, B., et al., *Community blood supply model: Development of a new model to assess the safety, sufficiency, and cost of the blood supply.* Medical Decision Making, 2005. **25**(5): p. 571-582.
168. Schreiber, G.B., et al., *Convenience, the bane of our existence, and other barriers to donating.* Transfusion, 2006. **46**(16584430): p. 545-553.
169. Schreiber, G.B., et al., *First year donation patterns predict long-term commitment for first-time donors.* Vox sanguinis, 2005. **88**(15720609): p. 114-121.
170. Poisson, J.L., C.W. Tuma, and I.A. Shulman, *Inventory management strategies that reduce the age of red blood cell components at the time of transfusion.* Transfusion, 2016. **56**(7): p. 1758-1762.
171. Maria, A., *Introduction to modeling and simulation*, in *Proceedings of the 29th conference on Winter simulation*1997, IEEE Computer Society: Atlanta, Georgia, USA. p. 7-13.
172. Elston, R.C. and J.C. Pickrel, *GUIDES TO INVENTORY LEVELS FOR A HOSPITAL BLOOD BANK DETERMINED BY ELECTRONIC COMPUTER SIMULATION.* Transfusion, 1965. **5**(5): p. 465-&.
173. Prastacos, G.P. and E. Brodheim, *PBDS - A DECISION SUPPORT SYSTEM FOR REGIONAL BLOOD MANAGEMENT.* Management Science, 1980. **26**(5): p. 451-463.
174. Brodheim, E. and G.P. Prastacos, *A regional blood management system with prescheduled deliveries.* Transfusion, 1979. **19**(4): p. 455-62.
175. Jennings, J.B., *An analysis of hospital blood bank whole blood inventory control policies.* Transfusion, 1968. **8**(5743048): p. 335-342.
176. Pereira, A., *Economies of scale in blood banking: a study based on data envelopment analysis.* Vox sanguinis, 2006. **90**(16635074): p. 308-315.
177. Pitocco, C. and T.R. Sexton, *Alleviating blood shortages in a resource-constrained environment.* Transfusion, 2005. **45**(7): p. 1118-1126.
178. Brodheim, E., C. Derman, and G. Prastacos, *EVALUATION OF A CLASS OF INVENTORY POLICIES FOR PERISHABLE PRODUCTS SUCH AS BLOOD.* Management Science Series a-Theory, 1975. **21**(11): p. 1320-1325.
179. Pegels, C.C. and A.E. Jelmert, *EVALUATION OF BLOOD-INVENTORY POLICIES - MARKOV CHAIN APPLICATION.* Operations Research, 1970. **18**(6): p. 1087-&.

180. Cumming, P.D., et al., *COLLECTIONS PLANNING MODEL FOR REGIONAL BLOOD SUPPLIERS - DESCRIPTION AND VALIDATION.* Management Science, 1976. **22**(9): p. 962-971.
181. Nahmias, S. and W.P. Pierskalla, *2-PRODUCT PERISHABLE NONPERISHABLE INVENTORY PROBLEM.* Siam Journal on Applied Mathematics, 1976. **30**(3): p. 483-500.
182. Hemmelmayr, V., et al., *Vendor managed inventory for environments with stochastic product usage.* European Journal of Operational Research, 2010. **202**(3): p. 686-695.
183. Federgruen, A.W.I., G. Prastacos, and P.H. Zipkin, *AN ALLOCATION AND DISTRIBUTION MODEL FOR PERISHABLE PRODUCTS.* Operations Research, 1986. **34**(1): p. 75-82.
184. Kendall, K.E., *MULTIPLE OBJECTIVE PLANNING FOR REGIONAL BLOOD CENTERS.* Long Range Planning, 1980. **13**(4): p. 98-104.
185. Elston, R.C. and J.C. Pickrel, *A STATISTICAL APPROACH TO ORDERING AND USAGE POLICIES FOR A HOSPITAL BLOOD BANK.* Transfusion, 1963. **3**(1): p. 41-&.
186. Frankfur.Gm, K.E. Kendall, and C.C. Pegels, *MANAGEMENT CONTROL OF BLOOD THROUGH A SHORT-TERM SUPPLY-DEMAND FORECAST SYSTEM.* Management Science Series B-Application, 1974. **21**(4): p. 444-452.
187. Critchfield, G.C., et al., *Automatic prediction of platelet utilization by time series analysis in a large tertiary care hospital.* American journal of clinical pathology, 1985. **84**(5): p. 627-31.
188. Cohen, M.A., et al., *An overview of a hierarchy of planning models for Regional Blood Bank Management.* Transfusion, 1979. **19**(505525): p. 526-534.
189. Bosnes, V., M. Aldrin, and H.E. Heier, *Predicting blood donor arrival.* Transfusion, 2005. **45**(2): p. 162-170.
190. James, R.C. and D.E. Matthews, *Analysis of blood donor return behaviour using survival regression methods.* Transfusion medicine, 1996. **6**(8696444): p. 21-30.
191. Godin, G., et al., *Determinants of repeated blood donation among new and experienced blood donors.* Transfusion, 2007. **47**(9): p. 1607-1615.
192. Godin, G., et al., *Factors explaining the intention to give blood among the general population.* Vox sanguinis, 2005. **89**(16146506): p. 140-149.
193. Prastacos, G.P., *OPTIMAL MYOPIC ALLOCATION OF A PRODUCT WITH FIXED LIFETIME.* Journal of the Operational Research Society, 1978. **29**(9): p. 905-913.
194. Prastacos, G.P., *Allocation of a Perishable Product Inventory.* Operations Research, 1981. **29**(1): p. 95-107.
195. Cohen, M.A. and W.P. Pierskalla, *Target inventory levels for a hospital blood bank or a decentralized regional blood banking system.* Transfusion, 1979. **19**(4): p. 444-54.
196. Yu, P.L.H., et al., *Predicting potential drop-out and future commitment for first-time donors based on first 1.5-year donation patterns: the case in Hong Kong Chinese donors.* Vox sanguinis, 2007. **93**(17547566): p. 57-63.
197. Rabinowitz, M., *Blood bank inventory policies: a computer simulation.* Health services research, 1973. **8**(4206023): p. 271-282.
198. Cumming, P.D., et al., *Cost effectiveness of use of frozen blood to alleviate blood shortages.* Transfusion, 1977. **17**(413224): p. 602-606.
199. Friedman, B.A., R.D. Abbott, and G.W. Williams, *A BLOOD ORDERING STRATEGY FOR HOSPITAL BLOOD-BANKS DERIVED FROM A COMPUTER-SIMULATION.* American journal of clinical pathology, 1982. **78**(2): p. 154-160.
200. Sirelson, V. and E. Brodheim, *A computer planning model for blood platelet production and distribution.* Computer methods and programs in biomedicine, 1991. **35**(4): p. 279-91.

201. van Dijk, N., et al., *Blood platelet production: a novel approach for practical optimization.* Transfusion, 2009. **49**(3): p. 411-420.
202. Haijema, R., J. van der Wal, and N.M. van Dijk, *Blood platelet production: Optimization by dynamic programming and simulation.* Computers & Operations Research, 2007. **34**(3): p. 760-779.
203. Haijema, R., et al., *Blood platelet production with breaks: optimization by SDP and simulation.* International Journal of Production Economics, 2009. **121**(2): p. 464-473.
204. Katsaliaki, K., *Cost-effective practices in the blood service sector.* Health Policy, 2008. **86**(2-3): p. 276-287.
205. Graf, S., E. Morse, and A. Katz, *BLOOD INVENTORY DISTRIBUTION BY TYPE IN A TOTAL SUPPLY-SYSTEM.* Transfusion, 1972. **12**(3): p. 185-&.
206. Kamp, C., et al., *Management of blood supplies during an influenza pandemic.* Transfusion, 2010. **50**(1): p. 231-239.
207. Kopach, R., B. Balcioglu, and M. Carter, *Tutorial on constructing a red blood cell inventory management system with two demand rates.* European Journal of Operational Research, 2008. **185**(3): p. 1051-1059.
208. Asllani, A., E. Culler, and L. Ettkin, *A simulation-based apheresis platelet inventory management model.* Transfusion, 2014. **54**(10pt2): p. 2730-2735.
209. Baesler, F., et al., *Analysis of inventory strategies for blood components in a regional blood center using process simulation.* Transfusion, 2014. **54**(2): p. 323-330.
210. Jagannathan, R. and T. Sen, *STORING CROSS-MATCHED BLOOD - A PERISHABLE INVENTORY MODEL WITH PRIOR ALLOCATION.* Management Science, 1991. **37**(3): p. 251-266.
211. Goedertier, S., et al., *Process discovery in event logs: An application in the telecom industry.* Applied Soft Computing, 2011. **11**(2): p. 1697-1710.
212. Mahendrawathi, E.R., H.M. Astuti, and A. Nastiti, *Analysis of Customer Fulfilment with Process Mining: A Case Study in a Telecommunication Company*, in *Third Information Systems International Conference 2015*, G. Bordea, Editor 2015, Elsevier Science Bv: Amsterdam. p. 588-596.
213. Mans, R.S., et al., *Application of Process Mining in Healthcare - A Case Study in a Dutch Hospital*, in *Biomedical Engineering Systems and Technologies*, A. Fred, J. Filipe, and H. Gamboa, Editors. 2008, Springer-Verlag Berlin: Berlin. p. 425-438.
214. Yang, W.S. and S.Y. Hwang, *A process-mining framework for the detection of healthcare fraud and abuse.* Expert Systems with Applications, 2006. **31**(1): p. 56-68.
215. Jans, M., et al., *A business process mining application for internal transaction fraud mitigation.* Expert Systems with Applications, 2011. **38**(10): p. 13351-13359.
216. Jans, M., M. Alles, and M. Vasarhelyi, *The case for process mining in auditing: Sources of value added and areas of application.* International Journal of Accounting Information Systems, 2013. **14**(1): p. 1-20.
217. Rozinat, A., et al., *Process Mining Applied to the Test Process of Wafer Scanners in ASML.* Ieee Transactions on Systems Man and Cybernetics Part C-Applications and Reviews, 2009. **39**(4): p. 474-479.
218. Paszkiewicz, Z., *Process Mining Techniques in Conformance Testing of Inventory Processes: An Industrial Application*, in *Business Information Systems Workshops, Bis 2013*, W. Abramowicz, Editor 2013, Springer-Verlag Berlin: Berlin. p. 302-313.
219. in *Merriam-Webster Online Dictionary*.
220. in *Dictionary.com*.

221. Brocke, J.v., *Handbook on Business Process Management*, in *International Handbooks on Information Systems Ser.*2010, Springer: New York. p. 600 p.
222. Cheng, C.K., et al., *Methods to create an evidence-based maximum surgical blood order schedule.* Transfusion, 2006. **46**(9): p. 150A-150A.
223. Services, W.S.D.o.E.; Available from: http://www.des.wa.gov/services/Risk/AboutRM/enterpriseRiskManagement/Pages/rootCauseAnalysis.aspx.