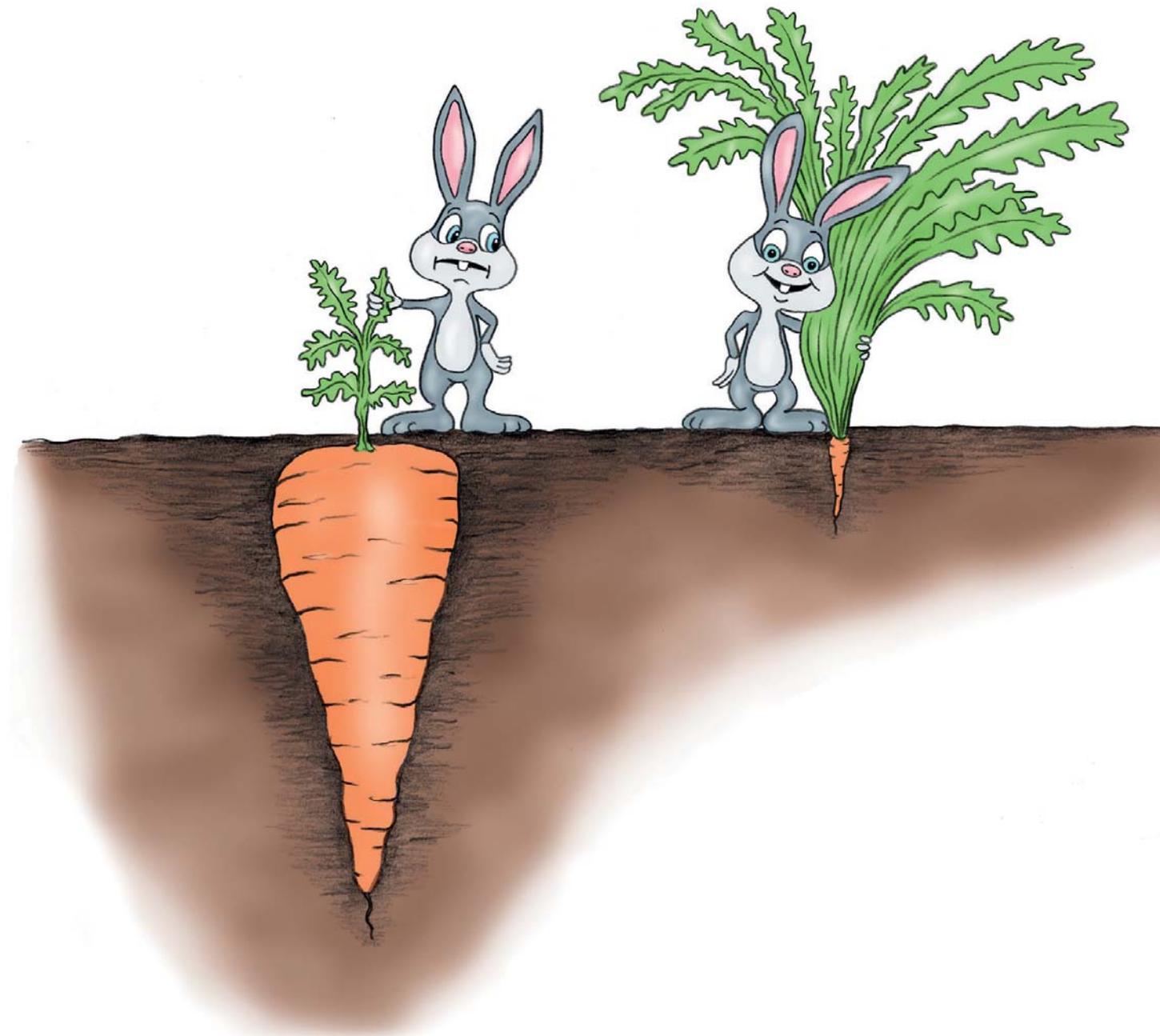


Value-Based Health Care: Challenges in moving forward



David Ebbevi



**Karolinska
Institutet**

From DEPARTMENT OF LEARNING, INFORMATICS,
MANAGEMENT AND ETHICS
Medical Management Centre
Karolinska Institutet, Stockholm, Sweden

VALUE-BASED HEALTH CARE

CHALLENGES IN MOVING FORWARD

David Ebbevi



**Karolinska
Institutet**

Stockholm 2016

All previously published papers were reproduced with permission from the publisher.
Cover illustration by *Lena Hellström Sparring, Din tecknande sparringpartner*. Inspired by internet meme, “Success: it’s not always what you see”.

Published by Karolinska Institutet.

Printed by EPrint AB 2017

© David Ebbevi, 2017

ISBN 978-91-7676-567-8

Value-Based Health Care: Challenges in Moving Forward

THESIS FOR DOCTORAL DEGREE (Ph.D.)

Publicly defended in CMB Lecture Hall, Karolinska Institute, Solna:

Friday, 3 February 2017, at 9:00 a.m.

By

David Ebbevi

Principal Supervisor:

Dr. Helena Hvitfeldt Forsberg
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Medical Management Centre

Co-supervisors:

Associate Professor Anna Essén
Stockholm School of Economics
Center for Human Resource Management and
Knowledge Work

Professor Staffan Lindblad
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Medical Management Centre

Dr. Sofia Ernestam
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Medical Management Centre

Opponent:

Associate Professor Charlotta Levay
Lund University
Department of Business Administration

Examination Board:

Professor Boel Andersson-Gäre
Jönköping University
Jönköping Academy for Improvement of Health
and Welfare

Associate Professor Magna Andréén Sachs
Karolinska Institutet
Department of Learning, Informatics,
Management and Ethics
Medical Management Centre

Professor Birger Forsberg
Karolinska Institutet
Department of Public Health Sciences
Division of Global Health

Associate Professor Lene Lindberg
Karolinska Institutet
Department of Public Health Sciences
Division of Applied Public Health

Associate Professor Erik Malmqvist
Linköping University
Department of Thematic Studies
Division of Technology and Social Change

To you;
the past, current or future patient,
my partner in care
and my teacher of life.

“Measure what is measurable,
and make measurable what is not so”

– Galileo Galilei

“A great building must begin with the unmeasurable,
must go through measurable means when it is being designed
and in the end must be unmeasurable”

– Louis Kahn

“I don’t follow orders,
and I don’t lie about what I do to manage my condition;
what I do I do for myself”

– Dominick L Frosch

ABSTRACT

Background. Value-based health care (VBHC), a strategic framework launched in 2006, suggests restructuring health care to maximize health outcomes in relation to societal costs. The framework builds upon a tradition of outcome assessment in clinical work and health economic evaluation. In suggesting maximization of outcomes in relation to costs, the framework makes assumptions about what constitute a valuable outcome to persons with a disease, and how this should be operationalized. The framework, for example, assumes: (A) only outcomes matter to patients; (B) outcomes that matter to patients are separable into three categories: health status achieved, time to health status achieved, and sustainability of health; and (C) standardized instruments are preferable when operationalizing a certain outcome. Critique has been raised towards VBHC as being ill adapted to the chronic care setting.

Aim. The aim of this thesis is to assess assumptions (A-C) of VBHC in a chronic care setting.

Methods. This thesis uses a mixed-methods approach. First, the current understanding and application of VBHC in the research literature is assessed using literature citation data. Secondly, empirical consequences of the assumptions (A-C) are assessed in the setting of rheumatology, using a randomized controlled trial (A) and semi-structured interviews (B-C). The research participants had rheumatoid arthritis (RA).

Findings. This thesis shows that the literature has not assessed assumptions of VBHC empirically; non-outcomes can have intrinsic value to persons with RA (A); outcomes not found in the outcome categories can have intrinsic value to persons with RA (B); and a common functional measurement in rheumatology, Stanford's Health Assessment Questionnaire-Disability Index, is not aligned with the perspective of persons with RA (C).

Discussion. The framework assumptions did not find support in the setting of rheumatology. This thesis does not address settings other than rheumatology and not all aspects of the assumptions. However, the findings show that the assumptions are not universally applicable. There are several possible reasons for the disagreement between the assumptions and the findings. The most reasonable is that VBHC addresses care from the perspective of the provider. Although the needs and challenges of the provider might be fully addressed by VBHC, the perspective of patients with RA is not. The disagreement between the assumptions and the findings, in combination with the lack of assessment in the scientific literature, raises the question whether VBHC is a scientific or a non-scientific framework.

Conclusion. The thesis shows limitations of VBHC in one specific setting, using data from persons with RA. Those specific limitations can most likely be avoided if VBHC were adjusted and implemented in patient-professional partnership.

PREFACE

To sustainably change how the world is lived, one must start with the world as it is. My scientific endeavor began out of this necessity—to maximize clarity of thought and correspondence of my mental models to this (perceived) external reality. I started in mathematics but my inner norms pushed me into medicine. There, my passion grew from the idea of improving life for persons who sought my advice as a physician, and science became the method simply because it is the best known way of tackling challenges. Although working as a physician can be rewarding, interesting, and meaningful, the deed rapidly cools as you gain a system perspective of health and suffering.

When I summarized the achievements—six years after I planned the first study in this thesis—I reconciled with the idea that science has indefinite room for improvement. For me, becoming a researcher has been about accepting the limitations of research design and conclusions, and as such learning to smile at the endless iterations. The challenge was always to know when a study is—as the management consultants would phrase it—“good enough.”

Enjoy the reading, insights, and failures. I did.

A handwritten signature in black ink, appearing to read 'David Ebbevi', with a stylized flourish at the end.

David Ebbevi

LIST OF SCIENTIFIC PAPERS

- I. Fredriksson JJ*, **Ebbevi D***, Savage C. Pseudo-understanding: an analysis of the dilution of value in healthcare. *BMJ Qual Saf.* 2015 Jul 1;24(7):451-7.
- II. Fredriksson C*, **Ebbevi D***, Waldheim E, Lindblad S, Ernestam S. Patient-initiated appointments compared with standard outpatient care for rheumatoid arthritis: a randomised controlled trial. *RMD Open.* 2016 Mar 1;2(1):e000184.
- III. **Ebbevi D**, Hvitfeldt Forsberg H, Essén A, Ernestam S. Value-based health care for chronic care: aligning outcomes measurement with the patient perspective. *Qual Manag Health Care.* 2016 Oct;25(4):203.
- IV. **Ebbevi D**, Essén A, Hvitfeldt Forsberg H. Persons with rheumatoid arthritis challenge the relevance of the Health Assessment Questionnaire: A qualitative study of patient perception. Preprints. 2017, DOI: 10.20944/preprints201701.0013.v1. Manuscript.

*Equal contributions

CONTENTS

Dedication	v
Abstract	vii
Preface.....	ix
List of Scientific Papers.....	x
List of Figures	xiii
List of Tables	xiv
List of Abbreviations	xv
1 Background.....	1
1.1 Value-Based Health Care.....	1
1.2 Pseudoinnovation	1
1.3 Value Frameworks in Health Care.....	2
1.3.1 Cost-effectiveness analysis to quantify value	2
1.3.2 VBHC: What is it?	3
1.4 VBHC in a Chronic Care Setting.....	8
1.5 VBHC Critique and Assumptions.....	9
1.6 Patient Participation.....	9
2 Aim	11
3 Methodology.....	13
3.1 Philosophical Assumptions	13
3.2 Overarching Methodology	15
3.3 Design	17
3.3.1 Summary of methods	17
3.4 Characterizing the Most-Cited VBHC Article	18
3.5 Empirical Setting and Study Participants	20
3.6 Linking Assumptions, Design, Data collection and Analysis.....	21
3.6.1 Outcomes and only outcomes have intrinsic value to patients	21
3.6.2 Outcomes important to patients can be subsumed under the three-tier model.....	23
3.6.3 Standardized outcome instruments measure outcomes that matter to patients.....	24
3.7 Applying Criteria for Research Quality.....	25
3.8 Ethical Considerations.....	27
3.8.1 Informed consent.....	27
3.8.2 Data collection.....	28
3.8.3 Data handling	28
4 Findings	31
4.1 VBHC is Not Empirically Assessed in the Literature.....	32
4.2 Non-Outcomes can have Intrinsic Value to Persons with RA	32
4.3 Outcomes not in the Three-Tier Model can have Intrinsic Value to Persons with RA	33

4.4	HAQ does not Align with the Perspective of Persons with RA	35
5	Discussion.....	37
5.1	Interpretations of the Findings	37
5.1.1	How well does VBHC acknowledge patient perspective?	37
5.1.2	Are only outcomes important to persons with RA?.....	39
5.1.3	Is VBHC a progressive research program?	40
5.1.4	Is VBHC a pseudoinnovation?	42
5.2	A New Substratgy Moving Forward?.....	43
5.3	Transferability	43
5.4	Strengths and Limitations	44
6	Conclusion.....	47
6.1	Implications	47
6.1.1	For patients	47
6.1.2	For policy makers and managers	47
6.1.3	For healthcare professionals	47
6.1.4	For outcome researchers	47
7	Acknowledgements	49
8	References	53

LIST OF FIGURES

Figure 1. Outcomes in Rheumatoid Arthritis.....	3
Figure 2. Three-Tier Model for Outcomes.	5
Figure 3. Philosophical and Methodological Positioning of the Present Thesis	16
Figure 4. Method Used in Substudy I.	18
Figure 5. Mean DAS28 in Intervention Group.....	33
Figure 6. Simplified Version of Outcome Model Developed in Substudy III.....	34

LIST OF TABLES

Table 1. The Three-Tier Model expanded for the Chronic Care Setting.....	6
Table 2. Specific Aims Investigated in Substudies I to IV.....	11
Table 3. Outline of Substudies in the Thesis	17
Table 4. Study Participants' Characteristics	21
Table 5. Overview of Substudy Findings	31
Table 6. HAQ Experience for Persons with RA.....	35

LIST OF ABBREVIATIONS

DAS28	28-joint disease activity score
HAQ	Stanford's health assessment questionnaire – disability index
ICF	international classification of function
IQR	interquartile range
RA	rheumatoid arthritis
RCT	randomized controlled trial
SOLO	structure of observed learning outcomes
TQM	total quality management
VAS	visual analog scale
VBHC	value-based health care

1 BACKGROUND

1.1 VALUE-BASED HEALTH CARE

In the past century, many attempts have been made to improve care while containing societal costs (1). Increased life span, technological improvement, and better-informed patients have increased the demand for care (2, 3). However, despite increased spending, health care systems are struggling to cover the increased demand (1).

Attempts to improve care while containing costs have resulted in an explosion of approaches (e.g., (1, 3-5)) and strategic frameworks, but no consensus on what it means to “improve” care. One of these strategies has been termed value-based health care (VBHC). Porter, its main proponent, outlined VBHC in a series of publications as a strategy for *providers* (4, 6-12). The main principle of VBHC is that the purpose of health care is to maximize results (*outcomes*) in relation to the money spent.

This thesis compares underlying assumptions of outcome measurements in VBHC with the perspective of RA patients. Other aspects of VBHC, such as costs, will not be addressed.

The many frameworks for improving care share similar traits. Before diving deeper into VBHC and its specific similarities to other frameworks, the next section (1.2) addresses a concept useful in understanding how such similarities arise: *pseudoinnovation*.

1.2 PSEUDOINNOVATION

The many views on what it means to *improve* care could give the impression that there is an infinity of ways to manage care. However, Bohmer suggested the tools available to health care managers are in fact limited to managing through values, resources, outcomes, processes, and decision systems (1). This limited amount of tools would imply that all strategic frameworks are restricted to different configurations of those tools. In addition, many specific quality-improvement tools (e.g., Lean (13) or Total Quality Management, TQM (14)) are conceptually similar to each other (15).

Despite the limited tools available and the conceptual similarity among the specific tools, new quality-improvement tools are often launched as different from predecessors despite their strikingly likenesses. The phenomena giving rise to the conceptual similarity between new and existing management ideas has been termed *the fashion perspective on innovation* (16) or *pseudoinnovation* (15). The new ideas are “pumped” for three to five years before they are “dumped” to the benefit of another management idea (15). Generally the upswing is characterized by emotional and unreasoned discourse, whereas the downswing is characterized by qualified and unemotional discourse (17). Some studies have argued that this cycle of management ideas is driven by the knowledge producer (i.e., the framework *creator* (15, 18)), whereas others argued that the cycle happens in interaction with the managers who *implement* the idea (19). Pseudoinnovation is reinforced by competition among knowledge producers to predict and use trends in management (20). To launch a theory as new and

different gives the appearance that the knowledge producer stays at the forefront of management development. Abrahamson suggests that scholars should assess the management innovation process critically and intervene against pseudoinnovation to make management innovation more useful (20).

1.3 VALUE FRAMEWORKS IN HEALTH CARE

Different stakeholders define the term *value* differently and in different contexts (21). The focus of this thesis, VBHC (first proposed by Porter and Teisberg in 2006 (4)), is only one of recent developments. Other value interpretations, such as triple aim (22), triple value care (23), value compass (24), value of co-creation (25(p32)), and the vast philosophical literature on value theory have been intentionally omitted from this thesis to retain its focus and clarity. The following section (1.3.1) briefly covers a method used to compare value in health economic evaluation before describing VBHC.

1.3.1 Cost-effectiveness analysis to quantify value

Scarce health care resources force policymakers and professionals to make trade-offs in resource allocation. Such trade-offs can be structurally assessed through health economic evaluations, which focus on relating costs to *results* such as alternative costs, number of procedures, lifespan, or quality of life. *Cost-effectiveness* is defined as the cost of a specific outcome; that is, a specific program's cost in relation to its natural units, such as the cost to find one case of a disease or cost per blood pressure improvement (26):

$$\text{cost effectiveness} = \frac{\text{incremental cost}}{\text{outcome}}$$

1.3.1.1 Outcomes in history and present-day health care

The idea of measuring health care outcomes predates health economic evaluations and its applications go beyond economic appraisal. Codman first proposed outcome measurement in the beginning of the 20th century. He called it the *end results idea* and described it as a practice wherein “every hospital should trace each patient with the object of ascertaining whether maximum benefit has been obtained and to find out if not, why not....The results shall be constantly analyzed and possible methods of improvement constantly considered” (27).

These “end results” came to be known as *outcomes* and gained increased attention following the work of Donabedian (28). The practice was first implemented in clinical settings during the 1980s (1) and later garnered even greater focus (29, 30). The interest led to developing specific models for separating outcomes (e.g., the ECHO model (31) separates outcomes into economic, clinical, and patient-reported categories). Outcomes can encompass any health-related result of health care, such as mortality, functional ability, or pain. As an interesting

side note, a study reported that breast cancer patients preferred the terms *medical results* or *what my life will be like* over the term *outcomes* (32).

This measurement trend caught momentum in rheumatology (33), most notably in the outcome model published by J. F. Fries in 1983 (34) and shown in Figure 1. Fries' model postulated important outcomes in five areas. He proposed measuring the displayed outcomes over time with the goal to minimize the area under the outcome curve.

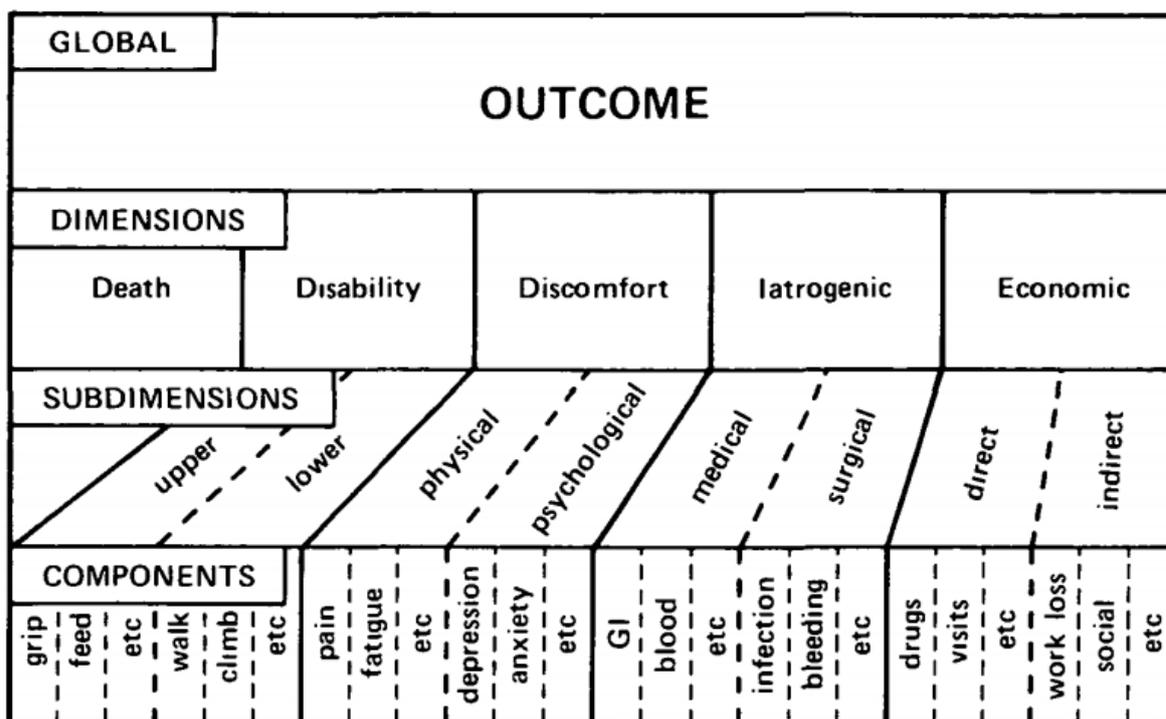


Figure 1. Outcomes in Rheumatoid Arthritis As described by J. F. Fries in 1983. Reproduced with permission from J. F. Fries, The assessment of disability: from first to future principles. *Rheumatology*, 1983;22(suppl 1):48-58(34). Copyright Oxford University Press.

1.3.2 VBHC: What is it?

Value-based health care has been marketed as “the strategy that will fix health care” (10). This overarching strategy of VBHC can be summarized in three principles (6).

First, the goal of health care is value for patients. Health care should focus on outcomes—what matters to patients—and not on the process or the structure of health care facilities. It is argued that every meaningful improvement to the process of health care will be reflected by a measurable improvement in patient outcome (4). This outlook that processes are only instruments used to improve outcomes runs at odds with some researchers' views. For

example, Donabedian argued that care cannot focus only on improving outcomes because many outcome responses are delayed or difficult to collect (35).

To highlight the importance of outcomes, Porter defined value as outcome in relation to costs (4):

$$Value = \frac{Outcome}{Cost}$$

Porter's equation is similar to an inverted version of the health economic evaluation equation (section 1.3.1), but the primary purpose of Porter's equation is not economic appraisal. The ratio outcome/cost is not to be taken literally or calculated into a single number. In fact, the outcomes are seen as multidimensional. Consequently, Porter's equation illustrates the *relationship* between outcomes and costs with the idea that to maximize value, providers should improve outcomes while maintaining or decreasing costs. In addition, Porter asserted that because the goal of value aligns with what is important to patients, it would feel more meaningful to providers to work towards value than towards other potential goals. Professionals would therefore be motivated more by potential value improvement than by other incentives (12).

Second, care is organized over medical conditions and full care cycles. Here, *medical condition* is defined as a patient's state best treated in an integrated manner (4). In practice, this means a medical condition is either a diagnosis including common complications or, as in primary and preventive care, a segment of conditions with similar complexity (11). For each medical condition, instruments should be selected to measure outcome. In VBHC, measurements must be broad in the sense that they should include all parts of care for a certain medical condition. Thus, improvement of care at early stages affects the measurements even if the effect of the improvement is only seen at later stages. Conversely, the measurements need to be narrow in the sense that no measurement should include patients with different medical conditions (4) because outcome and cost are assumed to differ between medical conditions.

Third, costs and outcomes have to be quantitatively measured to assess value. (For a detailed description on how to measure costs, see (36).) Outcomes are here distinguished from other potential measurements, such as process-, structural-, clinical-, or patient-reported experience measurements. This thesis refers to such measurements as *non-outcomes*. (Note that patient satisfaction with care is included in Donabedian's definition of outcomes(35) but excluded here). Details on how to measure outcomes in accordance with VBHC are covered in the next section (1.3.2.1).

1.3.2.1 Three-tier model for separating outcomes

To support choice of outcome measurements in VBHC, Porter developed a separate model (Figure 2). The model includes three dimensions (*tiers*) of outcomes (9): *health status*

achieved or retained, process of recovery, and sustainability of health. Each tier is subdivided into two levels. These categories are not novel. They were applied in evaluations before Porter's publication of the three-tier model (e.g., for migraine (37)). However, Porter made a number of claims in connection with publishing the model. Specifically, he assumed all relevant outcomes fit into the categories; there is a hierarchical relationship between the tiers where tier one is most important; and if one measurement is chosen per level, then improvement in those outcomes will *invariably* decrease costs (38). The specific examples in Figure 2 and Porter's other applications are not meant to be exhaustive. They demonstrate only the categories used to separate outcomes. The arrows indicate causal connections among outcomes. However, it is unclear how this model would apply to chronic diseases (39). Porter proposed an adaptation to the setting of chronic, primary, and preventive care (11) wherein the levels would be preserved but their interpretation expanded.

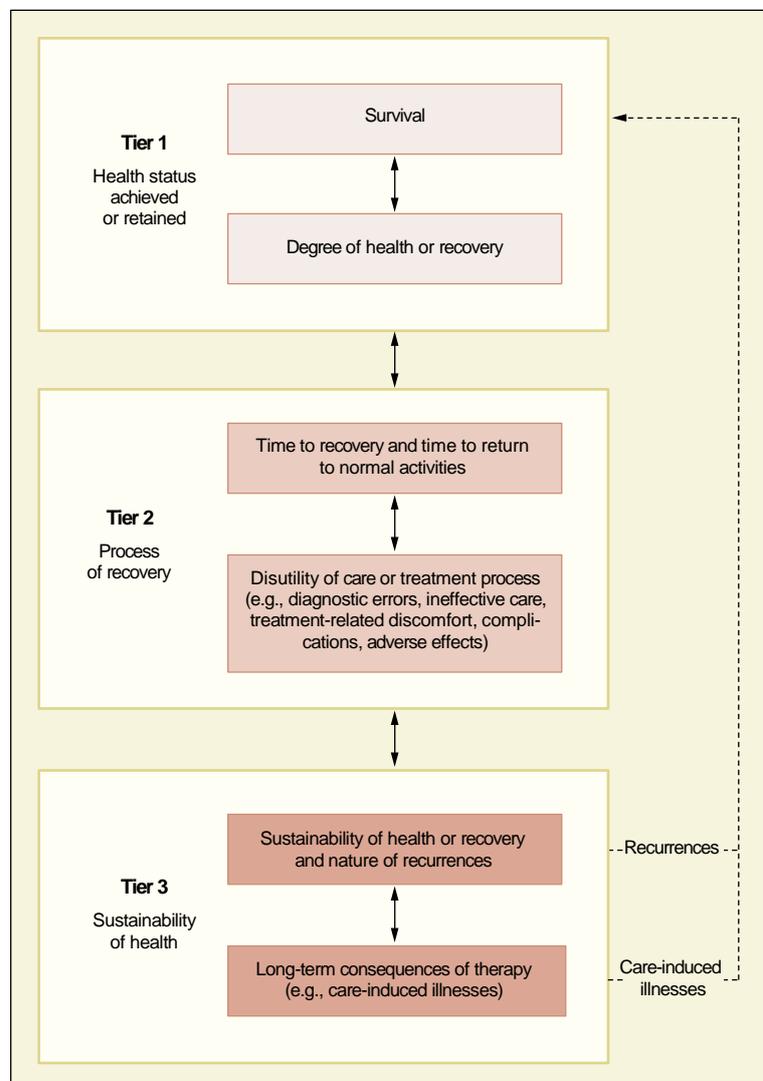


Figure 2. Three-Tier Model for Outcomes. Reproduced with permission from M. E. Porter, What is value in health care? *New England Journal of Medicine*, 2010;363(26):2477-81. Copyright Massachusetts Medical Society.

Table 1 this expanded interpretation of the three-tier model for chronic care settings. When choosing specific instruments to measure outcomes, Porter preferred standardized scales (e.g., 36-Item Short-Form Health Survey (SF-36), Western Ontario and McMaster Universities Osteoarthritis Index, EuroQol Group 5-Dimension Self-Report Questionnaire, Beck Depression Index) when available (38). A few publications about VBHC also mentioned a need to create indicators (4(p128), 12) and incentives to do so (40(p96)), but did not further develop the argument.

Table 1. *Interpretation of the Three-Tier Model in the Chronic Care Setting*

Level	Subcategories suggested by Porter (9) and Porter et al. (11)
Survival	Mortality (1-yr, 3-yr, 5-yr, longer)
Degree of recovery or health	Physical and mental functional status Control of chronic disease complications Pain level achieved Ability to return to work
Time to recovery or to return to normal activities	Time to treatment or definitive diagnosis Time to access specialist treatment for more complicated or urgent issues Time spent accessing treatment Workdays missed or time to return to work
Disutility of care or treatment process	Pain and anxiety before and during treatment Care complications Length of hospital stay Need for emergency department visits or hospitalizations
Sustainability of recovery or health over time	Maintenance of functional level Frequency of urgent care issues Acuity of chronic conditions and complications Ability to live independently Need for revision or reoperation
Long-term consequences of therapy	Long-term side effects or care induced illnesses

1.3.2.2 Comparison of Fries' and Porter's outcome models

Comparing Fries' outcome model (Figure 1) to Porter's (Figure 2), the similarities are striking. Fries separated outcomes in terms of *discomfort* and *disability*, whereas Porter distinguished between the health statuses *achieved* and *sustained*. In addition, Fries' category *Iatrogenic* equals Porter's categories of *process disutility* and *long-term health*. The specific content of these categories are the same. Further, both models recognize the importance of time but deal with it differently. Fries' model measured time as the area under the outcomes curve, whereas Porter's model measured it as separate data points (i.e., time to recovery). Thus, Fries' model aggregated outcome and time into one number and, in doing so, assumed the outcome is a ratio scale. As such, Fries' model more easily applies to chronic conditions that do not have clear endpoint towards which time can be measured. Further, Fries included costs as an outcome, but Porter kept outcome separate from the outcome model. Explicitly explaining this choice, Porter stated, "Treating cost as an outcome will only obscure value by confusing the process of care with the results of care" (41).

Despite the models' similarities, Porter made sparse reference to any earlier work on outcomes in his 2010 publication (9). Assuming Porter was unaware of Fries' outcome model, this similarity could be seen as a sign of reliability (42) or investigator triangulation (43). That is, the fact that two researchers ended up with similar conclusions supports the idea that the models include outcomes *non-patients* believe are important to patients. Whether the models reflect outcomes *patients themselves* think are important is not known.

1.3.2.3 VBHC substrategies to maximize value

Clinical care determines only 20% of health outcomes (excluding genetics). The rest of the outcomes are determined by health behaviors (e.g., smoking and exercise), social and economic factors (e.g., education and income), and the physical environment (e.g., air and water quality) (44). Initially, preventive care was a core part of VBHC (8), but later publications instead focused on clinical care delivery. Thus, when Porter et al. suggested six substrategies to maximize value, their suggestions concerned how to structure clinical care (10):

1. Administratively (and preferably physically) organize care as *integrated practice units* with all resources necessary to treat a medical condition.
2. Measure costs and outcomes, as introduced in the section above.
3. Reimburse health care costs through bundled payments for care cycles. This strategy differs from paying per patient (global capitation) or per service (fee-for-service). Instead, bundled payments are allocated per patient *and* medical condition. (For more detailed discussion, see (45).)
4. Integrate care delivery across separate facilities. This means providers who do not cover the full cycle of care should form partnerships with organizations and locations that cover the other parts of the cycle.

5. Expand services that deliver high value to cover greater geographical areas by establishing satellite facilities that handle less complicated parts of the care cycle. Preferably, health care professionals rotate among these locations.
6. Support the substrategies listed above through a health care information technology platform.

The empirical data presented is limited to cases illustrating each separate substrategy. No case illustrates the combined effect of all substrategies. Therefore, the questions of whether the substrategies interact and would work as expected when put together remains unanswered. Porter asserted they are “mutually reinforcing” (10), but no study has demonstrated a combination of all the substrategies.

1.3.2.4 Tactical and operational tools

Although some specific tools may generally be better at achieving value than others (46), VBHC does *not* prescribe any particular operational tools. Therefore, providers applying VBHC must themselves decide what they should do to improve outcomes without increasing cost. These decisions include, for example, which processes to measure internally; how to manage clinical practice guidelines and human resources; selecting quality improvement tools such as Lean (13), TQM (14), or quality registers (47); and whether to include other approaches such as person-centered care (48).

1.4 VBHC IN A CHRONIC CARE SETTING

The chronic care setting constitutes a challenge when delivering care in general and value-based care in particular. Chronic disease differs from acute episodes in its complexity, need for self-care, and importance of family support (49). Complexity makes it more difficult to standardize and execute care in an integrated fashion. Instead an iterative approach based on interaction between professionals and patients can be more suitable (1). Care is often fragmented and provided by many actors (50); thus, integrating chronic care requires substantial effort (51). This fragmentation encourages short-term goals (51) that could be at odds with the patient’s goal of long-term health (49). To support self-care, approaches such as person-centered care (48), education in self-management, or patient-formed networks to support or mentor each other (52) are applicable.

In VBHC, outcomes are defined in a way that requires an endpoint at which the measurement is performed. Thus, for example, it is not obvious when to measure health status achieved for persons with chronic diseases. Studies have raised criticism concerning how to apply VBHC to unpredictable health conditions or persons with multiple active comorbidities (53), and specifically how to apply it to chronic and palliative care (49). The critiques primarily argued that patients with chronic conditions might value more than sustainability (such as family support), and thus it would be better to focus on the whole trajectory of illness than to emphasize health status achieved (49). Further, questions about how to bridge VBHC with patient-centered care have been raised elsewhere (54, 55).

1.5 VBHC CRITIQUE AND ASSUMPTIONS

The assumptions of VBHC were not addressed in the theoretical framework (4) or later assessed analytically or empirically. As such, no empirical evidence supports the declaration of outcomes in general or the outcomes specified in the three-tier model as all that matter to patients (9). It is common for theories to lack explicit statements of assumptions (56). However, given the lack of empirical data and transparency, it is reasonable to suggest those assumptions are not backed by patient perspectives, statements, or cooperation.

Despite the lack of empirical support, Porter maintained that VBHC is “the best way, and the only way, to drive sustained improvements” (4). If his statement is correct, such a framework would be of great interest, and indeed VBHC has had a tremendous influence in the policy debate. The global organization, International Consortium for Health Outcomes Measurement, is working to operationalize outcomes for a majority of medical conditions. Particularly in Sweden, several hospitals partly implemented VBHC (e.g., Karolinska, Sahlgrenska, and Uppsala University Hospitals). These ongoing implementations mandate a more thorough examination of VBHC.

As with all frameworks (56), VBHC rests on several assumptions not explicitly stated. With the lack of evidence to support VBHC, it is reasonable to wonder if an independent researcher empirically or theoretically assessed any of these assumptions. In addition, a thorough examination of some underlying assumptions of VBHC might provide valuable insight on its scientific status and how to proceed with the ongoing implementations.

A complete analysis of all VBHC assumptions is well beyond the scope of this thesis. The assumptions assessed herein concern the statements that outcomes are what matters to patients (41); the three-tier model specifies these outcomes, whereas the examples made are not meant to be exhaustive (9); and when choosing instruments, “standardized scales such as the SF-36 or the Beck Depression Index are preferable when available” (38). These were addressed earlier in this chapter. Thus, for the purpose of this thesis, the assumptions are interpreted as:

- A. Outcomes, and only outcomes, have intrinsic value to patients;
- B. Outcomes important to patients can be subsumed under the three-tier model; and
- C. Standardized outcome instruments measure the outcomes that matter to patients.

Before proceeding with the aim of this thesis and how it relates to the assumptions discussed above, the next section (1.6) describes a concept important to assess these assumptions: *patient participation*.

1.6 PATIENT PARTICIPATION

At the same time health care struggles to improve care while maintaining costs, patients are gaining influence in health care at a system level (57) and in clinical care (58). Some publications have interpreted the increased influence as causing increased demand for

services (2) and others as a resource that mitigates the strain on the health care system (59). Several frameworks have detailed how to view or interact with patients in health care (eg. 60, 61). Efforts have been made to bridge one such patient framework, patient-centered care, with VBHC. The analysis highlighted conflicting aspects where patient-centered care focused more on patients' rights and needs and, as such, its connections were closer to quality than to costs (55). This thesis will not discuss specific patient *frameworks* and their relation to VBHC. Instead, it addresses the *patient experience* of care and health itself (62) in relation to VBHC.

Patients have actively pursued efforts to gain influence, power, and autonomy on all levels of health care operations (59). Several perspectives have defended the idea of patient influence in care: as a civil rights issue by patient representatives (59), as a consequence of service dominant logic by economists (63), and as a consequence of social cognitive theory originating from psychology (64).

More recently, patient-professional *partnership* (65-69) gained greater attention, emphasizing patients and professionals equally partaking in all parts of the processes in the health care operations. Instead of inviting patients to respond to predetermined questions and evaluate measurement design post-hoc, patients participate from the start. Patients influence questions, designs, and frameworks to the same extent as their professional partners.

2 AIM

The aim of this thesis is to assess assumptions (A-C) of VBHC in a chronic care setting. This overarching aim is addressed in four substudies, each with a specific aim. For clarity, these aims are shown with their corresponding assumption and setting information in Table 2.

Table 2. *Specific Aims Investigated in Substudies I to IV*

Substudy	Assumption	Rationale	Specific aim
I.	N/A	Assumptions of VBHC have not been systematically assessed	To characterize how VBHC is diffused and understood in the health care literature
II.	Outcomes and only outcomes have intrinsic value to patients	Patient initiation of care is valued by persons with RA (70, 71), but is not an outcome and it is unknown if it improves outcomes	To test the hypothesis that implementing a patient-initiated system of care could improve clinical outcome in RA
III.	Outcomes important to patients can be subsumed under the 3-tier model	The 3- tier model has been criticized as less suitable for chronic care (49) in general and RA in particular (72)	To test the 3-tier model against patients' views of value* in a chronic care setting
IV.	Standardized outcome instruments measure outcomes that matter to patients	HAQ is the most commonly used instrument to assess <i>function</i> in RA (73)	To explore how persons with RA experience the use of the HAQ in care

Note. *"Value" here does not refer to outcome/costs, but to the subjectivist notion of "what is important to me." Substudy I investigated how earlier academic literature assessed VBHC and therefore lacks an assumption; Substudies II to IV were designed to empirically assess assumptions in rheumatology (the relevance of this setting is described in Section 3.4). VBHC = value-based health care; RA = rheumatoid arthritis; HAQ = health assessment questionnaire.

3 METHODOLOGY

This chapter briefly summarizes the methods of the respective substudies (Table 3, Section 3.3.1), but puts emphasis on methodological considerations and integration in relation to the assumptions and the overarching aim.

3.1 PHILOSOPHICAL ASSUMPTIONS

Philosophers have suggested several ways of describing what science is (74-77). As Popper famously argued, scientific theories cannot be verified, but only rejected. In his argument, even *one* counter example is enough to reject an entire theory (74). However, Popper's position has been criticized as idealized and disconnected from how science is really done (75). Particularity in social research, empirical findings commonly have multiple interpretations and therefore may only partially support a particular theory (78). Accordingly, in the substudies of this thesis, the word *test* was used both in the Popperian sense as a means *to refute* a hypothesis (II), but also in the meaning of *critically applying* a framework to suggest relevance to a setting (III).

In accordance with Gough et. al. (79), this thesis distinguishes between configurative and aggregative research as it highlights differences between the aims of the substudies. *Configurative* research aims to find a way of understanding data and typically results in a model or theory (79). *Aggregative* research attempts to bring data together to support decisions and is typically helpful when creating societal policy (79). Often, research is both configurative and aggregative. For example, it might aim to structure or separate data in a way that makes it easier to understand, but at the same time inform, decisions. This is called *integrative* research (79).

To evaluate VBHC, this thesis uses the term *research program*, which originated from Lakatos (76). A research program is a set of beliefs about reality and about how knowledge can be gained. In this thesis, the VBHC theory itself is referred to as a strategic framework, whereas the practice and the assumptions behind the strategic framework are referred to as a research program. Lakatos (76) exemplified research programs as Marxism, Freudianism, or the Newtonian view of gravity. Every research program includes a hierarchy of hypotheses. They have a core of central hypotheses and then layers of less stable (*auxiliary*) hypotheses that protect the central hypotheses from refutation. The assumptions of VBHC introduced in the background chapter can be seen as hypothesis in this sense. In VBHC, assumption A would be closer to the core than assumption C, because a refutation of assumption A would have greater consequences for the research program than would a refutation of assumption C. In addition, there are undoubtedly hypotheses in VBHC that are even more important than the assumptions assessed here. For example, a core hypothesis seems to be that "one cannot manage what one cannot measure." This core hypothesis is protected by auxiliary hypotheses such as defining measurements as any kind of quantization. If new evidence threatens to refute this core hypothesis, then VBHC researchers are more likely to revise the definition of measurement than to abandon the hypothesis that only measurable phenomenon can be

managed. As will be subsequently argued in the discussion section, similarly assumption A is affected by how outcomes are defined.

In addition, every research program has a specific method to resolve discrepancies, handle new data, or solve problems within the program. Lakatos called this problem-solving machinery “heuristic” (76). In VBHC, a common way to deal with problems is to look for cases that have solved the problem and then incorporate those solutions into the theory. A contrasting method common in biomedical science would be to pose hypotheses and then try to test their consequences of the hypothesis. As a final characteristic, the research program makes *predictions* based on its hypothesis and problem-solving machinery (76). For example, VBHC predicted that if all levels of outcomes were measured in the three-tier model, then health care costs would invariably decrease (38).

If one research program’s predictions are more correct than those of a competing program, then the program is said to be *progressive* (76). The concept of progression is important in Lakatos analysis because it explains that throughout the history of science, early models and hypotheses have needed time and patience to develop into full scientific theories. It is therefore reasonable, Lakatos asserted, that new theories fail to explain part of the available data. Most scientific theories were developed progressively through iterations while competing with alternative research programs. (76).

In parallel, Kuhn’s (75) paradigm theory was used in this thesis to account for the epistemological stance of the thesis. The paradigm theory is sometimes applied to frameworks of similar complexity as VBHC. It is, in addition, more popular and widespread than the theory of research programs. However, the paradigm theory fails to account for several aspects relevant to VBHC. It primarily lacks an understanding of multiple parallel research programs, structured refutation, and the idea of differences in progression. Some scholars have suggested a “post-Kuhnian” view of paradigm (80) in which several paradigms compete in a way similar to Lakatos’ (76) theory of research programs. To avoid confusion, this thesis uses the term *research program* to refer to VBHC and *paradigm* to refer to an epistemological stance.

The research program’s focus on prediction is at odds with the categorization between *configurative* and *aggregative* research. Configurative research does not serve to make predictions but rather to configure data in a way that *provides a feeling* of understanding. To bridge this disparity, this thesis takes a *pragmatic standpoint*. Specifically, this thesis assumes that the usefulness of a theory is a product of its “problem-solving capacity” (81). The parallels to the problem-solving machineries discussed above are obvious, even though pragmatism does not limit the purpose of a theory to that of making predictions. The pragmatic paradigm assumes the problems of the real world as they appear to the inquirer. It seeks to solve problems with the “best” possible method without considering the epistemological origin of the specific method (82). However, this approach assumes the usefulness of a design is known in advance of performing the study, which can be considerably difficult (83).

In line with this, this thesis views the dichotomy of qualitative and quantitative *methodology* (84) and paradigm (85) as misleading. Instead, references to *quantitative* or *qualitative* in this thesis should be seen as references to specific *data* rather than as packages including ontological and epistemological assumptions. However, for the purpose of communication, the terms qualitative and quantitative methods are used in this thesis to summarize the *methods* used for a particular type of data. This rejection of separate qualitative and quantitative paradigms aligns with pragmatism (81) but is not particular to a specific tradition. The same assertion has previously been made by researchers focused on mixed (86), qualitative (87), and quantitative (88) methods.

3.2 OVERARCHING METHODOLOGY

A majority of management research is framed by a knowledge gap that forms the basis and motivation for the research aim (56). Instead, an aim can be formulated based on *problematization* of the literature and its underlying assumptions. As described by Alvesson et al, “Problematization research typically involves a more narrow literature coverage and in-depth readings of key texts, with the specific aim of identifying and challenging the assumptions” (56). Assessing assumptions is particularly relevant when political factors or a social interest bias govern knowledge production (56). The assumptions are often unpostulated and unlabeled (89). As such, it has been suggested (56) that implicit or weakly articulated assumptions are of greater interest because they are more difficult for researchers employing the theory to notice. Further, assumptions that are not instrumental in explaining known empirical findings are more likely to be false (56).

The problematization method (56) consists of identifying relevant literature, articulating assumptions implicit in the literature, evaluating assumptions in terms of likeliness to be false, developing alternative assumptions through comparison with other literature, discussing assumptions with stakeholders and assessing who will be affected by changed assumptions, and evaluating whether alternative assumptions are likely to generate a well-spread theory. In moving iteratively through these steps, assumptions relevant to the assessment emerge (56).

Problematization of the VBHC and adjacent literature was assessed in the background chapter of this thesis. Although discussions with stakeholders (patients, professionals, civil servants, politicians, management consultants, and academics) contributed to shaping the literature included in this thesis, their respective contributions were omitted. Detailing how this thesis arrived at those specific assumptions and tracing it through all the iterations would have been a study of its own. The method’s final step: evaluating whether alternative assumptions are likely to generate a well-spread theory was not performed in this thesis because the aim was not to produce a well-spread theory or new assumptions.

The problematization methodology does not prescribe how assumptions should be evaluated (56). In assessing the assumptions of VBHC, this thesis took a mixed-methods approach. This decision was a consequence of rejecting the distinction between qualitative and quantitative paradigms. The specific methods were selected, as described from the pragmatic standpoint

(81), simply because they were assumed to provide the most illuminating answer based on the specific aim of the inquiry. Further, as aligned with Tashakkori and Teddlie’s (90) recommendation, the thesis sought to integrate qualitative and quantitative findings and present them in a combined fashion whenever feasible. This highlights contradictions and alignments of findings from different methods (i.e., methodological triangulation). A summary, the hierarchy of the philosophical assumptions, and the methodology are illustrated in Figure 3.

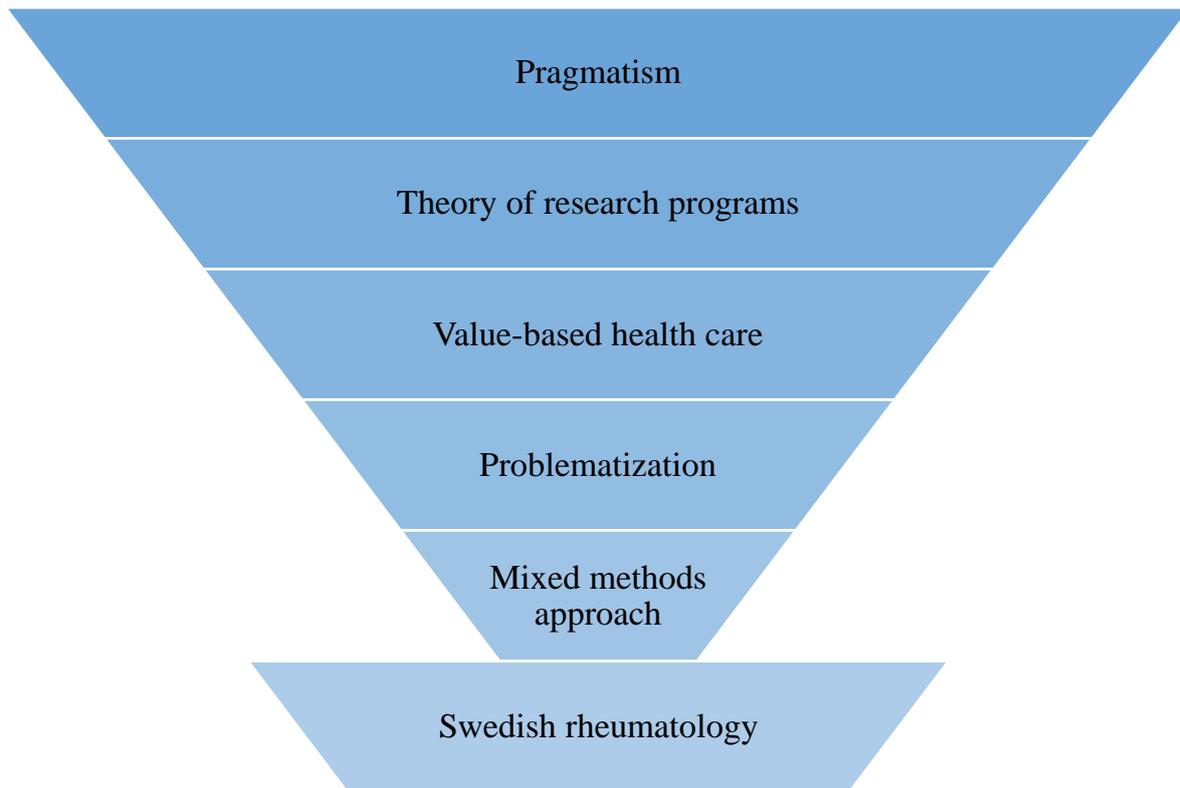


Figure 3. Philosophical and Methodological Positioning of the Present Thesis

3.3 DESIGN

The designs of the substudies are outlined in Table 3.

Table 3. *Outline of Substudies in the Thesis*

Substudy	Perspective	Design	Sample	Data source	Analysis
I	Retrospective Configurative	Sequential mixed methods	255 citations	Literature, citation registry data	Qualitative and quantitative content analysis, non- parametric tests
II	Prospective Aggregative	Randomized controlled trial	131 persons with RA	Blinded assessor, patient reported, laboratory tests	Longitudinal mixed models, non- parametric tests
III	Retrospective Integrative	Qualitative Abductive	22 persons with RA	Semi-structured interviews	Qualitative content analysis
IV	Retrospective Configurative	Qualitative Inductive	40 persons with RA	Semi-structured interviews	Thematic analysis

Note. RA = rheumatoid arthritis

3.3.1 Summary of methods

This thesis used quantitative and qualitative data. Quantitative data were used to investigate causes in an randomized controlled trial (RCT) (II), correlations between understanding of VBHC and journal impact (I), as well as to *describe* citation counts and the study population in the RCT (I, II). Qualitative data were used to capture the understanding of VBHC (I), patient conception of value (III), and patient experiences of an outcome measure (IV). Substudies I to IV were conducted independently, even though Substudies III and IV partially shared data.

In Substudy I, the most cited VBHC article was categorized into conceptual aspects using content analysis. Articles citing the VBHC article were categorized to evaluate how the conceptual aspects were understood according to the structure of observed learning outcomes (SOLO) taxonomy. Substudy II was an 18-month, controlled blinded endpoint, two-center randomized study with 131 RA patients. The intervention group individually initiated

appointments when they perceived a flare. 28-joint disease activity score (DAS28), satisfaction with care, confidence in care, and number of appointments with the rheumatologist was evaluated in comparison with standard care. In Substudies III and IV respectively, persons with RA were interviewed to analyze abductively the three-tier model and to evaluate a commonly used outcome measure. Substudy I did not test a VBHC assumption and is therefore presented first, followed by setting and Substudies II to IV.

3.4 CHARACTERIZING THE MOST-CITED VBHC ARTICLE

A sequential mixed method was developed, and used to characterize how VBHC was understood. The term *exploratory* is often used for mixed-methods studies that start with a qualitative investigation, whereas the term *explanatory* describes those that end with a qualitative investigation (91). This substudy started with a quantitative exploration and then alternated between qualitative and quantitative data, as outlined in Figure 4. It is worth noting that Substudy I took on a broader aim than would have been justified by the overarching aim of this thesis. It not only reviewed assumptions of VBHC in a chronic care setting, but also characterized understanding including and beyond assumptions in *all* settings where VBHC had been applied.

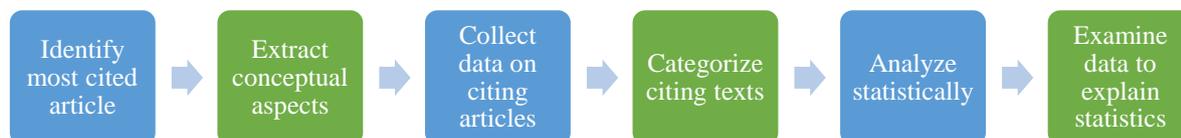


Figure 4. Method Used in Substudy I. Blue signifies analyzing quantitative data and green, qualitative data.

The most-cited VBHC article written by Porter was identified in Web of Science. If citations monotonically correlated to trend impact, then this article could be seen as the main cause of VBHC interest from other researchers. Making this assumption, the most-cited article was referred to as the *trend-starting article*. However, the assumption, and thus this label, are questionable and are not used here. In addition, a qualitative analysis of *all* VBHC articles would certainly have provided more solid data on which article had the greatest trend impact. However, the most cited article turned out to be the main VBHC article on outcomes, and an in-depth analysis thus aligns with this thesis’ focus on outcomes.

From the most cited article, a nominal scale was developed using qualitative content analysis (92), by openly coding the text in the article and then inductively categorizing the codes. The

categories were formed to maximize inner homogeneity and other heterogeneity (82). This analysis rendered five categories in which everything written in the most-cited article could be sorted. It also meant that every article citing this article would refer to one of the five concepts represented in the categories, assuming the citations were correct. In the substudy, the method was referred to as *thematic* analysis according to Braun and Clarke (93). Their step-by-step guide coincides with the method employed. However, in terms of tradition and assumptions, the term *content* analysis would have more appropriately named the process (discussed further at the end of section 3.3.3).

Next, all articles citing the most-cited article (as of 2 May 2014) were classified in terms of concept and understanding. This classification was important to enable analysis of longitudinal trends to achieve a broader understanding of frequencies unattainable by qualitative analysis alone. To enable this classification, text referring to the most-cited article was extracted using a semi-automated text-mining technique that flagged all parts of the articles referring to the most-cited article (NVivo v. 10.0). Quantitative data on publication year and impact factor were extracted from public registries (94). All extracted citations were quantified using the nominal scale developed with content analysis and SOLO, an ordinal five-step scale from education assessment (95). In this phase, the method has a similarity to a meta-analysis method in case studies—quantitative case survey (96, 97)—which seeks to formally code and quantify qualitative cases in order to apply statistical methods.

There are two well-recognized taxonomies for understanding: Bloom's taxonomy (98) and Bigg's SOLO taxonomy (95). Bloom's taxonomy groups learning into categories labeled *remember*, *understand*, *apply*, *analyze*, *evaluate*, and *create*, whereas SOLO distinguishes between *no idea*, *one idea*, *many ideas*, *relate*, and *extend* (99). Thus, SOLO categorizes understanding as *misunderstanding a concept*, *understanding one concept*, *understanding more than one concept*, *understanding the relations between concepts*, or *extending and improving the concepts*. For the purpose of Substudy I, Bloom's taxonomy proved too granular to capture data. For example, it was difficult to distinguish between remembering and understanding based on the citing texts. In addition, the substudy primarily concerned understanding, whereas Bloom's taxonomy also addressed thinking strategies such as *analyzing* and *creating* (98). Because the SOLO taxonomy did not include these, it had a clearer progression of understanding between the levels, making possible treating it as an ordinal scale, and thus performing more advanced statistical analysis.

The data were analyzed with descriptive and inference statistics using medians, interquartile ranges, and nonparametric tests because the scales were nominal and ordinal (78). A χ^2 test for trend was used for longitudinal analysis, and Spearman's Rho when analyzing non-normally distributed correlations. To explain the quantitative findings in greater depth, the coded data were reexamined for additional insight such as what kind of misunderstandings of VBHC occurred, rather than just the frequency of misunderstanding.

3.5 EMPIRICAL SETTING AND STUDY PARTICIPANTS

RA was chosen as a case for analyzing the assumptions of VBHC. This section describes this setting and the study participants. Because the purpose of this thesis was not to develop VBHC specifically within rheumatology, or adjust it to rheumatology, both the empirical setting and the medical condition (RA) are addressed here rather than in the background chapter.

Rheumatoid arthritis is a chronic inflammatory disease with loss of function due to structural joint damage (100). Remission and prevention of joint damage can be achieved with early discovery and treatment (101). The disease affects the ability to work and many other aspects of life, such as functional, sexual, social, and emotional (102). Patients are frequently assessed and therapy adjusted until remission has been reached (103).

Outcome measurements are registered in the Swedish Rheumatology Quality Registry (104) (e.g., DAS28, EuroQol five dimensions questionnaire, Stanford's health assessment questionnaire—disability index (HAQ), work ability, pain according to visual analog scale (VAS)). Patients seeking care enter these data through a computer interface before their visits (105) with a nurse or, if necessary, a physician (106). Interprofessional care and rehabilitation are important for successful care (107, 108). The presence of RA increases the risk of other diseases such as osteoporosis (109), and certain treatments aggravate this effect (110). Other treatments increase the risk of cardiovascular mortality and morbidity (111), serious infections (112), and nonmalignant skin cancers (113). (For the purpose of this thesis, RA including these complications is assumed to be a single medical condition.) These risks mean that patients seem to experience many of the aspects detailed in the categories of the three-tier model, including long-term consequences of care.

Treatment guidelines are based on primarily assessment of DAS28. The DAS28 is a composite index summarizing inflammatory activity in joints and blood with the patient's reported global health. It is calculated using graded tenderness and swelling in 28 joints (excluding the feet), erythrocyte sedimentation rate, and patient-reported VAS-general health (114). Values below 3.2 are considered low disease activity (115); and values below 2.6, disease remission (116). The least clinically significant difference of DAS28 is 0.6 (117). Another frequently used registry outcome, HAQ, is addressed in section 3.2.3.

In Sweden, RA care is highly relevant to questions about VBHC applications. When Boston Consulting Group rated countries ready for a “VBHC transformation,” Sweden came out on top (118). Porter et. al. also argued that the open publication of outcome and process measurements (Öppna jämförelser) already increased value (4(p376-7)). The specific setting of rheumatology was interesting due to criticism of how VBHC applied to the care of persons with chronic diseases (49) in general and RA in particular (72). The rationale in this sense was to assess VBHC in one of the settings where it is most difficult to apply (i.e., using critical case sampling—“If it works here, it will work anywhere” (82(p266))). Due to the specific design of this thesis, Swedish rheumatology is suitable for two additional reasons:

- 1) Swedish rheumatology has a well-developed quality registry that forms a basis for evaluating outcomes. This registry covers process measures, physician-judged variables, and laboratory data as well as patient-reported outcome measures. The registry includes patients in all socioeconomic and demographic groups.
- 2) Rheumatology patients are in the forefront of patient participation, ensuring rich accounts of the patient perspective is available. In addition, rheumatology has a strong tradition in adjusting outcomes to better suit the patient perspective (119).

Study participant demographics are outlined in Table 4.

Table 4. *Study Participants' Characteristics*

Study	Age <i>Mdn</i> (IQR, range)	Number female (%)	Disease activity ^a <i>Mdn</i> (IQR, range)	Disease duration <i>Mdn</i> (IQR, range)	Satisfaction with care ^b <i>Mdn</i> (IQR, range)
RCT (II)	64 (55-71)	41 (61)	3.4 (2.7-4.7)	5 (2-12)	93 (82-97)
Three-tier evaluation (III)	55 (30-82)	17 (77)	Low (low-high)	11 (2-48)	Satisfied (not satisfied-satisfied)
HAQ evaluation (IV)	55 (27-82) ^c	32 (80)	Low (low-high)	11 (1-48) ^c	Satisfied (not satisfied-satisfied)

Note. *Mdn* = median; IQR = interquartile range; *Mdn* (IQR) is displayed for the RCT, and *Mdn* (range) for the evaluations; ^adisease activity in the RCT shows DAS28 and in the evaluations, patient-perceived disease activity on the three-step scale of low, medium, or high; ^bsatisfaction in the RCT shows VAS and in the evaluations, dichotomous patient assessments of generally satisfied or generally not satisfied; ^cage data on one patient and disease duration data on four patients are missing because they were not systematically collected.

3.6 LINKING ASSUMPTIONS, DESIGN, DATA COLLECTION AND ANALYSIS

3.6.1 Outcomes and only outcomes have intrinsic value to patients

To assess the assumption, “Outcomes and only outcomes have intrinsic value to patients,” it was not enough to show that specific non-outcomes are valuable. Specifically, VBHC makes the argument that non-outcomes can be perceived as valuable because they affect outcomes that in turn have intrinsic value. Assessing this assumption was therefore a matter of

assessing if a valuable non-outcome would affect outcomes (as predicted by VBHC). Thus, it was a matter of causality. Will a valuable structural change such as allowing patients to initiate their own appointments (70, 71) cause a change in outcomes? For this purpose, RCT was ideal because randomization mitigates the effect of confounders.

The RCT designed to test this causality was an 18-month, controlled blinded endpoint, two-center randomized trial. Treatment followed local guidelines primarily based on DAS28 assessment. Another frequently used registry outcome, HAQ, is addressed in Section 3.2.3. Patients were recruited for inclusion at change of attending rheumatologist because these were the only patients naïve to the intervention (convenience sampling). Using a computer generated pseudorandom sequence, 131 persons with RA were randomized to intervention ($n = 64$) or control ($n = 67$) groups. The intervention group was guaranteed appointments with a rheumatologist within 10 working days if they subjectively experienced a flare in disease activity. Control group participants were booked for appointments in advance, according to standard guidelines. Independent assessments were performed in the two groups at 0, 3, 6, 12, and 18 months. The DAS28 was used as a primary outcome. In addition, the study collected HAQ, data on satisfaction with and confidence in care measured on the VAS, and the number of visits.

Quantitative data were analyzed using non-parametric tests and linear mixed models. Non-parametric tests are applicable to non-normally distributed data, as compared with parametric tests that assume the data are normally distributed. Whereas parametric tests analyze differences between group means, non-parametric tests examine differences in distribution. Descriptive statistics of normally distributed interval data can be characterized using means and standard deviation, whereas medians and ranges apply to non-normally distributed ordinal scales (78).

The primary outcome of the RCT—DAS28—was analyzed longitudinally. This can be done with mixed model, ANOVA-tests, and χ^2 test for trend. The ANOVA tests cannot use missing data, and the χ^2 test for trend is suitable for nominal scales (it was used when assessing the longitudinal change of citations of VBCH, described in the section above). Because DAS28 was normally distributed and some participants had missing values, a longitudinal mixed model was most suitable. There are several types of mixed models, and here an unstructured mixed model was selected based on an *information-theoretic approach*. That is, the longitudinal model was an average of several feasible models. This average compensated for any model-specific problems and provided the most robust inference. By using this method, it was less necessary to account for influence of individual variables (120). However, to further increase accuracy, the model was nevertheless adjusted by three variables that demonstrated differences between the intervention and control groups when analyzed categorically: sex, disease activity, and age. The Mann-Whitney U test was used to compare ordinal and interval scales, and Pearson's χ^2 test for nominal scales in exploratory analyzed secondary outcomes. Significance level was $\alpha = 0.05$.

In the study plan for the RCT, the primary outcome was the area under the curve of DAS28 over time. The intention was to emphasize the time dimensions of outcomes (similarly to Fries' outcome model, Figure 1). However, this approach failed to consider the problem of missing data. Area calculations under the curve required more than one data point and were not feasible because data were missing. To avoid excluding patients with missing data, the actual DAS28, unaccounted for in the area under the curve, was used as a primary outcome. This way, the longitudinal mixed model could account for missing values and more data could be included. However, the purpose of choosing *one* primary outcome and letting the others be secondary (or exploratory) in advance was to reduce the risk of false positives (Type 1 error), because the risk increases with the number of performed analyses (78). In this case, the area under the curve for DAS28 was never calculated and thus never tested. Therefore, the risk of false positives did not increase even though the final analysis deviated from the study protocol.

3.6.2 Outcomes important to patients can be subsumed under the three-tier model

The assumption, “Outcomes important to patients can be subsumed under the three-tier model” concerns whether patient conception of value aligns with the three-tier model categories. This model could have been evaluated through a patient survey. However, to assess the assumption and intricacies of value conceptions, qualitative data were assumed to provide greater depth in the mechanisms of how potential outcomes interact and in the possibility to explore new hypotheses iteratively.

An interview guide was designed based on previous interviews performed in another research project and literature from the health care management field. The guide was expanded iteratively as new topics emerged. Such an iterative revision was important for the results because the interview guide was originally intended to study patient experience of value in the care process. After six interviews, the guide was broadened to include value outside of care. This change was made because it became obvious that most value to persons with RA was created outside of health care. Although this expansion required a renewed ethical approval and thus slowed reporting of the results, it allowed interesting findings to appear. The interview guide was consequently expanded to include more aspects of life outside of care. Interviews continued until perceived saturation, rendering 22 interviews.

The interviews were analyzed with abductive content analysis (92) in relation to the three-tier model, but thematic analysis (93) was also considered. The distinction between content analysis and thematic analysis is not self-evident in the literature. *Content analysis* is usually concerned with capturing the totality of the data (121) and with the tradition of using categories for frequency counts (93). Capturing the totality of data content is often analyzed by condensing meaning units—an attempt to summarize sentences in a more condensed form but with preserved meaning. In addition, content analysis often distinguishes between *manifest* (“obvious”) and *latent* (“between the lines”) interpretations (122). In contrast, *thematic analysis* does not make this distinction (123). It integrates the data with knowledge

about the setting (93, 123) and generally focuses on larger meaning units such as full paragraphs (93). Further, content analysis often concerns identifying mutually exclusive categories whereas themes are generally assumed to have common denominators (121). Content analysis was suitable for this substudy because the analysis was performed abductively (124) to assess the three-tier model, which consisted of mutually exclusive categories, and to encompass the totality of the data. Similarly, the substudy investigating concepts in the most-cited VBHC article also fit better with content analysis, but for a different reason. In that substudy, where the interest was *manifest content*, it was crucial to include the totality of data. In addition, the developed nominal scale was created to enable quantification.

In pursuit of more fully understanding participant conception of value in the interview data, two additional analytical frameworks were considered alongside the three-tier model: the international classification of function (ICF) (125) and the value compass (24). These considerations were made before the thesis project was solely directed at VBHC. The ICF was omitted because several publications already used it to classify value according to persons with RA (126-128). The value compass was omitted because it separated outcomes into functional health status, clinical outcomes, satisfaction, and costs. Its emphasis on satisfaction itself, costs, and clinical outcomes was less applicable to the data than was the three-tier model. Most notably, the data on costs were very shallow despite costs being a topic in the interview guide, and clinical outcomes were not valuable to the participants.

3.6.3 Standardized outcome instruments measure outcomes that matter to patients

The assumption, “Standardized outcome instruments measure outcomes that matter to patients,” concerns outcome instruments used in care delivery. For this study, the choice was made to look at experiences of measurements in care rather than perceptions of a specific instrument. Although this decision limited the instruments available for evaluation, it focused the evaluation towards measurements seen as feasible enough to be incorporated in everyday care. The specific setting of RA is discussed in Section 3.4. The next paragraph details the considerations made in choosing the specific instrument in RA.

Of the outcomes covered for RA in the Swedish Rheumatology Quality Registry, two were extensively used, agreed with the three-tier model, and have been confirmed to be valuable to patients: global pain and functional level (129). In the registry, these two outcomes are measured with VAS for global pain level and HAQ for functional level. Whereas VAS is unidimensional, HAQ consists of 20 items that encompass different functional aspects. Therefore, valuing functional ability might not be the same as valuing every functional dimension measured in HAQ. For example, one dimension of HAQ covered opening cans; it would be perfectly possible to value functional ability as a whole, while not valuing the ability to open a can. That is, persons with RA might interpret functional ability differently than did the HAQ developer. In addition, HAQ has been suggested as an outcome suitable for VBHC (130). For these reasons, exploring HAQ was deemed to be of greatest relevance to

the assumption and to generate the highest impact in this setting (i.e., high-impact case (82(p266))).

This specific aim was then addressed using the semi-structured interview data collected to assess the three-tier model (22 interviews) and interview data from a separate research project that investigated the role of decision-support systems in the care process (18 interviews). These datasets were selected because both contained detailed accounts of the HAQ. Often, HAQ was brought up as a topic by the respondents themselves, which supported the interpretation of HAQ as an instrument of relevance to patients and thus relevant to investigate. The interview data were analyzed by the thesis author with thematic analysis (93) using the computer program NVivo v. 10.0. The aim was to explore the experience of HAQ in interaction with the care setting. From that exploration, the study captured relevant findings inductively—it became less important to capture all data in the generated themes. For example, some participants had difficulty discussing some aspects of HAQ during the interview; these data were not formed into a theme (such as “difficulty discussing HAQ”) because it would not contribute to the research question. As was reasoned, thematic analysis is generally more applicable when the generated findings have common denominators and do not encompass all collected data.

To aid analysis of HAQ experiences, two separate theoretical frameworks were explored: *institutional ethnography* (131) and *concept theory* (132). These frameworks provided valuable insights concerning the findings. Institutional ethnography highlighted how clinical encounters could be controlled by the communication that HAQ allowed, and concept theory showed how specific functional aspects could become more important because of the attention given to them through HAQ. However, both analyses were removed from the substudy manuscript to adhere to word limits and because they were not necessary for the substudy’s central message.

3.7 APPLYING CRITERIA FOR RESEARCH QUALITY

In problematization methodology, quality is determined by the assumptions selected for problematization. Similar to the pragmatic paradigm idea of *usability*, an assumption is more interesting if it contributes significantly neither to understanding of a theory nor to explaining empirical finding (56). The VBHC assumptions were selected for their lack of theoretical contribution to the overall framework and expected practical implications. This section describes the steps taken in this thesis to ensure the empirical findings were trustworthy. Further, the different research traditions tied to quantitative and qualitative data have their own criteria for quality but have similar traits (133) and will therefore be presented thematically.

The first trait concerns the extent to which the findings agreed with the object of study. This means taking necessary steps to ensure it is reasonable to believe the findings are “true.” When using quantitative data, this is often called *internal validity* (134), which concerns choice of statistical methods and power calculations and avoiding confounders (78). In the

RCT, power calculations (power 80%) ensured a sufficient sample size and the mixed model was adjusted for confounders. In addition, scales with non-equal intervals were treated as such by avoiding parametric tests. When handling qualitative data, the relation to truthfulness is often called *credibility* and concerns choices made to ensure trustworthiness in the interpretation. Realist researchers in general would argue for member checks; that is, asking research participants to verify data and interpretations (133). This member-check technique was employed in the interview study of the three-tier model, where some of the respondents left comments about the finished model. Constructivist and critical researchers in general would instead argue for extended engagement with the data and collaboration with participants in interpretation, respectively (135).

The second trait concerns the extent to which the findings are applicable outside the object of study. In the tradition of working with quantitative data, this is usually termed *external validity* or *generalizability* (134). This validity is achieved mainly by avoiding systematic errors when collecting samples from the studied population (78). The RCT used convenience sampling, which is common in RCTs but limits generalizability. However, this RCT was a two-center trial, which made the findings less dependent on the specific setting of one center. In addition, the RCT employed usual treatment guidelines instead of a fixed protocol. That type of pragmatic RCT closer resembled the real situation of everyday care (“ecological validity”). Qualitative researchers also have the ambition of applying findings outside the study setting. This applicability is often called *transferability* and concerns *reporting* the findings in such a way that a reader can judge if the findings apply in a different context (133). This divide between qualitative and quantitative terminology is sometimes characterized by the terms *contextualized data*, meaning data with a detailed description of the setting, and *decontextualized data*, often referring to randomly selected samples (136). The interview substudies were contextualized by reporting on all factors of the setting believed to impact the findings. Similar to contextualization, reporting of the RCT involved a description of the population (“baseline data”) and setting, as is the quantitative convention (134). Further, despite the divide of terminology, researchers working with quantitative data also attempt to transfer findings between different populations or settings.

The third trait considers the extent to which the researcher affects the findings. Quantitative methods often call this *objectivity* (133) and see it as a product of the methods used—including stating the hypothesis and α -values in advance, as well as blinding researchers, research participants, and assessors (137). Therefore, in the RCT, the qualitative assessment of the joints necessary to calculate the DAS28 was conducted by a blinded assessor. Similarly, the categorization of articles citing the most-cited VBHC article was performed independently by two researchers blinded to each other’s interpretations. In the qualitative tradition, this is often called *confirmability* and can concern actively looking for mistakes and data contradicting the findings (78) or comparing the findings of different researchers (“investigator triangulation” (43)). In the interview substudies, confirmability was ensured by alternating between the perspectives of making the data fit and looking for data that contradict the findings. The literature showed findings similar to the substudy of the three-tier

model and this strengthens confirmability. In addition, the critical tradition has suggested detailed reporting of the researcher's preconceptions ("reflexivity") (135), nowadays used across many qualitative traditions (82).

The final trait addresses the extent to which the findings are affected by factors that are not the object of study. It concerns whether the study is *repeatable*. For quantitative data, this could be measured through *reliability*—the idea of conducting multiple measurements and then assessing agreement (78). Thus, it represents whether the instrument is reliable. For qualitative data, it is more common to speak of *dependability*, meaning close attention is paid to factors outside the object of study and their interactions. (These factors are often called *context*; this thesis instead uses the term *setting*, but in this thesis the terms would be interchangeable). Observations of these factors should be closely reported (138). Some studies have argued that dependability concerns an in-depth description of changes in the *researcher* during the project because the researcher is affected by the study and the context (133). This concept parallels the idea of reliability because some qualitative researchers would argue that the researchers themselves are the research instrument (139), suggesting the major part of the analysis and interpretation take place in the researcher's mind. Thus, consistency in "the instrument" is consistency in study design, data collection, and analysis. Hence, dependability concerns the practice of research, differing from confirmability that concerns interpretation of the findings. Dependability in the present thesis is judged by the reader based on the substudy reports.

3.8 ETHICAL CONSIDERATIONS

This section focuses on ethical aspects of the studies involving research persons. The Regional Ethics Board in Stockholm, Sweden granted formal approval (2008/1226-32/3, 2009/895-31/5 and 2012/1911-31/5).

3.8.1 Informed consent

Study participants made the decision to participate in the substudies without the presence of anyone involved in their care. This separation was important to mitigate the risk they would feel pressured to participate. The researcher provided them information in a nontechnical way, and participants could retract their consent at any time (140). Decisions to participate in the RCT (II) were made with a nurse who was not involved in the patients' care. The decisions to participate in the interview studies (III, IV) were made with the interviewer over the phone after the participants' physicians informed patients in advance that a researcher would contact them. If their physicians (or anyone else involved in care) directly asked for their participation, it could give the impression that declining to participate might affect the physician-patient relationship or result in worse care. In chronic care, patients are often followed by the same physician and, because of this, relationships with the physician can be of extra importance. These considerations were made before the start of the studies, but continuity and relationship were later confirmed to be important in the findings of this thesis (III).

The RCT (II) concerned an intervention in care that could possibly harm patients. This particular intervention put demands on the patient's autonomy and transferred responsibility for care to the patient. Therefore, it was important to exclude persons who, due to impaired judgment or reasoning, could not be expected to assess the information given about the study and might not be able to assess their own health during the trial. In the case of impaired judgement, any consent given would not be considered *informed* consent (140). This can be especially tricky when the research concerns interventions in the care processes because research subjects have been demonstrated to have difficulty distinguishing between care and research (141). In addition, the checkpoints that assessed patients every third month (initially) acted as a safety net and identified any patient who had not sought care despite high disease activity.

3.8.2 Data collection

When digging deeply into experiences of care, accounts of misconduct in care may arise. Gaining this information during an interview could in turn infer a duty to act. However, due to the limited resources available for this research, it would not be possible to engage in such accounts. Instead, participants with such experiences would be referred to other functions in place to handle those issues, such as the Patient Claims Panel. Further, it could be argued that engaging in such accounts could harm the participant because researchers seldom possess the expertise required to act.

The interview guide concerned very personal areas such as death and sexuality. As such, the interview process could evoke bad memories or discomfort. Old memories could be brought back to life with potential psychological suffering. This could be problematic because the researchers did not have therapeutic experience to account and handle the memories and their consequences. Instead, they would advise professional help if any such harm occurred. Further, even though this is an important risk, the research project's expected benefits outweighed this risk.

For reasons of equity, the burden of research should be evenly distributed among research subjects (140). One, for example, should not perform more research on people who are close by for convenience reasons. In addition, such convenience sampling decreases the quality of the findings (82) and makes it less justifiable to take the time of and expose research subjects to risks. However, the RCT used convenience sampling. This was acceptable because the trial could not have been performed by using random sampling of *all* persons with RA.

3.8.3 Data handling

Interpretation of the interviews could affect the patient perspective, mixing views that originated from the researcher with the patients' opinions. This issue concerns research *quality* (discussed above as *confirmability*), but also has an ethical aspect. Qualitative researches often give voice to unheard or marginalized groups, and any clouding of their experience would affect how the group is understood and conceptualized, with potential effects on, for example, stigma. However, the particular population of this thesis—persons

with RA in Sweden—has a strong influence in the policy and health care debates due to their influential patient association. Thus, this issue was handled by a careful examination of the data, assuming different perspectives (III, IV), and asking participants for clarification when necessary (III).

Speaking of “persons with RA” as influential assumes the RA population is relatively homogeneous. However, that is unlikely to be the case. Perhaps subgroups of patients are not represented in the interviews. If that is the case, the research could be a disadvantage to the subpopulations not represented. Presenting the normative and aggregating nature of the data as a model or as themes rather than different narratives could make alternative perspectives less salient. This would not necessarily harm those subgroups physically or mentally, but it could harm their interests in terms of influencing care.

Handling the interview and RCT material involved trade-offs between *integrity* and efficiency, or *validity*. If the only priority were integrity of the research, then only one researcher would handle all the data. In the case of interviews, this would reasonably be the interviewer. However, that would be at odds with the validity gained from sharing data for investigator triangulation (III, IV) or from asking a statistician for advice (II). A further efficiency aspect comes in hiring external transcribers to process the interview data. This thesis used investigator triangulation, statistical advice, and professional transcribers. Integrity was protected by anonymizing data to the greatest extent possible, using secure encrypted data storage, and ensuring only one researcher had access to the identity of the research persons.

4 FINDINGS

The findings, grouped by substudy, are outlined in Table 5.

Table 5. *Overview of Substudy Findings*

Specific Aim	Main Result	Substudy Conclusion
I. To characterize how a nascent management concept (VBHC) is diffused and understood in the health care literature	25% of the citing articles demonstrated no understanding of VBHC; 50% demonstrated a superficial understanding	VBHC may be undergoing a process of dilution rather than diffusion.
II. To test the hypothesis that implementing a patient-initiated system of care could improve clinical outcomes for RA	No significant differences in disease activity, satisfaction, confidence in care, or number of appointments	Patient-initiated care was not better than traditional care in terms of outcomes analyzed.
III. To test the 3-tier model against patients' views of value in a chronic care setting	The 3-tier model lacks dimensions of information, social health, predictability, and continuity.	With modifications, the model is applicable to RA for research, clinical management, and outcome development.
IV. To explore how persons with RA experience the use of the HAQ in care	Persons with RA experience limitations of individual items, search for meaning in a partly faulty summative score, and wrestle with the effects of using HAQ.	HAQ adds perceived value to clinical care, but needs either revising with patients or individualized prioritization among items.

Note: VBHC = value-based health care; RA = rheumatoid arthritis; HAQ = health assessment questionnaire

The studies outlined in Table 5 mainly correspond to the specific assumptions detailed earlier; however, the studies' exploratory aspects make several substudies relevant to *each* assessed VBHC assumption. The substudies, therefore, are below presented thematically, with Sections 4.2 to 4.4 in correspondence with the respective assumptions.

4.1 VBHC IS NOT EMPIRICALLY ASSESSED IN THE LITERATURE

Value is the most commonly referenced concept of VBHC. Of the articles that addressed Porter's most cited article, 51% referred to the concept of value. The proportion of articles referring to different aspects did not change significantly over time. Commonly, the concepts of value were understood at a SOLO level between *no idea* and *many ideas*, with *many ideas* being most common (I).

However, at the time of the assessment of the most cited article, there had been no empirical evaluation of the assumptions or theoretical development, and only five empirical applications. There was no longitudinal change of this trend in the literature (I).

4.2 NON-OUTCOMES CAN HAVE INTRINSIC VALUE TO PERSONS WITH RA

This section concerns the assumption, "Outcomes and only outcomes have intrinsic value to patients."

Outcomes was the best-understood concept in articles stating VBHC trends (I). Specifically, 23% of the articles used the concept of outcomes when citing VBHC. Most authors were able to relate this concept to other concepts (*Mdn* = many ideas; *IQR* = one idea to relate). Studies that misunderstood outcomes (*no idea*), for example, confused it with patient-reported experience measures.

In the interviews, *outcomes* had relevance to patients. When asked about what they valued, persons with RA overwhelmingly referred to outcomes (III). Sometimes, HAQ was even used as a tool to achieve a specific outcome (IV). The HAQ reporting was a way of communicating to the physician that a certain outcome was important to the participant. In addition, others expressed the value of aspects such as continuity, attitude of health care personnel, and time to diagnosis.

Implementing a patient-initiated system of care did not demonstrate improvement in processes or outcome. The primary outcome, DAS28 (Figure 5) showed no significant differences over 18 months ($p > 0.050$). With a 95% confidence interval, the difference was -0.01 to 0.91 at 18 months. The HAQ, patient satisfaction, and confidence in care showed no significant difference throughout the trial or between groups (II).

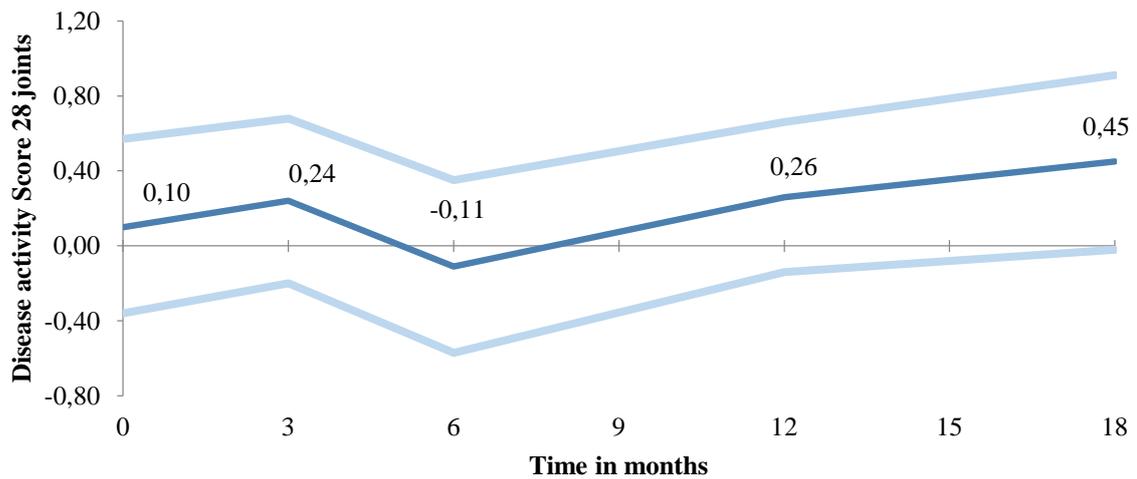


Figure 5. Mean DAS28 in Intervention Group In relation to control, with 95% confidence interval. A positive value corresponds to the intervention group having higher DAS28, meaning higher disease activity. Scores were calculated with mixed model analysis adjusted for sex, disease onset, and age. Black numbers represent DAS28 means at measurement points.

4.3 OUTCOMES NOT IN THE THREE-TIER MODEL CAN HAVE INTRINSIC VALUE TO PERSONS WITH RA

This section concerns the assumption, “Outcomes important to patients can be subsumed under the three-tier model.”

Outcomes important to persons with RA include all aspects of the three-tier model, apart from “Need for emergency department visits or hospitalizations” (III). When the outcome HAQ was used in care, it was sometimes seen as a way to bring the intricacies and complexity of life together into one holistic assessment (IV). However, participants differed markedly on which specific outcomes they valued. For example, some valued survival highly, while others emphasized function—explicitly disregarding survival.

The same person could change the most important aspects between situations. For example, when seeking care for acute problems such as swollen joints, *time to health status achieved* would be most important; whereas when seeking care for chronic ailments, *sustainable function, continuity, and attitude of health care personnel* rose in importance. *Reducing time spent accessing treatment* became particularly important if an appointment had no consequence for the treatment (III). The RCT’s exploratory analysis showed that appointments in the intervention group more often resulted in a change in treatment ($p < 0.001$) (II).

Several outcomes not covered by the VBHC outcome model were important to participants (III). They valued feeling sufficiently informed, social support from family and health care, and predictability of disease. Information was important to accept the disease, mitigate long-lasting questions, and enable self-care or self-control. In terms of social support as an outcome, participants made no distinction between support from family and support from health care. Predictability of disease was important to enable planning everyday life—sometimes it was given more importance than low disease activity.

The VBHC outcome model can be adjusted to better reflect outcomes in relation to the perspectives of persons with RA (III). Participants experienced different outcomes as important at different time points in their care process. The outcomes did not have hierarchical importance common to all. Further, some participants saw the most important outcome as being cured, even though there is no known cure for RA. Figure 6 illustrates the three-tier model from the perspective of persons with RA.

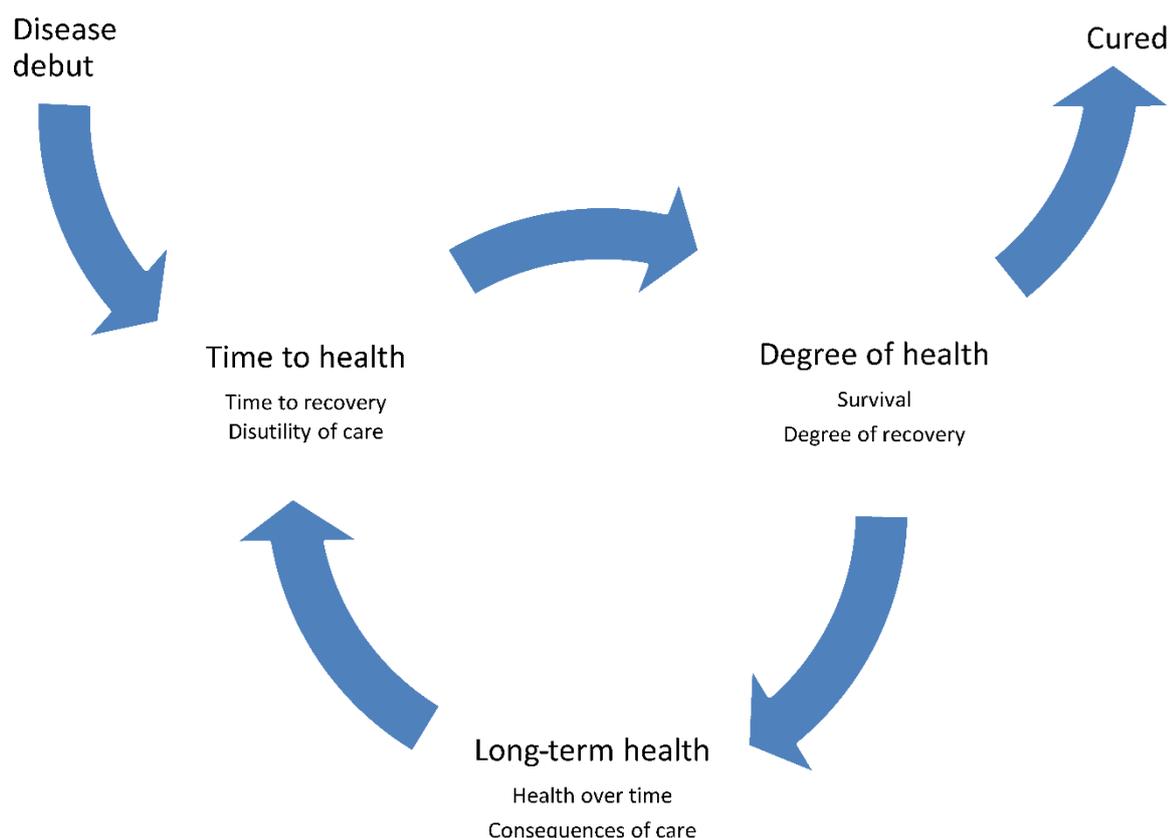


Figure 6. Simplified Version of Outcome Model Developed in Substudy III Illustrates outcomes relevant from the perspective of persons with RA. “Cured” is included in the model because it would be highly valued by participants, even though there is (at present) no cure for RA.

4.4 HAQ DOES NOT ALIGN WITH THE PERSPECTIVE OF PERSONS WITH RA

This section concerns the assumption, “Standardized outcome instruments measure outcomes that matter to patients.”

The HAQ did not capture relevant functional aspects for persons with RA (IV). Participants experienced the instrument as useful and practical but did not think it measured all functional domains important to them. The instrument was important because of the effects it had on care and health perceptions. The three themes and their respective subthemes are outlined in Table 6. Participants reported problems concerned interpreting items, wanting to add new aspects of function, or wanting to remove “unimportant” items (i.e., items they found unnecessary or static). That is, the items either never applied to them or became irrelevant because they were used multiple times.

Table 6. *HAQ Experience for Persons with RA*

	Theme		
	Problem with Individual Items	Meaning of Summative Score	Effects on Care and Health Perceptions
Subtheme:			
	Missing	Capturing holistic perspectives	Effects on physician behavior
	Unclear	Reflecting a temporary state	Effects on understanding RA
	Unnecessary	Requiring strategies for interpretation	
	Static		

5 DISCUSSION

In summary, VBHC had been applied but not tested as of May 2014 (date of substudy I) with any longitudinal change. Further, in the rheumatology setting, the assessed assumptions did not find support. The substudies indicated that non-outcomes can have intrinsic value to persons with RA, and outcomes on the three-tier model level and on an instrument level of HAQ did not necessarily align with the perspectives of persons with RA.

5.1 INTERPRETATIONS OF THE FINDINGS

This section addresses three ways to consider the disagreement between the assumptions of VBHC and the findings in rheumatology: first, as determined by the VBHC stakeholder perspective; second, whether the discrepancy can be explained by how “outcome” is defined; and third, if the findings indicate lack of VBHC progression or pseudoinnovation.

5.1.1 How well does VBHC acknowledge patient perspective?

The developers of VBHC did not explicitly describe their method, and little evidence indicated it was made by or informed by patients. In response, a refined three-tier model suggested from patient interviews (Figure 6) highlights the feedback loop (dotted line in the original three-tier model of Figure 4) and removes the hierarchical aspects. The change better acknowledges that the outcomes important at the onset of disease are also important at recurrences. This is in line with Porter’s idea of the original model: “When recurrences or new illnesses occur, some higher-tier outcome dimensions such as survival, degree of recovery from the recurrence, and so on will also apply to measuring the outcome of these recurrences or illnesses” (38(p4)). The interview findings about value show that different aspects were important at different timepoints from the perspective of persons with RA. From a provider perspective, different priorities at different timepoints in the care cycle might be less obvious because providers treat patients at different stages of the care cycle simultaneously and thus measure all outcomes in parallel.

Lately, studies in non-chronic care settings also tried to integrate the three-tier model with the patient perspective. Similar to the findings of this thesis (III), a study recognized the importance of the “time burden on patient and caregivers, and travel burden for patient and caregivers”(142). They also identified patient-provider communication as valuable, but the participating patients saw it as a driver of outcome rather than as an outcome itself. Correspondingly, participants in this thesis research ranked “attitude of staff” and “communication with staff” secondary to “time to health status achieved” if symptoms were non-chronic (III). However, patients considered these items valuable as an outcome when their symptoms were chronic.

The aspects found important in this thesis (III) include continuity, information, and cure. Those aspects were also included in a recently developed, expanded value definition for primary care (143). This expanded definition encompasses health and “healing” (i.e., health status achieved) as seen in Porter’s three-tier model, but also includes cure and experience of

care (143). *Experience of care* aligns with the categories of continuity and information from the interviews in this thesis (III). In addition, the new definition addresses preconditions support—that is, factors facilitating everyday life such as social care (143). This support was not seen in the analysis of the three-tier model (III). However, when the interviews were analyzed in relation to HAQ (IV), a related finding surfaced about how problems can be solved in different ways—and “achieving the result on your own” was not always the patients’ top priority. For example, the valued outcome was not necessarily “being able to buy groceries,” but “to get groceries.” That is, going to the store would not be important (especially with assistance), but obtaining food would be. This prioritization aligns with the intention of VBHC. The existence of preconditions support such as social care would merely be a structural or process measurement and a means of achieving the true end (i.e., the outcome).

Obviously, alignment with patient perspective is important to a framework that claims to focus on what really matters to patients. However, VBHC was not created primarily to increase patient influence or participation. Instead, its purpose was to solve the “cost crisis” and lack of quality improvement in health care (4). This thesis does not address those aspects of VBHC. Porter, originally a competition researcher (144) specializing in strategy, naturally addressed health care as an industry (145) from the provider (and payer) perspective. This can be counterintuitive because traditionally, public organizations focus more on cooperation than competition (146) and not all market principles apply to health care (147). However, a strategic approach could provide important insight, even though it affects the focus of VBHC. To illustrate this provider focus, Porter wrote, “Patient use [of outcomes] is secondary. The most important benefits of outcome measures are improved value in care delivery and information for referrals to other providers” (6). He argued that consumers (i.e., patients) could not use outcome information efficiently and instead medical teams would create improvements based on outcomes (12). In further examples, prevention has a peripheral role despite reasonably being the most cost-effective way to improve outcomes; in terms of experience, the focus is on provider experience rather than patient experience; self-care has no place in the framework (to date); the costs addressed concern only provider costs and not, for example, patient fees or patients’ lost working time; and so forth.

In addition, the lack of alignment can also be seen in HAQ, the instrument studied in this thesis (IV). The findings show that applying readymade instruments does not guarantee measuring patient-perceived value (i.e., face validity in the care context). The instrument can still be valuable for other purposes such as research, but validity in research does not necessarily imply validity in care. In addition, instruments in research are analyzed on an aggregated level, whereas outcome instruments for care are used for individual patients.

Porter previously argued that the instruments can be changed along the way, as long as organizations established the habits of learning and improvement (4(p136), 12). However, it is difficult to imagine by what mechanisms professionals are supposed to discover that the instruments they used for so long lacked relevance to patients. On the contrary, if the

instruments were used extensively in research without professionals evaluating the value to patients, then it would be doubtful this would change because the instrument was also introduced in care. In addition, the more data collected with an instrument, the greater advantage the instrument will have in terms of historical and international comparability. This thesis does not address any setting other than rheumatology, and then only one instrument in rheumatology. However, considering this counter example, one cannot assume alignment between instruments and the patient perspective. It would be reasonable to conclude that showing agreement between patient perceptions and outcomes on a model level (that is, illustrating contextual relevance of the three-tier model) is not sufficient evidence to build VBHC. Agreement with the patient perspective should be shown also on an instrument level.

Alignment of the patient perspective and outcomes measured could also have consequences for the motivation of professionals. Porter asserted that measuring outcome makes work more meaningful to professionals because they see the improvement they create for the patients. This might be the case, but it requires measuring outcomes important to patients, and not only those endorsed by professionals. An implementation study in Sweden showed a risk that professionals understand VBHC as measuring what *they* think patients should value (148). Professionals' motivation to improve outcomes can be considered less important when they are not motivated by outcomes with actual value to patients.

5.1.2 Are only outcomes important to persons with RA?

Donabedian's idea of sectioning data into structure, process, and outcomes (28) had a deep impact on management research. With its profound focus on outcomes, VBHC argues that structure and process are of importance only if they lead to improvement in outcomes. The RCT in the present thesis showed that a structural change of value to patients did not correspond to an improvement in outcome. Neither secondary outcomes nor patient satisfaction showed improvement. However, specific instruments to measure patient participation, such as *patient activation measure* (149) were not used during the trial. Although change in patient activation could be seen as an outcome or "result" of care, changing the system to patient-initiated care is not an outcome. It is a structural aspect.

Similarly, aspects valuable in the interviews (continuity and attitude of health care personnel) (III) are traditionally considered process measures (150, 151). In the substudy, the argument was made that these aspects in themselves had such a profound effect on participants' everyday life, they should be considered outcomes for chronic diseases. For example, lack of continuity would bother the participants beyond being a mechanism for worry or anxiety. With the same line of argument then, does this imply even patient initiation could be seen as an outcome due to its potentially profound importance in everyday life?

The interview data do not show this result and reasonably, if this were the case, the *profound effect* would be the outcome not the patient initiation. However, this approach to redefining *outcome* means making the three-tier model, by definition, applicable to everything patients value. Indeed, as Porter asserted, the specific dimensions in the three-tier model were not

intended to be exhaustive (38). Nevertheless, it raises the question whether it is *useful* (152) to protect the core hypothesis (76) by adding auxiliary hypotheses (e.g., the additional hypothesis “continuity is an outcome in RA” protects “only outcomes are valuable”). Perhaps the data in the present thesis would be better represented by rejecting the assumption that persons with RA value only outcomes.

In contrast to VBHC, the co-creation framework provides an alternative perspective on outcomes. In this framework, the co-creation experiences of individuals—their experiences of the process or of the outcome—determine value (25(p32)). This would mean that value also entails organizational improvement and personal, economic, or societal gains that are not limited to optimizing health *outcomes* (153).

In addition, if one instead accepts the assumption that only outcomes have importance, the wide confidence interval in the RCT (II) and the (perhaps consequent) lack of effect illustrates how difficult managing based on outcomes can be in practice. Measuring impact on outcomes requires big data materials. The failure to show statistical impact parallels Donabedian’s assertion that relying solely on outcomes is not practical in the management of everyday care (35).

5.1.3 Is VBHC a progressive research program?

The lack of development, empirical tests, or critical application of the framework means VBHC is not currently advancing its position in relation to other strategic frameworks for health care. The empirical consequences of VBHC (i.e., predictions made by assumptions when formulating the specific aims) do not correspond to empirical findings in RA for the assumptions studied in this thesis. Considering that VBHC targets a generally evidence-based arena (i.e., health care), this section discusses what these findings and the current literature suggest for VBHC’s scientific status.

Using Lakatos’ terminology (Section 3.1), this thesis could not show that VBHC is a progressive research program. Progressiveness, according to Lakatos, is the criterion that distinguishes science from pseudoscience. Does this mean VBHC is pseudoscientific? This thesis has not assessed all predictions made by VBHC; it has only showed that no other academic work has assessed prediction, and the thesis itself has failed to find support for predictions based on three critical assumptions of VBHC. In addition, progressiveness itself is not important, but how progressive the program is relative to competing programs. Certainly other programs, such as the broader quality improvement in health care, have seen progression and critical application (154), but other solutions with complexity similar to VBHC, such as Lean or TQM, have also seen degenerating popularity (15). The assessment of which is the “most progressive” program is therefore not obvious.

5.1.3.1 Developments in VBHC after the research project start

The trend in the VBHC literature (I) suggested VBHC would see no development in the upcoming years. However, a few developments of VBHC have been suggested recently,

some of which were discussed above in relation to the patient perspective. In addition, one theoretical analysis suggested measuring different aspects on a population level rather than on an individual level—to measure chronic diseases with an emphasis on disease severity on the population level and patient experience on an individual level (155). Another study based on empirical data suggested developing VBHC by merging it with operational concepts. The purpose was to develop an instrument to assess changes when implementing VBHC (156). Their operationalization addresses all VBHC substrategies except spreading services geographically. As the only empirical study outside of this thesis (to the author's best knowledge) that sought to develop VBHC, the discussion and an empirically founded development of VBHC were most welcome.

However, despite the few recent suggestions, an argument could also be made that a lack of progression might be necessary in VBHC's early stages—an argument that aligns with the position of the VBHC publications. Porter argued that we need to pass beyond the current focus on process measurements as soon as possible (9). The finding that participants in the present studies care *primarily* about outcomes (III) supports Porter's argument. However, the relatively low use of outcomes compared with other measurements does not represent what patients value (e.g., among indicators in National Quality Measures Clearinghouse, only 8% were outcomes (157)). In line with this argument, it would be reasonable to accept lack of progression because testing the theory would result in lost implementation momentum. That is, because it is *so important* to measure outcomes, there is no time to improve the theory, test the framework, or think critically about the details. Despite the focus on criticism in this thesis, the interviews (III to IV) found support for measuring outcomes. In addition, early research programs require time to establish themselves (76). That is, VBHC requires patience towards its progressiveness—considering it was launched in 2006 (4). Although development of the theory is less visible, it seems applications are at full scale. The next paragraphs summarize recent applications.

5.1.3.2 *Applications of VBHC after project start*

Application of VBHC has mainly concerned developing outcome scorecards. Longitudinal invariance in understanding of the most-cited article led to the conclusion that the VBHC concepts would continue to be misused. Since the time of the review of articles citing the most-cited VBHC article (I), value and outcomes continue to be used improperly. For example, studies did not collect data on any outcome mentioned in the most-cited article (158), use a unidimensional outcome in relation to costs (159), or consider adherence to medicine as an outcome (155). Others confused clinical measurements with outcomes (160). In addition, an implementation study of VBHC in Sweden showed misunderstandings of VBHC among managers (148) similar to the ones seen in the literature.

However, outcome scorecards that align with the three-tier model have also been developed with extensive patient participation (e.g., head and neck cancer (161) or breast cancer (32)) or without patients (e.g., in vascular surgery (162), bacteremia (163), spinal surgery (164), Parkinson's disease (164), and thyroid cancer (165)). Interestingly, Fayanju et. al noted that

patients who were asked for outcomes named what the authors called “non-outcome entities,” such as “experts” or “CT scan.” That is, when asked for outcomes, patients named entities the three-tier model does not consider outcomes. However, because the authors used the three-tier model rigidly, the studies discarded these results as not applicable to the model. Hence, those results were not included in their study findings (32). In contrast, the present thesis study instead criticized the three-tier model (III) when it did not agree with the patient perspective.

The nongovernmental organization, International Consortium for Health Outcomes Measurement, has facilitated several outcome scorecards for VBHC. However, the scorecards they published thus far were not separated on the three-tier model tiers. In addition, the scorecards prioritized scientifically rigorousness with an emphasis on standardized instruments rather than patient perceived value. So far, they have included only one to three patient-participants in a much bigger committee of professionals (e.g., 166-169). Nevertheless, their outcomes have already been employed in care evaluation (e.g., for prostate brachytherapy (170, 171)).

These implementations and the argument preceding them suggest there is a disproportion between critical assessment of VBHC and the extent of implementations. Progressiveness, according to Lakatos, concerns whether changes of VBHC lead to predictions that are more accurate. From this follows that in order to be progressive and hence scientific, the research program must be ready to revise itself. Obviously, *scientific* theories are preferable in health care. Even though the literature citing VBHC showed no trend in testing or developing VBHC (I), there is still time for VBHC to pursue progressiveness.

5.1.4 Is VBHC a pseudoinnovation?

Leaving the rather abstract question of VBHC’s scientific status for a more empirical analysis, the disproportion of implementation and critique suggests an uncritical implementation of VBHC. This characterizes *pseudoinnovation* (17). Implementation of pseudoinnovations generates small technical benefits to the system; therefore, uncovering pseudoinnovations is important (20). There are no fixed criteria for categorizing innovations as pseudoinnovative (15). This section will therefore only address aspects *indicating* that VBHC is a pseudoinnovation.

Although the literature on VBHC refrained from positioning itself in relation to other strategic health care frameworks, several striking similarities with earlier research literature emerged in the background chapter. For example, similarity exists between Fries’ 1983 outcome model (Figure 1) and Porter’s 2010 three-tier model (Figure 1), as well as the value equation (Section 1.3.2) to the cost-effectiveness analysis of classical health economy (Section 1.3.1).

However, one could argue that the differences constitute improvements in making the models more practically relevant or easier to apply in daily clinical practice and health care governance as opposed to the other frameworks used more often in research. Instances

include the decision to separate costs from outcomes and highlight the relationship between them, or the much more intuitive and practical equation of outcomes/cost compared to the policy-centered and abstract equation, cost-effectiveness = cost/outcomes. Nevertheless, other VBHC concepts are largely the same as ‘competing’ concepts. For example, the concept of cycle of care is similar to clinical pathways (172) and value stream in Lean. If small differences make the framework more practical than its alternatives, then this implies the focus should be on these differences in order to bring innovation through VBHC in implementation.

5.2 A NEW SUBSTRATEGY MOVING FORWARD?

The critique of VBHC in this thesis concerned relevance of the outcomes to the patients and progressiveness of VBHC. Because patients are the only source used for assessment of VBHC assumptions in this thesis—and simultaneously the target of most health care development—the critique can be subverted by including patients in VBHC implementation. Perhaps it would be reasonable to add a seventh substrategy to the ones described in section 1.3.2.3: “New measurements (and care processes) should be developed in partnership with patients.” There is no *theoretical* conflict between working in partnership with patients and working with VBHC. However, it could be that patient perspectives and VBHC assumptions do not agree in *practice*, as seen in the RA setting in this thesis. This suggests that if partnership was included as a substrategy, there would be trade-offs between the substrategies (e.g., patient perceived value and measurement of outcomes). Although adding this substrategy eliminates the critique of this thesis, the feasibility of this seventh substrategy is, like the feasibility of the other substrategies, a suggestion that needs empirical study.

5.3 TRANSFERABILITY

The substudy designs and findings closely tie to patient perspectives characteristic in chronic care settings. Patient initiation of care (II) requires, among other things, that patients are able to notice when they should seek care and have had the medical condition for a sufficiently long time to learn the characteristics of their disease. In RA, a disease flare is noticeable to the patient; other chronic conditions (such as diabetes or high blood pressure) require medical equipment for patients to judge their disease status. Further, acute diseases might be difficult to arrange with patient initiation in a safe manner because recognizing a flare might require experience. In addition, the interviews highlight continuity, attitude of health care personnel (III), understanding the meaning of HAQ, and the value of holistic assessment (IV). Those aspects might be more important in chronic diseases compared to acute episodes because chronic diseases have a greater need for coordination and complex interventions (51). Further, the emphasis put on life quality over life expectancy (III) can also be seen in other incurable diseases (173). However, the patients’ rich opinions about HAQ seemed to be affected by longitudinal involvement with the measurement and its impact on care. The influence of measurements is prevalent in the Swedish rheumatology setting (105), but it is (currently) rare in other settings. Swedish registry based data collection is supported by high degree of generalized trust, unique personal identification numbers and publicly funded

health care (174). For this reason, Swedish rheumatology was suitable for the aim of this thesis, but transferring the findings to other settings could be considerably difficult.

5.4 STRENGTHS AND LIMITATIONS

This thesis tackles VBHC from several methodological stances. Specific methods were chosen based on their applicability to specific aims, and multiple methods were only used in Substudy I. The general limitations of the specific methods are mitigated by several approaches (i.e., intermethod reliability (42) or methodological triangulation (43)). However, the findings from the interviews are affected by the chosen method. That is, analyses of qualitative data aim for depth but reduce attainable breadth. Although depth was the purpose of the interviews, an exploratory mixed-methods design (where the interview results are used to construct a survey) would have generated greater generalizability. Similarly, the substudy evaluating the three-tier model (III) compared the results with qualitative studies. However, quantitative studies of relevance to the three-tier model were not addressed (eg. 175). Even though comparing qualitative findings with quantitative studies can be very challenging, it likely would have generated additional insights.

In tackling VBHC *assumptions* rather than *implementation*, different knowledge was gained. The overarching question of whether VBHC works as intended is empirical and could best be answered by studying the consequences of implementing VBHC. Such a study was not included in the present thesis, primarily because data on controlled implementations or access to multiple cases for cross-case analysis is rare. Further, such implementations in Sweden would hardly be transferable to other settings at present, considering the advantages the Swedish setting has in implementing VBHC (118). Instead, assessing the assumptions—as in this thesis—is more feasible. Further, such assessment provides knowledge of the mechanisms of the theory and may give insight into flaws that would not become obvious in a noisy empirical case study. As such, the theoretical assessment complements an implementation study. In conjunction with this assertion, the setting of rheumatology serves as a setting for all the assessments made, although multiple settings would have provided greater transferability or generalizability. For this reason, one must be careful about transferring any of the findings to other medical conditions or to non-Swedish settings. In addition, this thesis does not *fully* cover the rheumatology setting or even RA. For example, this thesis does not address instruments commonly used for outcome assessment in RA (other than HAQ). However, the purpose of this thesis did not aim to adjust VBHC to the RA setting, but rather to show how assumptions made by the theory are not *universally* true.

Although previous studies criticized VBHC for lacking patient centeredness (54), no study systematically assessed VBHC in relation to the patient perspective as extensively as this thesis. This study covers only outcomes in VBHC but is, to date, the most comprehensive assessment of VBHC in relation to the patient perspective. The RCT and interview studies (II to IV) of this thesis were informed by, not coproduced with, patients. Critical theory argues that it is not enough to *ask* participants, because their perspectives would be framed by the researcher-chosen methods and questions (135). A full *partnership* with patients is seen

as the best way to mitigate this risk. Although this thesis was not designed from the perspective of critical theory, the criticism still holds.

The outcome used in the RCT—DAS28—is the most commonly used global assessment outcome in RA trials (73) but has not been confirmed to be valuable to patients (129). The RCT showed no improvement in HAQ either, but as demonstrated from the interviews (IV), HAQ does not capture functional aspects of value. When this thesis assessed the assumptions, it assumed they were independent. This means that the substudies designed to assess the assumptions assumed the other VBHC assumptions are true. When the studies were designed, it was not known that the assumptions would not find support. This approach neglects any interaction effects and means the substudies results are not connected to only one specific assumption. The lack of effect on outcomes in the RCT could also be explained by the fact that the outcomes measures were not in fact outcomes valuable to patients. This implies that the intervention actually may have improved outcomes after all, but the right outcomes were not measured. In retrospect, it is difficult to suggest what would have been a suitable primary outcome for the RCT, as the commonly used outcome measures, DAS28 and HAQ, are not valuable to persons with RA. The EuroQol five dimensions questionnaire, measured in Swedish rheumatology and among the instruments suggested by Porter, was not studied in this thesis. However, a review from 2013 suggested patient reported health related quality of life is not affected by patient-initiated appointments (70).

The findings in the RCT (II) did not fulfill the criterion of non-inferiority analyses (176). The *substudy* conclusion (i.e., that the intervention is as safe as standard care) is based on lack of significant difference in DAS28. However, the risk of false *negative* when assessing only the *p*-value is 20%, because it corresponds to the statistical power of 80%. The 5% significance level only determines the risk of false *positives*. This is relevant because the significance level required to implement an intervention is usually $p < 5\%$. Instead, under the principles of a non-inferiority analysis (176), the *p*-value is disregarded in favor of the 95% confidence interval. In the RCT, the confidence interval was -0.01 to 0.91. This entails both 0.0 and 0.6 (least clinically significant effect of DAS28 in RA), which means the study sample is not large enough to conclude that patient-initiated care was *not inferior* to standard care. It only demonstrates that it was *not better* than standard care. Despite this limitation, the substudy still supports the argument made in this thesis concerning the assumption assessed, because the purpose of this thesis was only to show that patient-initiated appointments were not superior in outcomes.

A parallel qualitative analysis of the RCT intervention would help understand the underlying mechanism of the intervention. Based on the data from the RCT, it was difficult to explain why it did not demonstrate effects on DAS28. A concurrent qualitative study employing observations or interviews might have rendered explanatory data on the mechanisms active during the RCT.

To develop the three-tier model (III), an approach closer to grounded theory (177) would have been more applicable than was content analysis. A grounded theory approach could

have led to a more elaborate development of the hypotheses that emerged late in the analysis process (e.g., time spent on self-care as a consequence of care, and whether time spent accessing treatment should be moved to consequences of care or included in both time to health and consequences of care). The interview guide was developed iteratively, and the first six interviews completely analyzed before subsequent interviews were performed. However, as saturation was reached, the deeper analysis (and re-analysis of interviews) was not conducted until all the data were collected. A fully iterative approach and iterative model development might have been more suitable. However, “by-the-book” grounded theory (177) is not suitable to model evaluation because grounded theory strives to build a model without preconceived ideas.

6 CONCLUSION

This thesis shows potential limitations concerning assumptions about outcomes made by VBHC in an RA setting and simultaneously illustrates sparse empirical development of VBHC. In RA, aspects other than outcomes can be important, the three-tier model does not address all important outcomes, and instruments are not necessarily aligned with important aspects of outcomes. Although the empirical evidence in this thesis is not sufficient to suggest specific adjustments to VBHC, it is obvious more research should precede full-scale VBHC implementation. Researchers seeking to develop VBHC could start with the assumptions assessed in this thesis because these areas seem to lack alignment between what VBHC assumes and what persons with RA value. Further, studies of assumptions not assessed in this thesis and consequences of implementations are warranted.

The criticism of VBHC in this thesis concerns outcomes in relation to the patient perspective (using RA as a model for chronic disease). This means the criticism would fall if VBHC were implemented in partnership with patients and with a willingness to change aspects of VBHC based on the patient assessments. Such a change would correspond to adding a seventh substrategy to the six described by VBHC: “New measurements (and care processes) should be developed in partnership with patients.”

6.1 IMPLICATIONS

6.1.1 For patients

Outcomes of VBHC are not aligned with what persons with RA perceive as valuable. This might be the case for other chronic medical conditions as well. Patients have reason to assess outcomes suggested by VBHC critically, asking themselves, “Is this *really* valuable to me?”

6.1.2 For policy makers and managers

The VBHC is often misunderstood and lacks empirical evidence. Any implementation should be carefully studied and performed in partnership with patients. Large-scale implementations should be avoided until more evidence is gathered.

6.1.3 For healthcare professionals

Outcomes currently used in care might not be relevant in improving value as perceived by patients. In the RA setting, continuity of care and attitude towards patients are still important, contrary to what VBHC asserts. This might be the case for other chronic diseases as well.

6.1.4 For outcome researchers

Instruments developed for research might not be relevant to patients for use as outcomes in care. Achieving relevance requires assessment of standardized instruments and, if necessary,

development of new instruments together with patients. Many medical conditions lack outcomes known to be important to patients.

7 ACKNOWLEDGEMENTS

First of all, thanks to all the participants who made this research possible! Your willingness to share your experiences and participate in experimental care processes has been mind-blowing.

My main supervisor, Helena Hvitfeldt Forsberg: You took me in early on and this really made a difference when I started out as a researcher. You helped me see that a PhD can be managed as one would manage any project. There were times in this project when I would not have made it through without you, and I'm very happy for the support you showed.

My co-supervisor, Anna Essén: Thank you for all the exciting discussions about theory and management. You always respected my opinions, treated me as though I were a senior researcher, and listened closely whenever I asked you to. As I have realized throughout this project, my heart lies much more in theory than in practice, and therefore it has been delightful to have you as a guiding light and role model in the kingdom of science.

My co-supervisor, Staffan Lindblad: Thanks for all the tips on how to navigate the administrative landscape. Even when you are super busy, your enthusiasm is always there. I remember when we first planned the thesis and I asked you if it was good enough for you—you said, "I'm just happy I can be part of this."

My co-supervisor, Sofia Ernestam: Thanks for picking me up for my master's thesis. Even though your heart is primarily in development, you spent time with me trying to detangle the research challenges. I truly appreciated that. A creative mind like mine needs to be balanced by a realist and your competence on what was doable in practice was very helpful.

My mentor, Dan Grandér: You've had the ungratifying function of picking up the pieces when I was torn apart. We have not met very often, but it has always been very meaningful to me when we did. Thank you for a lot of wise and thoughtful questions and advice. It's been a privilege having you as a resource during this project.

My funders, Stockholm County Council, Vetenskapsrådet and Forte (the Swedish Research Council for Health, Work Life and Welfare: 2012-1688, 2014-4238): Thank you!

Carolina "Stubinen" Wannheden: Awesome roomie! Thanks for your patience with my complaints, pragmatism (when wisdom was not enough), and humor. We started out as roomies and happily for this final marathon, we are back there again. Thank you!

The research group, Elena, Åsa, and Julia: Best colleagues I could imagine. Such a fun crew to be around and very good support when the walls were falling.

All my PhD colleagues. Extra thanks to: Jens-Jacob (one of the few who "got me." Man, you saved me!), Vibeke (thanks for being the most relaxed PhD ever), Caroline L (thanks for the *invaluable* humor and the company at the coffee machine when I was new at LIME), Sara T (my Alvedon® in the PhD world. Thanks for your ear and thoughts), Carl (the most creative

management researcher I know), Pamela (known from TV – thanks for emotional support, ideas, and the wicked humor), Emma (thanks for making teaching an absolute delight, you were just what I needed in a teaching partner), Sara R (you taught me more about what it means to live with a chronic disease than any study I've ever read!), Sara K (few people pull off being wise and cool at the same time, but you are the proof of concept), and Abb & Linda (the LDN crew – thanks for great brainstorming sessions).

My first contact with science, Melissa Nordström: You showed me science in a vastly different area than the topic of this thesis – the patience you showed with my clumsiness was very important in my decision to pursue a PhD. Thank you!

KVALNET: Hallelujah! Such a resource. You've given me more detailed feedback than anyone I've worked with. It's been a pleasure - Matilda, Per, Hedvig, and Nicklas – I owe you everything!

My non-PhD colleagues Cecilia, Elin, Karolina, Inga, Lotta, and Maud. Thank you for all the fun coffee breaks. Ja, gud ja (i London), tänk nu är det klart! Sandra, one of the kindest I've met – and a great course leader. In the branch of teaching, I should also mention Therese – great co-teacher.

As every manager knows, nothing can be achieved without support functions: Mats, Gert, Ingrid, Pia. Thank you!

An extra shout out to those who read drafts of this thesis: the research group, KVALNET, Anna, Mats, Pamela. Thank you!

Also, thanks to Michael Porter who made this work possible. I hope you know that the only reason I would spend this much time critiquing VBHC is that I care deeply about it.

Outside of research, my colleagues at Åland: Malin, Matti, Victor, Jonas, Lidia, Åsa, Marina, Marina, Carro, Linnea, Marlene, Sussi, and Josefin, who gave me courage to endure clinical work. Snart blir det flådiga drinkar och trendigt bryggkaffe!

Also outside of research, the people who shaped my thinking about management – the minds of the Swedish Federation of Young Scientists—so many brilliant people. A special thanks to: Sanna Holm, Andreas Svensson, Guncha Ekheden, Anders Lundberg, Alexander Sehlström, Daniel Langkilde, and Diana Diez. What you taught me might have never given me a job, but on every single job, the knowledge has saved me so many times. Another important management companion – Leo Silvén Möller – *the* partner in the Nobel Night Cap darkness, wise and humble. Thank you!

Mamma, Pappa, Elisa, Kattis – I love you.

My friends during this time: Angelica Joelsson, Johanna Hansson, Caroline Bergqvist, Fredrika Svahn, Isak Eliasson, Linnea Rusek, and Filip Gloria. I have too much to say to fit on this page. Also some additional friends I've lost contact with but who meant a lot to me in

the beginning of this project: Christina Ehrenborg, Ida Havland, Karolin Planting, Gustaf Prag, Hedvig Löfdahl, Anna Lennartsson, Elisabet Ålander, Sana Amin, and Anne Örtqvist.

Maníh – most of this book I wrote before I knew you, but I could not write this section without mentioning you. When one looks back on love, it is often the happiest moments that are remembered. But the truly great moments, the moments when love shines with the most respectful intensity, is love in the face of struggle. My inability to handle failures has tested me, but in no way close to the way in which it has tested your resilience. Under pressure great character can arise, and yours has arisen in the most beautiful way. Beijo!

8 REFERENCES

1. Bohmer R. Designing care. Boston: Harvard Business School Publishing; 2009.
2. Economist Intelligence Unit. Global outlook: Healthcare. 2014, March. Retrieved from <http://www.eiu.com/industry/Healthcare>.
3. Institute of Medicine, Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century. Washington DC: National Academy Press; 2001.
4. Porter ME, Teisberg EO. Redefining health care: creating value-based competition on results. Brighton, MA: Harvard Business Press; 2006.
5. Christensen CM, Grossman JH, Hwang J. The innovator's prescription: a disruptive solution for health care. New York: McGraw-Hill; 2009.
6. Porter ME. Value-based health care delivery. *Ann Surg.* 2008;248(4):503-9.
7. Porter ME. Defining and introducing value in healthcare. In: McClellan MB, McGinnis JM, Nabel EG, Olsen LM, editors. Evidence-based medicine and the changing nature of health care (IOM roundtable on evidence-based medicine). Washington DC: National Academies Press;2008. p. 161-71.
8. Porter ME. A strategy for health care reform: Toward a value-based system. *NEJM.* 2009;361(2):109-12.
9. Porter ME. What is value in health care? *NEJM.* 2010;363(26):2477-81.
10. Porter ME, Lee TH. The strategy that will fix health care. *Harv Bus Rev.* 2013;91(12):24.
11. Porter ME, Pabo EA, Lee TH. Redesigning primary care: a strategic vision to improve value by organizing around patients' needs. *Health Aff.* 2013;32(3):516-25.
12. Porter ME, Teisberg EO. How physicians can change the future of health care. *JAMA.* 2007;297(10):1103-11.
13. Mazzocato P, Savage C, Brommels M, Aronsson H, Thor J. Lean thinking in healthcare: a realist review of the literature. *Qual Saf Health Care.* 2010;19(5):376-82.
14. Øvretveit J. Total quality management in European healthcare. *Int J Health Care Qual Assur.* 2000;13(2):74-80.
15. Walshe K. Pseudoinnovation: The development and spread of healthcare quality improvement methodologies. *Int J Qual Health Care.* 2009:mzp012. DOI: 10.1093/intqhc/mzp012.
16. Birkinshaw J, Hamel G, Mol MJ. Management innovation. *Acad Manage Rev.* 2008;33(4):825-45.
17. Abrahamson E, Fairchild G. Management fashion: lifecycles, triggers, and collective learning processes. *Admin Sci Q.* 1999;44(4):708-40.
18. Jackson B. Management gurus and management fashions: a dramatisic inquiry. London: Psychology Press; 2001.
19. Heusinkveld S, Sturdy A, Werr A. The co-consumption of management ideas and practices. *Manage Learn.* 2011;42(2):139-47.
20. Abrahamson E. Management fashion. *Acad Manage Review.* 1996;21(1):254-85.

21. Antoñanzas F, Terkola R, Postma M. The value of medicines: a crucial but vague concept. *Pharmacoeconomics*. 2016;34(12):1227-39.
22. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff*. 2008;27(3):759-69.
23. Muir G, Anant J. Promoting triple value healthcare in countries with universal healthcare. *Healthc Pap*. 2016;15(3):42-8
24. Nelson EC, Mohr JJ, Batalden PB, Plume SK. Improving health care, part 1: the clinical value compass. *Jt Comm J Qual Improv*. 1996;22(4):243-58.
25. Ramaswamy V, Ozcan K. The co-creation paradigm. Palo Alto, CA: Stanford University Press; 2014..
26. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2015.
27. Codman EA. *The shoulder; rupture of the supraspinatus tendon and other lesions in or about the subacromial bursa*. Boston: T. Todd; 1934.
28. Donabedian A. Evaluating the quality of medical care. *Milbank Q*. 1966;44(3):166-206.
29. Huber M. Measuring medicine's effects: The outcomes management system. *J Occup Environ Med*. 1991;33(3):272.
30. Steinwachs DM, Wu AW, Skinner EA. How will outcomes management work? *Health Aff*. 1994;13(4):153-62.
31. Kozma C, Reeder C, Schulz R. Economic, clinical, and humanistic outcomes: a planning model for pharmacoeconomic research. *Clin Ther*. 1992;15(6):1121-32.
32. Fayanju OM, Mayo TL, Spinks TE, Lee S, Barcenas CH, Smith BD, et al. Value-based breast cancer care: a multidisciplinary approach for defining patient-centered outcomes. *Ann Surg Oncol*. 2016;23(8):1-6.
33. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum*. 1980;23(2):137-45.
34. Fries JF. The assessment of disability: from first to future principles. *Rheumatol*. 1983;22(suppl 1):48-58.
35. Donabedian A. The quality of care: how can it be assessed? *JAMA*. 1988;260(12):1743-8.
36. Kaplan RS, Porter ME. How to solve the cost crisis in health care. *Harv Bus Rev*. 2011;89(9):46-52..
37. Antonaci F, Sances G, Guaschino E, De Cillis I, Bono G, Nappi G. Meeting patient expectations in migraine treatment: what are the key endpoints? *J Headache Pain*. 2008;9(4):207-13.
38. Porter ME. Measuring health outcomes: the outcomes hierarchy. *NEJM*. 2010;363:2477-81.
39. Stuart B, Chernof BA. What is value in health care? *NEJM*. 2011;364(13):e26; author reply e.
40. Teperi J, Porter ME, Vuorenkoski L, Baron JF. The Finnish health care system: a value-based perspective. *Sitra reports*. 2009;82.

41. Porter ME. What is value in health care? *NEJM*. 2010;363(26):2477-81.
42. Trochim WM, Donnelly JP. *Research methods knowledge base*. 2nd edition. Boston: Cengage Learning; 2001
43. Denzin NK. *Sociological methods: a sourcebook*. New Brunswick, NJ: Aldine Transaction; 2006.
44. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. *Am J Prev Med*. 2016;50(2):129-35.
45. Porter ME, Kaplan RS. How to pay for health care. *Harvard Bus Rev*. 2016;94(7-8):88.
46. Bohmer RM. The four habits of high-value health care organizations. *NEJM*. 2011;365(22):2045-7.
47. Larsson S, Lawyer P, Garellick G, Lindahl B, Lundström M. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. *Health Aff*. 2011;10.1377/hlthaff.2011.0762.
48. Lepage A, Gzil F, Cammelli M, Lefevre C, Pachoud B, Ville I. Person-centredness: conceptual and historical perspectives. *Disabil Rehabil*. 2007;29(20-21):1555-65.
49. Stuart B, Chernof B. What is value in health care? *NEJM*. 2011;364(13):e26.
50. Espallargues M, Serra-Sutton V, Solans-Domènech M, Torrente E, Moharra M, Benítez D, et al. Development of a conceptual framework for the assessment of chronic care in the Spanish national health system. *Revista Española de Salud Pública*. 2016;90.
51. Norris SL, Glasgow RE, Engelgau MM, O'Connor PJ, McCulloch D. Chronic disease management. *Dis Manag Health Out*. 2003;11(8):477-88.
52. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. *JAMA*. 2002;288(19):2469-75.
53. Harris A, Carney K, Volpicelli FM. Value in hospital medicine. *Hosp Med Clin*. 2017;6(1):131-45.
54. Tilburt JC, Montori VM, Shah ND. What is value in health care? *NEJM*. 2011;364(13):e26; author reply e.
55. Tseng EK, Hicks LK. Value based care and patient-centered care: divergent or complementary? *Curr Hematol Malig R*. 2016:1-8.
56. Alvesson M, Sandberg J. Generating research questions through problematization. *Acad Manage Rev*. 2011;36(2):247-71.
57. Crawford MJ, Rutter D, Manley C, Weaver T, Bhui K, Fulop N, et al. Systematic review of involving patients in the planning and development of health care. *BMJ*. 2002;325(7375):1263.
58. Chewning B, Bylund CL, Shah B, Arora NK, Gueguen JA, Makoul G. Patient preferences for shared decisions: a systematic review. *Patient Educ Couns*. 2012;86(1):9-18.
59. Lind C. Patient-professional partnership: Examples, best practices and recommendations [Report for QRC, June 2014, accessed March 2016]. Available from: <http://qrcstockholm.se/wp-content/uploads/2015/02/Patient-professional-partnership-examples-best-practices-and-recommendations.pdf>.

60. Holmström I, Röing M. The relation between patient-centeredness and patient empowerment: a discussion on concepts. *Patient Ed Couns*. 2010;79(2):167-72.
61. Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? *Patient Educ Couns*. 2007;66(1):13-20.
62. Cornwell J. Reframing the work on patient experience improvement. *PXJ*. 2015;2(1):11-4.
63. Verma R, Elg M, Engström J, Witell L, Poksinska B. Co-creation and learning in health care service development. *JOSM*. 2012;23(3):328-43.
64. McAlister AL, Perry CL, Parcel GS. How individuals, environments, and health behaviors interact. In: Glanz K, Rimer BK, Viswanath K, editors. *Health behavior and health education: theory, research, and practice*. 4th edition. San Francisco: Jossey-Bass;2008. p. 169-188.
65. Fumagalli LP, Radaelli G, Lettieri E, Masella C. Patient empowerment and its neighbours: clarifying the boundaries and their mutual relationships. *Health Policy*. 2015;119(3):384-94.
66. Development PFM. Collaborative patient engagement: mapping the global landscape; a first step in co-creating an action-orientated framework for patient engagement. 2016.
67. Dyrstad DN, Testad I, Aase K, Storm M. A review of the literature on patient participation in transitions of the elderly. *Cogn Technol Work*. 2015;17(1):15-34.
68. Jones EL, Williams-Yesson BA, Hackett RC, Staniszewska SH, Evans D, Francis NK. Quality of reporting on patient and public involvement within surgical research: a systematic review. *Ann Surg*. 2015;261(2):243-50.
69. Gallant MH, Beaulieu MC, Carnevale FA. Partnership: an analysis of the concept within the nurse-client relationship. *J Adv Nurs*. 2002;40(2):149-5.
70. Whear R, Abdul-Rahman A-K, Thompson-Coon J, Boddy K, Perry MG, Stein K. Patient initiated clinics for patients with chronic or recurrent conditions managed in secondary care: a systematic review of patient reported outcomes and patient and clinician satisfaction. *BMC Health Serv Res*. 2013;13(1):501.
71. Lempp H, Scott DL, Kingsley GH. Patients' views on the quality of health care for rheumatoid arthritis. *Rheumatol*. 2006;45(12):1522-8.
72. Desai SP, Solomon DH. A new paradigm of quality of care in rheumatoid arthritis: how our new therapeutics have changed the game. *Arthritis Res Ther*. 2013;15(5).
73. Kalyoncu U, Dougados M, Daures J, Gossec L. Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review. *Ann Rheum Dis*. 2009;68(2):183-90.
74. Popper K. Philosophy of science. In: Mace CA, editor. *British philosophy in the mid-century*. London: George Allen and Unwin; 1957.
75. Kuhn TS. Logic of discovery or psychology of research. In: Lakatos I, Musgrave A, editors. *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press; 1970. p. 1-23.
76. Lakatos I. *Science and pseudoscience: Philosophical papers*. Cambridge: Cambridge University Press;1978; p. 1-7.

77. Thagard PR, editor. Why astrology is a pseudoscience. PSA: proceedings of the biennial meeting of the Philosophy of Science Association; 1978. JSTOR.
78. Bryman A. Social research methods. Oxford: Oxford University Press; 2015.
79. Gough D, Thomas J, Oliver S. Clarifying differences between review designs and methods. *Syst Rev*. 2012;1(1):1.
80. Hjørland B. Concept theory. *J Am Soc Inf Sci Tech*. 2009;60(8):1519-36.
81. McDermid D. Pragmatism [Internet]. Internet encyclopedia of philosophy, n.d. Available from <http://www.iep.utm.edu/pragmati/>.
82. Patton MQ. Qualitative evaluation and research methods. 4th edition. Thousand Oaks, CA: Sage;1990.
83. Hall R. Mixed methods: in search of a paradigm. In Lê T, Lê Q, editors. *Conducting research in a changing and challenging world*. New York: Nova Science Publishing; 2013. p. 71-78.
84. Åsberg R. Det finns inga kvalitativa metoder–och inga kvantitativa heller för den delen: det kvalitativa-kvantitativa argumentets missvisande retorik. *Pedagogisk forskning i Sverige*. 2001;6(4):270.
85. Lilford JPR. Qualitative methods: an alternative view. *BMJ*. 2011;342:957.
86. Teddlie C, Tashakkori A. *Foundations of mixed methods research: integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage; 2009.
87. Glaser B, Strauss A. *The discovery of grounded theory*. London: Weidenfeld & Nicholson; 1967. p. 288-304.
88. Reichardt CS, Cook TD. *Beyond qualitative versus quantitative methods: qualitative and quantitative methods in evaluation research*. Beverly Hills, CA: 1979; p. 7-32.
89. Wax ML, Gouldner AW. The coming crisis of western sociology. *Kansas Journal of Sociology*. 1970;3(4):217-219. JSTOR.
90. Tashakkori A, Teddlie C, editors. *SAGE handbook of mixed methods in social & behavioral research*. JSTOR; 2011.
91. Klassen AC, Creswell J, Plano Clark VL, Smith KC, Meissner HI. Best practices in mixed methods for quality of life research. *Qual Life Res*. 2012;21(3):377-80.
92. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs*. 2008;62(1):107-15.
93. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.
94. Thomson Reuters. *ISI web of knowledge: journal citations reports*; 2014. Available from: <http://apps.isiknowledge.com>. 2014.
95. Biggs J. Individual differences in study processes and the quality of learning outcomes. *High Educ*. 1979;8(4):381-94.
96. Yin RK, Heald KA. Using the case survey method to analyze policy studies. *Admin Sci Q*. 1975:371-81.

97. Pelz D. Use of innovation in innovating processes by local governments. Ann Arbor: University of Michigan, CRUSK Institute for Social Research; 1981.
98. Krathwohl DR. A revision of Bloom's taxonomy: an overview. *Theory Pract.* 2002;41(4):212-8.
99. Hattie JAC, Brown GTL. Cognitive processes in asTTle: the SOLO taxonomy. asTTle Technical Report #43. Auckland: University of Auckland/Ministry of Education; 2004.
100. Majithia V, Geraci SA. Rheumatoid arthritis: diagnosis and management. *Am J Med.* 2007;120(11):936-9.
101. Schipper LG, Kievit W, den Broeder AA, van der Laar MA, Adang EM, Fransen J, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: starting with methotrexate monotherapy is cost-effective. *Rheumatol (Oxford)*. 2011. DOI: 10.1093/rheumatology/ker084.
102. Barlow J. Living with arthritis. Oxford: John Wiley & Sons; 2009.
103. Schipper LG, van Hulst LT, Grol R, van Riel PL, Hulscher ME, Fransen J. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. *Rheumatol (Oxford)*. 2010;49(11):2154-64.
104. Lindblad S, Essén A, Stawiarz L. Svensk Reumatologis kvalitetsregister-Årsrapport. Svensk Reumatologis förening; 2009-2010.
105. Hvitfeldt H, Carli C, Nelson EC, Mortenson DM, Ruppert BA, Lindblad S. Feed forward systems for patient participation and provider support: adoption results from the original US context to Sweden and beyond. *Qual Manag Health Care.* 2009;18(4):247-56.
106. Ndosi M, Vinall K, Hale C, Bird H, Hill J. The effectiveness of nurse-led care in people with rheumatoid arthritis: a systematic review. *Int J Nurs Stud.* 2011;48(5):642-54.
107. Reid A, Brady A, Blake C, Mongey AB, Veale DJ, FitzGerald O, et al. Randomised controlled trial examining the effect of exercise in people with rheumatoid arthritis taking anti-TNF α therapy medication. *BMC Musculoskelet Disord.* 2011;12:11
108. Lineker SC, Bell MJ, Badley EM. Evaluation of an inter-professional educational intervention to improve the use of arthritis best practices in primary care. *J Rheumatol.* 2011;38(5):931-7.
109. Gabriel SE. The epidemiology of rheumatoid arthritis. *Rheum Dis Clin North Am.* 2001;27(2):269-81.
110. van Staa TP. The pathogenesis, epidemiology and management of glucocorticoid-induced osteoporosis. *Calcif Tissue Int.* 2006;79(3):129-37.
111. Sarzi-Puttini P, Atzeni F, Gerli R, Bartoloni E, Doria A, Barskova T, et al. Cardiac involvement in systemic rheumatic diseases: an update. *Autoimmun Rev.* 2010;9(12):849-52.
112. Leombruno JP, Einarson TR, Keystone EC. The safety of anti-tumour necrosis factor treatments in rheumatoid arthritis: meta and exposure-adjusted pooled analyses of serious adverse events. *Ann Rheum Dis.* 2009;68(7):1136-45.
113. Chakravarty EF, Michaud K, Wolfe F. Skin cancer, rheumatoid arthritis, and tumor necrosis factor inhibitors. *J Rheumatol.* 2005;32(11):2130-5.

114. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum.* 1995;38(1):44-8.
115. Mäkinen H, Kautiainen H, Hannonen P, Möttönen T, Korpela M, Leirisalo-Repo M, et al. Disease activity score 28 as an instrument to measure disease activity in patients with early rheumatoid arthritis. *J Rheumatol.* 2007;34(10):1987-91.
116. Fransen J, van Riel PL. DAS remission cut points. *Clin Exp Rheumatol.* 2006;24(6 Suppl 43):S-29-32.
117. Van Gestel A, Prevoo M, Van't Hof M, Van Rijswijk M, Van de Putte L, Van Riel P. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis: Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism criteria. *Arthritis Rheum.* 1996;39(1):34-40.
118. Soderlund N, Kent J, Lawyer P, Larsson S. Progress toward value-based health care: lessons from 12 countries. Boston: Boston Consulting Group; 2012.
119. Kvien TK, Heiberg T. Patient perspective in outcome assessments: perceptions or something more? *J Rheumatol.* 2003;30(4):873-6.
120. Anderson DR, Burnham KP. Avoiding pitfalls when using information-theoretic methods. *J Wildl Manag.* 2002:912-8.
121. Krippendorff K. Content analysis: an introduction to its methodology. Thousand Oaks, CA: Sage; 2012.
122. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures, and measures to achieve trustworthiness. *Nurse education today.* 2004;24(2):105-12.
123. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nurs Health Sci.* 2013;15(3):398-405.
124. Timmermans S, Tavory I. Theory construction in qualitative research from grounded theory to abductive analysis. *Soc Theory.* 2012;30(3):167-86.
125. World Health Organization. International classification of functioning, disability, and Health (ICF). World Health Organization: 22 July 2016. Available from: <http://www.who.int/classifications/icf/en>.
126. Sverker A, Thyberg I, Östlund G, Waltersson E, Thyberg M. Participation in work in early rheumatoid arthritis: a qualitative interview study interpreted in terms of the ICF. *Disabil Rehabil.* 2014;36:242-9.
127. Coenen M, Cieza A, Stamm TA, Amann E, Kollerits B, Stucki G. Validation of the international classification of functioning, disability, and health (ICF) core set for rheumatoid arthritis from the patient perspective using focus groups. *Arthritis Res Ther.* 2006;8:R84.
128. Stamm TA, Cieza A, Coenen M, Machold KP, Nell VP, Smolen JS, et al. Validation of the international classification of functioning, disability, and health (ICF) core set for rheumatoid arthritis from the patient perspective using focus groups. *Arthritis Rheum.* 2005;53(3):431-9.

129. Ebbevi D. Value based health care: operationalizing patient outcome using registry data; the case of the Swedish Rheumatology Quality Registry. Stockholm: Karolinska Institutet; 2011.
130. Catarinella F, Bos W. Digital health assessment in rheumatology: current and future possibilities. *Clin Exp Rheumatol*. 2016;34(5):2.
131. Smith DE. *The everyday world as problematic: a feminist sociology*. Toronto: University of Toronto Press; 1987.
132. Rosch E. Cognitive representations of semantic categories. *J Exp Psychol*: 1975;104(3): 192.
133. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*. 1981;29(2):75-91.
134. Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin; 1963.
135. Creswell JW, Miller DL. Determining validity in qualitative inquiry. *Theory Pract*. 2000;39(3):124-30.
136. Andersen RS, Risør MB. The importance of contextualization: anthropological reflections on descriptive analysis, its limitations and implications. *Anthropol Med*. 2014;21(3):345-56.
137. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16(1):62-67.
138. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. *Educ Inform*. 2004;22(2):63-75.
139. Denzin NK, Lincoln YS, editors. *Handbook of qualitative research*. Thousand Oaks, CA: Sage;1994.
140. Beauchamp TL, Childress JF. *Principles of biomedical ethics*. New York: Oxford University Press; 2001.
141. Pentz RD, White M, Harvey RD, Farmer ZL, Liu Y, Lewis C, et al. Therapeutic misconception, misestimation, and optimism in participants enrolled in phase 1 trials. *Cancer*. 2012;118(18):4571-8.
142. Wong Y-N, Roach N, Meropol NJ. Addressing patients' priorities as a strategy to improve value. *Oncologist*. 2016;21(11):1279-82.
143. Rollow W, Cucchiara P. Achieving value in primary care: the primary care value model. *Ann Fam Med*. 2016;14(2):159-65.
144. Teisberg EO, Porter ME, Brown GB. Making competition in health care work. *Harvard Bus Rev*. 1993;72(4):131-41.
145. Rosenberg Hansen J, Ferlie E. Applying strategic management theories in public sector organizations: developing a typology. *Public Manag Rev*. 2016;18(1):1-19.
146. Bryson JM, Roering WD. Applying private-sector strategic planning in the public sector. *J Am Plann Assoc*. 1987;53(1):9-22.
147. Folland S, Goodman AC, Stano M. *The economics of health and health care*. Upper Saddle River, NJ: Pearson Prentice Hall; 2007.

148. Andersson AE, Bååthe F, Wikström E, Nilsson K. Understanding value-based healthcare: an interview study with project team members at a Swedish university hospital. *J Hosp Adm.* 2015;4(4): 64.
149. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the patient activation measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res.* 2004;39(4p1):1005-26.
150. Woodward CA, Abelson J, Tedford S, Hutchison B. What is important to continuity in home care? Perspectives of key stakeholders. *Soc Sci Med.* 2004;58(1):177-92.
151. Parasuraman A, Zeithaml VA, Berry LL. A conceptual model of service quality and its implications for future research. *J Marketing.* 1985:41-50.
152. Vaihinger H. The philosophy of “as if”: a system of the theoretical, practical, and religious fictions of mankind. Ogden CK, trans. London: Kegan Paul, Trench, Trubner; 1924.
153. Jackson CL, Janamian T, Booth M, Watson D. Creating health care value together: a means to an important end. *Med J Aust.* 2016;204(7):S3-S.
154. Speroff T, Miles PV, Dougherty D, Mittman BS, Splaine ME. Introduction to the Academy for Healthcare Improvement conference on advancing the methods for healthcare quality improvement research. *Implement Sci.* 2013;8(1):1.
155. Lorts A, Bucuvalas J. Identifying and improving value in patients with chronic disease. *Curr Treat Options Pediatr.* 2016;2(4):311-1-7.
156. Körber C, Strååt D, Henter J-I, Ringman-Uggla A, Dabhilkar M. Implementing value-based health care at the provider level: an operations management view. EurOMA conference paper; 2016.
157. Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. *NEJM.* 2016;374(6):504-6.
158. Colegate-Stone T, Tavokkolizadeh A, Moxham J, Sinha J. Increasing value: the King’s College Hospital Experience. *BJHCM.* 2016;22(6):326-34.
159. Rothes M, Laranja-Pontes J, Rocha-Gonçalves F. Value measurement approach in healthcare: meta-analysis of European health care systems against cancer, using published data. *Health Science Journal.* 2016;10(3):1-7.
160. Parra E, Arenas MD, Alonso M, Martínez MF, Gamen Á, Aguarón J, et al. Assessing value-based health care delivery for haemodialysis. *J Eval Clin Pract.* 2016.
161. Overveld L, Braspenning J, Hermens R. Quality indicators of integrated care for patients with head and neck cancer. *Clin Otolaryngol.* 2016.
162. Mofidi R, Wong P, Gatenby T, Milburn S. Value based healthcare and delivery of vascular surgery services in the United Kingdom. *Rev Vasc Med.* 2016;6:1-9.
163. Shah SB, Gupta T, Severinsen KD, McIlwain E, White CJ. Volume to value: defining the value of cardiovascular imaging. *Ochsner J.* 2016;16(3):203-7.
164. Porter ME, Guth C. Redefining German health care: moving to a value-based system. Berlin: Springer Science & Business Media; 2012.
165. Ying AK, Feeley TW, Porter ME. Value-based healthcare: implications for thyroid cancer. *Int J Endocr Oncol.* 2016;3(2):115-29.

166. Salinas J, Sprinkhuizen SM, Ackerson T, Bernhardt J, Davie C, George MG, et al. An international standard set of patient-centered outcome measures after stroke. *Stroke*. 2016;47(1):180-6.
167. Rodrigues IA, Sprinkhuizen SM, Barthelmes D, Blumenkranz M, Cheung G, Haller J, et al. Defining a minimum set of standardized patient-centered outcome measures for macular degeneration. *Am J Ophthalmol*. 2016;168:1-12.
168. Allori AC, Kelley T, Meara JG, Albert A, Bonanthaya K, Chapman K, et al. A standard set of outcome measures for the comprehensive appraisal of cleft care. *Cleft Palate Craniofac J*. 2016.
169. Rolfson O, Wissig S, van Maasackers L, Stowell C, Ackerman I, Ayers D, et al. Defining an international standard set of outcome measures for patients with hip or knee osteoarthritis: consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. *Arthritis Care Res*. 2016.
170. Thaker NG, Pugh TJ, Mahmood U, Choi S, Spinks TE, Martin NE, et al. Defining the value framework for prostate brachytherapy using patient-centered outcome metrics and time-driven activity-based costing. *Brachytherapy*. 2016;15(3):274-82.
171. Thaker NG, Ali TN, Porter ME, Feeley TW, Kaplan RS, Frank SJ. Communicating value in health care using radar charts: a case study of prostate cancer. *J Oncol Pract*. 2016;12(9):813-20.
172. Kinsman L, Rotter T, James E, Snow P, Willis J. What is a clinical pathway? Development of a definition to inform the debate. *BMC Med*. 2010;8(1):1.
173. Balducci L. Value, the new medical mantra: misunderstood, abused, and underused. *J Med Person*. 2013;11(1):3-4.
174. Levay C. Policies to foster quality improvement registries: lessons from the Swedish case. *J Intern Med*. 2016;279(2):160-72.
175. Hewlett S, Carr M, Ryan S, Kirwan J, Richards P, Carr A, et al. Outcomes generated by patients with rheumatoid arthritis: How important are they? *Musculoskeletal care*. 2005;3(3):131-42.
176. Hahn S. Understanding noninferiority trials. *Korean J Pediatr*. 2012;55(11):403-7.
177. Straus A, Corbin J. Grounded theory methodology: an overview. In: Denzin NK, Lincoln YS, editors. *Handbook of qualitative research*. Thousand Oaks, CA: Sage;1994. p. 273-85.