INVEST

INVEstigating Safety Tipping points in Swiss hospitals: assessing the causal effect of capacity utilization on in-hospital mortality using routine data

Inaugural dissertation

To be awarded the degree of Dr. sc. med.

Presented at
the Faculty of Medicine
of the University of Basel

by

Narayan Sharma
from Parbat, Nepal

Basel, 2021

Approved by the Faculty of Medicine

On application of

Prof. Dr. Michael Simon (primary supervisor)

Prof. Dr. René Schwendimann (second supervisor)

Dr. Dietmar Ausserhofer (further advisor)

Dr. med. Olga Endrich (further advisor)

Prof. Dr. Jack Needleman (external expert)

Basel, 30.08.2021

Prof. Dr. Primo Leo Schär

Dean

Table of Contents

To my parents who lighted my path of learning, even though they had no opportunity for their schooling.

List of Abbreviations

AND-PSO	American Data Network-Patient Safety Organization
AHRQ	Agency for Healthcare Research and Quality
CHOP	Schweizerische Operationsklassifikation (Swiss operation procedure classification)
CI	Confidence Interval
COVID-19	COrona VIrus Disease of 2019
DAGs	Directed Acyclic Graphs
DRGs	Diagnosis-Related Groups
ICD-10	International Classification of Diseases, 10th revision
INVEST	INVEstigating Safety Tipping points in Swiss hospitals
IOM	Institute of Medicine
IPTW	Inverse Probability of Treatment Weight
LOS	Length of Stay
MSM	Marginal Structural Model
PCCL	Patient Clinical Complexity Level
PSIs	Patient Safety Indicators
RCT	Randomized Controlled Trial
SD	Standard Deviation
SwissDRGs	Swiss Diagnosis-Related Groups
TCF	Treatment Confounder Feedback
WHO	World Health Organization

Acknowledgements

Learning is a continue process and everything is hard before it is easy. First and foremost, I would like to acknowledge and thank all the people who contributed directly or indirectly to my learning that resulted to this dissertation, even if I forget to mention anyone by name.

Particularly, I would like to thank my PhD committee, Prof. Dr. Michael Simon, Prof. Dr. René Schwendimann, Dr. Dietmar Ausserhofer and Dr. med. Olga Endrich, for giving me the opportunity to work on this exciting project (INVEST) to pursue a PhD at the Institute of Nursing Science (INS). My deepest gratitude goes to my primary supervisor, Prof. Dr. Simon who gave the beginning to work in MatchRN project, which eventually followed with the opportunity to undertake this PhD. His support thorough out the entire process of learning has been vital not only to accomplishing this dissertation but also developing my professional and personal skills. Prof. Dr Schwendimann was always supportive and provided his expertise on patient safety, academia, and successful development of the project. Dr. Ausserhofer has guided me through the review process, listening and discussing different issues during the journey to reaching this dissertation. Dr. med. Endrich brought her instrumental knowledge of medical controlling data in the project. Additional thank goes to Prof. Giusi Moffa providing technical knowledge on causal questions in the project. A special thanks goes to Prof. Dr. Jack Needleman who agreed to be the external expert for assessing my work. It was great learning by reading his papers and now, I feel privileged to have him being the part of my dissertation.

Furthermore, I would like to thank all the Swiss hospitals regularly reporting the data to Swiss Federal Statistics Office (FSO). A special thanks goes to the data management and administrative team of FSO providing the data for our analysis.

I also wish to thank all my colleagues at the INS and in particular Prof. Dr. Sabina de Geest. She was supportive and motivating not only on my academic carrier but also moral support to my family and my home country, Nepal. I am grateful to the A-team at INS (Greet Van Malderen, Cornelia Kern, Michael Huber, Juliane Barnick and Michelle Killenberger) supporting for all the administrative and IT knowledge. I wish to thank my fellow PhD colleagues, postdocs and seniors for their moral support, motivation and the important contributions to the success of this thesis.

My sincere thanks go to my family in Switzerland, and back in Nepal and friends all around. My deepest gratitude goes to my parents, providing the first insights of learning and who are proud of my achievements. A special thanks go to my wife Barbara Sharma-Grüter for her support and patience during the journey of my PhD. Last but not the least, thanks to our kids Asmita and Nabin for their smile.

Narayan Sharma, July 2021

Summary

Measuring patient safety is challenged by the differences in methods for measurement, variation in services provided, and care demands in hospitals. To minimize them health service providers, policymakers and researchers investigate potential causes of safety issues. Capacity utilization, (i.e., bed-occupancy rate), is one aspect of safety tipping points i.e., high capacity utilization levels in which patient safety incidents like in-hospital mortalities are more likely to occur. To see whether the entire hospital system can supply sufficient resources to match the care demand, several aspects of care like the volume of patients, turnover, and severity of patient's illnesses need to be considered and their link with outcomes such as mortality should be measured.

Previous studies have shown the association, particularly capacity utilization and in-hospital mortality. Yet, the time-varying nature of capacity utilization in hospitals with its time-varying confounders was not investigated at all. Even though, studies included other individual factors and comorbidities in the model of capacity utilization and mortality. The risk adjustment of both individual and time-varying covariates with a causal inference methodology in health service research could enhance our understanding based on observational data. Applying the causal perspectives in the routine data helps to understand patient safety in the absence of randomized experiments.

Therefore, the overall aim of this dissertation was to investigate different aspects of the safety tipping phenomenon. We first explored hospital care demand from a longitudinal perspective, mainly daily capacity utilization, patient turnover, and patient clinical complexity level (PCCL) of 102 Swiss general hospitals. Secondly, we derived Swiss comorbidity weights and compared these with existing weights to predict in-hospital mortality. Finally, a time-varying causal model with individual factors was explored via directed acyclic graph (DAG) and the potential causal effect of daily capacity utilization on in-hospital mortality was estimated.

This dissertation is embedded in the INVEST (INVEstigating Safety Tipping points in Swiss hospitals) study, focusing on in-hospital mortality as a patient safety indicator, which is all-cause mortality from the routinely collected data of all Swiss hospitals from 2012 to 2017. Overall, this dissertation is structured in six chapters.

Chapter 1 provides a general introduction in the field of patient safety, in-hospital mortality, system thinking, and the tipping point phenomenon in hospital care. Furthermore, a brief outline of comorbidity and weighting systems to predict in-hospital mortality is provided. Additionally, an introduction of causal inference with time-varying variables and the causal effect estimations are offered for capacity utilization and in-hospital mortality. **Chapter 2** states the aims of this dissertation.

The article presented in **Chapter 3** constitutes the findings of care demand variation in Swiss hospitals from the longitudinal perspective. We utilized one-year patients' data to explore the variation of each hospital care demand. Hospital care demand was longitudinally explored as a percentage of daily capacity utilization, percentage of daily patient turnover, and average daily patient clinical complexity level values per general hospital. Patient clinical complexity level (PCCL) was measured as a cumulative effect of a patient's clinical complexities or comorbidities (CC) for each episode of care and it ranged from 0 to 4, i.e., no CC to very severe CC. We used Swiss diagnosis-related groups version 6 for compiling clinical conditions reported in ICD 10 codes for generating each patient's PCCL value. The results indicate the average daily capacity utilization (87.7%) and average daily PCCL value (2.06) were highest in university hospitals but the patient turnover (22.5%) was lowest. Similarly, patient turnover (34.5%) was highest in medium basic hospitals but the lowest daily PCCL value (1.26). Moreover, the capacity utilization (57.8%) was lowest in small basic hospitals. The findings suggest there was pronounced variability of all three measures of care demand in Swiss hospitals and they distinctly varied between days, weeks, and seasons throughout the year. Variation and daily trends in care demand required a balanced care supply in Swiss hospitals.

Chapter 4 describes different comorbidities indices in Swiss hospitals, one of the measures of patient complexity and the predictive performance of different comorbidity weighting systems with in-hospital mortality. We used six years of Swiss general hospital data with a total of 6.09 million patients' cases for the analysis. The data were randomly split into two halves to derive and validate Swiss comorbidity weights from the Elixhauser comorbidity index. Both sets were used for validation of the new weights and to compare new weights with existing comorbidity weights to predict in-hospital mortality. Derivation and validation of weightings were conducted

with generalized additive models adjusted for age, gender, and hospital types, and the comparison was done using c-statistic and net reclassification improvement (NRI). Overall, c-statistic with Swiss weights (0.867, 95% CI, 0.865–0.868) was slightly higher than van Walraven's weights (0.863, 95% CI, 0.862–0.864) and Charlson's weights (0.850, 95% CI, 0.849–0.851). The NRI of new Swiss weights improved the predictive performance by 1.6% on the Elixhauser-van Walraven and 4.9% on the Charlson weights. The patient population-based Swiss weights support the analysis of in-hospital mortality and other health outcomes. Comorbidities/comorbidity scores are important individual factors for risk adjustment for hospital exposure and health outcomes.

Chapter 5 analyzed the causal effect of capacity utilization on in-hospital mortality in Swiss hospitals. Grounded on the variation of the time-varying measures of care demand (Chapter 3), we hypothesized exposure to daily capacity utilization has a causal effect on 14-days in-hospital mortality. Time-varying confounders and time-fix variables of capacity utilization and in-hospital mortality were linked along with an unmeasured variable using DAGs. Daily capacity utilization in Swiss hospitals was explored to observe safety tipping points at the 85th percentile of the distribution on the hospital level. To test our hypothesis, we used one-year patient data, adding 14-days exposure to capacity utilization, the time-varying confounders, patient turnover, and PCCL and time-fixed variables weekdays, hospital types, Elixhauser comorbidity index weighting score, age, and sex. Inverse probability of treatment weights (IPTW) was computed to balance the weight between high capacity utilization and low capacity utilization from the safety tipping points cutoff of 85th percentile and to eliminate treatment-confounder feedback. The computed IPTW was incorporated with a marginal structural model (MSM) to estimate the causal effect of capacity utilization on in-hospital mortality. The finding shows that daily exposure to high capacity utilization yielded a 2% increase in 14-days in-hospital mortality for each additional day of high capacity utilization (Odds Ratio (OR) 1.02, 95% CI: 1.01 to 1.03). Additionally, we found weekend effects and increased mortality risk with increasing comorbidity scores in the Swiss patient population. The variation of safety tipping points between hospitals suggests Swiss hospitals get strained at different points

ranging between 42.1% to 95.9%. Thus, hospitals require the planning of resources based on the intensity of the care demand.

Chapter 6 of this dissertation summarizes the key findings of all studies and discusses them in the context of the literature. Furthermore, the strengths and limitations of the studies are discussed, and implications of causal inference in an observational study and the application of evidence in hospitals are presented. This dissertation contributes to the current literature in the field of patient safety and the causal relationship between capacity utilization and in-hospital mortality in an observational study. It identifies safety tipping points variation in terms of capacity utilization in Swiss general hospitals. After reaching tipping points, hospitals are under strain, and matching hospital resources with care demands is required to improve patient safety and quality of care.

Chapter 1

Introduction

Patient safety is a major concern of the general public, health care services, and policymakers. The WHO (World Health Organization) describes patient safety as a health care discipline that emerged with the evolving complexity of health care systems and the resulting rise of patient harm in health care facilities [1]. Agencies working on patient safety and healthcare systems aim to prevent and reduce risks, errors, harm, or death that occur to the patients during the provision of healthcare [2-4]. Patient safety is nowadays widely accepted with a clear consensus to deliver quality health services effectively, safe, and people-centered across the world. However, the occurrence of adverse events due to unsafe care is likely one of the ten leading causes of death and disability in the world [5].

In all healthcare systems matching care demand and supply reduces unsafe care and safety incidents. However, it is challenging to meet the demand from the given supply, mainly due to triaging, rationing of care, and adverse events which might occur during the care process [6]. One of the key aspects of care demand is capacity utilization in hospitals. Safety tipping points could be from a change in some of the factors resulting in the significant impact on outcomes i.e., high capacity utilization that increases the risk of patient safety incidents, such as in-hospital mortality [7,8]. In hospitals, high capacity utilization, i.e., high bed occupancy rates, is provoking safety tipping points [7]. To understand whether the entire hospital system meets the care demand like the volume, turnover, and severity of patients, it is important to investigate safety tipping points from the care demand perspective and factors, such as comorbidities associated with negative health outcomes- like mortality, to improve patients' care.

1.1 Patient safety events and in-hospital mortality

The Agency for Health Research and Quality (AHRQ) has categorized patient safety events as incidents, near misses, and unsafe conditions [9]. A patient safety incident is an event that reached the patient, regardless of harm (including adverse and sentinel events) whereas a near miss is an event that did not reach the patient (without known harm), but has the potential for harm [9,10], where unsafe conditions are circumstances e.g. at the workplace that increase the probability of safety events. The most common patient safety events are related to blood and blood products,

healthcare-associated infections, falls, medications, devices, or surgery [9]. Moreover, there is a large group of other events/conditions, which might be unknown/unreported such as delay in treatment, high capacity utilization, specimen mishandling, telemetry issues (staff's perception of event categorization), or staff's target on high-risk patients [11]. Events classified as "others" are mainly due to indirect causes rather than disease severity itself.

Relating safety events to all-cause death in hospitals, the American Data Network-Patient Safety Organization (AND-PSO) reported 83% of all deaths occurred in three event types, specifically "others", healthcare-associated infections, and surgery or anesthesia [11,12]. More specifically, the others category holds more than 60% of all-cause death. In 1999 the Institute of Medicine's (IOM) report *To Err Is Human* estimated between 44 to 98 thousand deaths per year occurring from medical errors in the United States [13]. Similarly, previous studies have shown that up to 40% of patients experienced safety events in acute care hospitals resulting in prolonged hospital stays or mortality [14-20]. In-hospital mortality is the ultimate harm to the patients and their relatives can experience and is often due to the indirect/system-level factors which are less often investigated. Among these causes, high capacity utilization or safety tipping points in hospitals have been linked with in-hospital mortality [7,8].

1.2 Safety tipping points in hospitals

Hospitals experiencing safety issues are rarely investigated with an indirect cause like capacity utilization and safety tipping points. The theoretical foundation of capacity utilization was provided by Johansen in 1968 *as the maximum that can be produced by a production unit with fixed and variable inputs for a given period and provided that the availability of variable factors of production is not restricted* [21] and several studies have utilized Johansen's definition of capacity utilization in healthcare [22-24]. In the hospital context, capacity utilization is simply defined as the percentage of beds occupied per unit time, which is driven by key variables such as the number of admissions and discharges and the length of stay [24,25]. For example, when bed occupancy is low, staff experience less work pressure, and when bed occupancy gradually increases work pressure is also increased with relatively constant resources.

Subsequently, at very high occupancy levels, work pressure can exceed a critical point – also referred to as safety-tipping point [7,24,26]. At this point, managers are unable to respond adequately, and consequently, quality of care and patient safety deteriorates during periods of high occupancy [7].

The study from Kuntz et al. described the safety tipping point in terms of capacity utilization in German hospitals, where they reached the critical point at 92.5% of bed occupancy, reporting one in seven death at a bed occupancy rate above this critical point.[7]. No other papers have explained high capacity utilization in hospitals as tipping points, but few papers have reported the relation between high bed-occupancy and in-hospital mortality in Danish [8] and English hospitals [27,28]. In these studies, the tipping point phenomenon in hospitals involves two main components of the health service i.e., the demand of capacity utilization and the resources of providers influencing the outcomes (Figure 1.1). These parts are incorporated in system approach models [38] showing the interactions and the possible causal relations between them. However, the models do not necessarily show when and how these relations exist. With the different times and processes, when systems are under strain (e.g., high workload) the relation could show negative health outcomes such as mortality. This thesis describes safety tipping points in Swiss hospitals based on capacity utilization and how these might influence adverse outcomes such as mortality.

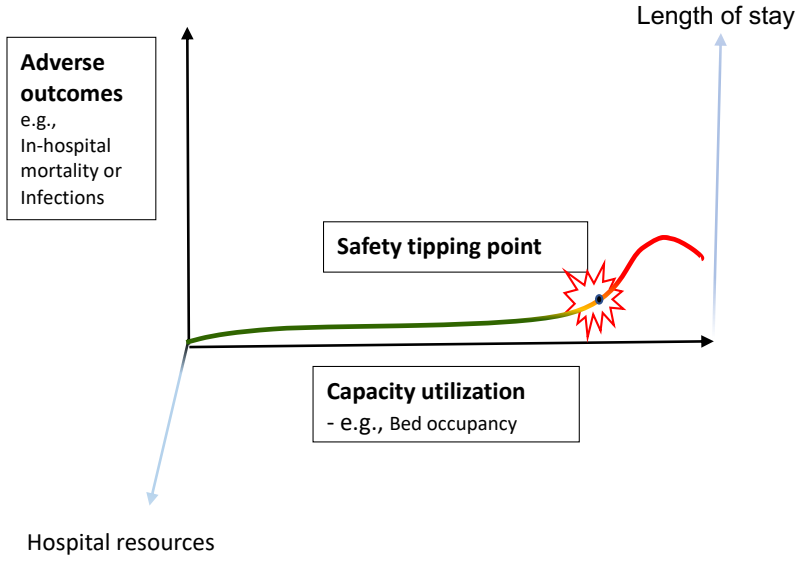


Figure 1.1 The safety tipping point phenomenon in hospitals, with the major components, capacity utilization and adverse outcomes (Adopted from Kuntz et al. (2014))

1.2.1 Factors influencing safety tipping points in hospitals

Patient safety and tipping points depend mainly on the safety practices in hospitals. The AHRQ defines patient safety practices as *a type of structure or process whose application reduces the probability of adverse events or in-hospital mortality resulting from exposure to the health care system across a range of diseases and procedures* [29]. Safety practice refers to the process and structure of hospitals (e.g., balancing higher skilled and lower skilled staff in a unit during the weekend), which aims in reduction of errors and maintaining the system under pressure. Overall safety practices could be described as the context (i.e., the structure of the hospitals), care demand, and supply. Context refers to the structure or the types of hospitals and the time-related changes in hospitals "patterns or trend", which varies with days of the week. For instance, in Switzerland there are five different types of general hospitals, depending on size, the number of cases treated, the combination of hospital units, and level of care delivery, described by the Swiss medical association [30,31]. Namely, they are university hospitals, tertiary care hospitals and large basic hospitals, medium basic hospitals, and small basic hospitals. Safety practices differ between hospital types, as fewer patients visit small hospitals and more patients in university hospitals depending on the resources and size [30]. Thus, this changes the tipping points of capacity utilization by hospitals and over time.

Care demand in hospitals refers to the actual workload in hospitals driven by the headcounts and severity (clinical complexity) of each case and varying with respect to time/days [32]. Headcounts refer to the number of patients throughout the day [7], and actual patients discharged or admitted in a day [33]. On the other hand, patient severity refers to the level of severity/complexity of patients depending on comorbidities and clinical complexity which are often reported in the form of ICD codes [34]. All these components of care demand are interrelated to increase or decrease the level of safety tipping points in hospitals.

Care supply directly refers to the level of services or resources available in hospitals, such as the staffing mix of nurses, doctors, and other care assistants during the patient's hospital stay. The change in the nursing staff's between a registered nurse and nursing assistant could influence patient safety outcomes like in-hospital

mortality [35]. Therefore, the level of care supply might influence safety tipping points in hospitals, however, care supply largely depends on care demand [32].

1.2.2 Variation of time-varying factors of safety tipping points in hospitals

Another important consideration of safety tipping point is the time-varying nature of care demand and supply in hospitals. The hospital system is a continuous process, where patients or staff enter and leave each hour or day, or day of the weeks throughout the year. The longitudinal variation of exposures and confounders of demand and supply in hospitals create strains at certain points in time (i.e., too high demand given the available resources) which could result in adverse outcomes. For instance, the daily change in capacity utilization leads to more patients per day during weekdays and fewer patients per day on a weekend in a normal working week [36]. Patient admission and discharge might also show a similar trend [37]. Similarly, the patient's severity of disease in hospital units and staffing mix varies constantly throughout the day and days of the week [32]. All these variations in demand should match with each hospital's supply/resources for safety practice. Thus, investigating tipping points through capacity utilization requires a description and understanding of all factors contributing to the demand-side variation over time.

1.3 System thinking representing tipping points phenomenon in hospitals

System thinking is a perspective that targets the understanding how elements of system influence one another. An example of nature is an ecosystem. System thinking in healthcare applies scientific insights to understand the elements that influence health outcomes; models the relationships between those elements (e.g., causes and effects); and alters design, processes, or policies based on the resultant knowledge to produce better health at lower cost [38]. System thinking brings most of the factors of safety tipping points into three interacting parts of a complex adaptive system in health; 1) provider system: a clinic or hospital where physicians or nurses provide care and support for patients; 2) healthcare system: a collection of clinical programs and centers that are the part of a large organization; 3) patient's system: patients interacting with families, friends, their providers, and healthcare system [38,39]. As explained above system thinking provides a perspective on the interplay between care

demand and supply in hospitals which could also be described with the iceberg model of system thinking. For instance, care demand factors and their patterns over time are the parts of the hospital system that influence daily capacity utilization and ultimately safety tipping points.

1.3.1 Iceberg model of system thinking in hospitals

The iceberg model of system thinking (Figure 1) visualized what you see above the waterline is only the tip of the problem (e.g., in-hospital mortality) [40-42]. However, what lies beneath the surface provides a much deeper understanding of the problem [41]. For example, high capacity utilization with comorbidities (trend or pattern) can reveal one of the underlining causes of safety events or mortality in hospitals. The model includes the structure of the hospitals or health system and values or beliefs or patient's experience of health service, also triggers of the problem [11,43]. The foundations of the problem presented under the water in the iceberg model are explained via causal loops, feedback, and variations in the healthcare system [40]. Although simplistic, the model provides a linkage within a hospital system to define the underline roots of in-hospital mortality through the safety tipping phenomenon, in particular the capacity utilization in hospitals.

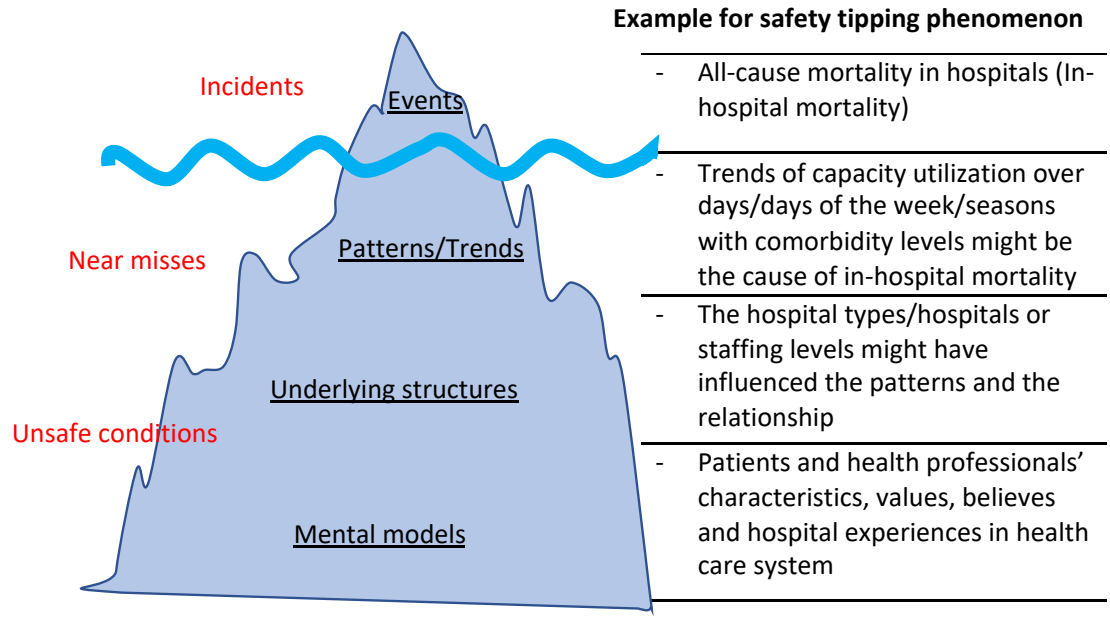


Figure 1.2. The iceberg model of system thinking for in-hospital mortality (Adapted from The Iceberg Model by Goodman, 2002)

1.4 Medical comorbidities in hospitals

Along with the hospital and healthcare system factors influencing the patient safety, patients with comorbidities are more at risk, influencing tipping points and particularly the outcome of in-hospital mortality. For example, acute renal failure in cancer patients is close call or near miss for in-hospital mortality. The concept of patient complexity has not been widely used in healthcare but has different meanings across medicine, nursing, and social work disciplines. Patients' complexity is represented or described with different terms, like comorbidity, multimorbidity, poly-pathology, dual diagnosis, or multiple chronic conditions [44]. Along with the different term's comorbidity is probably the oldest used in medicine to describe complexity, which is the presence of one or more additional clinical conditions often co-occurring with a primary condition [45]. Comorbidity is associated with worse health outcomes including mortality, more complex clinical management, and increased health care costs, which needs appropriate adjustment to reduce bias in observational studies [45,46]. Studies have suggested comorbidity is used as a personal-level construct with four major types of distinction: 1) the nature of health condition (e.g., ICD 10), 2) the relative importance of co-occurring conditions (e.g., mental illness and substance abuse), 3) the chronology of the conditions (e.g., patient with diabetes might diagnose with depression with time), 4) expended concepts (e.g., morbidity burden or severity/complexity) [45].

1.4.1 Comorbidities for risk adjustment of in-hospital mortality

Various approaches have been taken to characterize the combined burden of diseases or conditions as a single measure of scale or index. The Charlson index, (developed in 1987) [47] and the Elixhauser index, (developed in 1998) [48], both comorbidity indices are the most popular measures. The Charlson index is mainly used for chronic conditions and the Elixhauser index is used for both acute and chronic conditions. For hospital data, ICD codes are used to classify each of the comorbidities within each comorbidity index [49]. Moreover, the Elixhauser index is statistically superior to the Charlson index to predict in-hospital mortality [46,50].

Moreover, comorbidity weight differs from the comorbidity count because the odds of dying due to diabetes are not the same as the odds of dying from metastatic cancer.

Therefore, Charlson weight was developed and used mainly for the Charlson index, however different weights were developed for the Elixhauser index depending on the sample and context of the data [51,52]. van Walraven et al (2009) developed a composite score (VW score) for the Elixhauser comorbidities by modeling in-hospital mortality with inpatient admission data (Ottawa Hospital, Ontario, Canada, 1996–2008) [53]. In the same way, country or regional-specific comorbidity weights were developed and verified using local data [46,50]. A weighted sum score/summary measure provides an attractive advantage over plain dummy variables [46,54], as it reduces the overfitting risk of more parameters, unjustifiable in small datasets [46] and limits computational requirements in large ones [54].

Several studies are using comorbidities as important individual factors to predict in-hospital mortality or other outcomes like cancer, rehospitalization, length of hospital stay, hospital discharge using both comorbidity counts and weighting scores [46,50,51,55,56]. Furthermore, it is adjusted to eliminate confounding created by disease severity in the causal relationship between treatment and outcome [57-59] and support policymakers to control health care costs and improve quality via Diagnosis-related groups (DRGs) [60]. Thus, for the safety tipping phenomenon in hospitals, comorbidities are an important predictor influencing capacity utilization and in-hospital mortality including reimbursement policy, which needs appropriate adjustment.

1.5 Causal inference in an observational study

Knowing both the system-level variables and individual patient-level variables in hospitals, provide evidence of system dynamics offering a nexus between variables in a meaningful way. Moreover, safety tipping points could vary across hospitals depending on the size of the hospital and demand load and it is further challenged by the time-varying nature of variables (e.g., capacity utilization). Particularly, one of the causes of mortality in a hospital could be high capacity utilization which needs appropriate tools/instruments to define the nexus and the methods to estimate its effect with the outcome through causal inference perspective.

Causal inference stated by Hernán and Robins is a complex scientific task that relies on the triangulation of evidence from different sources and the application of a variety

of methodological approaches [59]. Moreover, Pearl describes causality, as a scientific friendly language for articulating knowledge and mathematical machinery for processing and combining knowledge and data for drawing new causal conclusions [61]. These definitions exhibit somehow differently from experimental to observation studies. Results from a well-conducted randomized controlled trial (RCT) should ideally inform on the comparative merits of treatment or exposure choices for a health outcome in an experimental study [62]. On the other hand, with observation studies, outcome comparison is much less convincing. A key reason not to endow observational associations with causal interpretation is the lack of randomized treatment/exposure assignments [59]. However, with an increase in methodological approaches and relative ease in computation, causal inference from observational data then revolves around the hope that the observational study can be viewed as a conditionally randomized experiment to draw causal relationships with assumptions, like ignorability and positivity [59,61-63].

The system approach in hospitals bridges patient and hospital characteristics with the context of the healthcare system. Within hospitals, some variables are exposures, some of them are confounders or mediators and some of them are outcomes. Particularly, capacity utilization is a natural exposure of patients in hospitals for the day's patient stays e.g., patients exposed to high or low-capacity utilization on each day, rather than randomized exposure in experimental study e.g., the number of patients with treatment A vs. a number of patients with usual treatment. In observational study causal interpretation is further challenged with the computation including confounders, methodology, and assumptions. However, hospitals generate large amounts of real data [64] regularly with different aspects of capacity utilization, other care demands, supply and outcomes variables, and causal inference provides means to enable researchers to identify links from the observational data [65].

For instance, inferring the effect of high and low-capacity utilization [8,27] and the difference in potential outcomes at a particular point of time could show the causal effect between the two exposures. For example, if a patient is exposed to high capacity utilization he/she is more likely to have negative health outcomes [7] and vice versa. After adjusting for system and individual factors which influence the relationship, the true difference in estimate between the high and low-capacity utilization shows the

causal effect between them. The time fixed or cross-sectional settings are helpful to generate causal hypotheses. However, in health service research, we have time-varying covariates of demand and supply, instrumental variables, individual and system-level variables confounding the relationship. Thus, it needs a causal inference perspective to represent the problem through directed acyclic graphs (DAG), to identify potential biases through time-varying covariates, and the appropriate methods to measure the causal effect

1.5.1 Representing causal inference through Directed Acyclic Graphs

Directed acyclic graphs (DAGs) are popular tools for identifying appropriate adjustment strategies for causal inference in epidemiological and applied health research [66,67]. DAGs provide a simple and transparent way for scientists to identify and demonstrate their knowledge, theories, and assumptions about the causal relationships between variables [67]. DAGs help to synthesis evidence to guide risk adjustment using (direction) d-separation rules given by Pearl in 1995 [68] following causal inference principles (define, assume, identify and estimate) [61,69-71]. Mainly, representing causal inference using DAG includes three main stages of mapping, translation, and integration [67].

Mapping includes defining the outcome, exposure, controlled, mediators, and instrumental variables (nodes) from the literature and expert knowledge [67,72]. The directed or undirected edges are drawn originating from control or confounders to exposures and outcomes (direction does not matter until the translation stage). The next stage is translation, where posited relationship and its reverse are both assessed, where the edges may be posited, reversed, or bi-directional [59]. If not, they are deleted. The third stage is integration, the translated DAGs are combined into one by synthesizing all indexed directed edges and the nodes are recombined to reduce complexity and to established consistency [67].

From the literature, studies have shown the association between capacity utilization and in-hospital mortality [7,8,27]. It is also known that comorbidity influences capacity utilization increasing length of stay and directly the mortality; staffing, patient turnover, hospital types, age, sex are (directly/indirectly) related to both capacity utilization and mortality [7,50,51]. These relations and factors influencing

overall patients safety in hospitals can be explored in DAGs, and the issues of time-varying covariates can be highlighted and explained through DAGs building process of mapping, translation, and integration [73].

For instance, locomotive diseases (e.g., musculoskeletal system and fractures) and/or respiratory diseases are independent causes of hospitalization, i.e., the collider (common effect-the two arrowheads collide into hospitalization) (Figure 1.3). If the collider or death is controlled by a researcher in a general population dataset, it induces collider bias and induces a distorted association between the locomotive and respiratory diseases. Since the path from locomotive to respiratory diseases is opened at hospitalization. The DAG in pink lines shows the error when adjusting the collider or descendent of the collider, which is a particular example of a bias provided by Sackett in his study [74]. The DAG visually represents the bias and suggests researchers exclude colliders or descendants of the collider to control during risk adjustment.

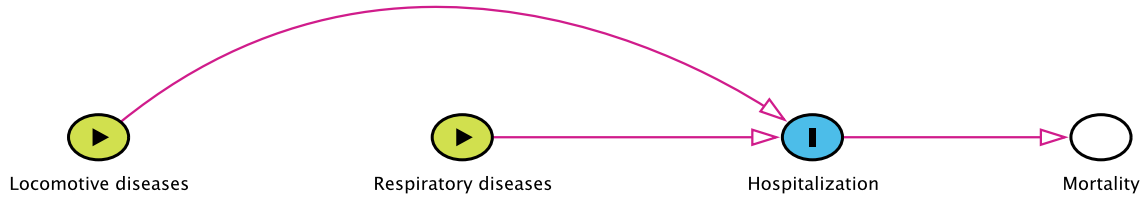


Figure 1.3 Illustration of collider bias using causal DAG between locomotive and respiratory diseases

1.5.2 Causal inference with time-varying variables

Time-varying covariance occurs when a given covariate changes over time during the follow-up period (e.g., hemoglobin level during pregnancy), which is a common phenomenon in clinical and health service research [75]. Under this condition, each of the time-varying variables could change over time, that could be treatment/exposure or confounders, or both. Essentially, we need to integrate these changes to assess the total causal effect of treatment/exposure on outcomes.

Many questions arise when treatment or exposure varies over time that is potentially causal. For example, medical treatment may take different values for a single individual over time, which would be described as **time-varying treatments**. Time-varying treatments are defined clearly with a plan, policy, protocol, or regimes a rule

to assign treatment at each time of follow-up [59,68]. Very popular medical treatment strategies are "always treat" and "never treat" [59]. This applies the same principle for time-varying exposures if they are binary. However, in the continuous exposure, the threshold of treatment strategy could be determined based on units or dose e.g., CD4 cell counts in HIV patients [59,76]. Time-varying treatments are also influenced by other variables called **time-varying confounders**. Many exposures of epidemiological interest are time-varying; for example, treatment dose, body mass index, smoking status, blood pressure, depression, air pollution, socioeconomic status [77]. For instance, in hospitals, daily capacity utilization might be influenced by daily staffing, patient turnover, and the severity of the patients.

The sequence of time-varying treatments influenced by respective time-varying confounders for each time point requires adjustment of time-varying confounders to validly estimate the causal effect of any treatment. However, time-varying confounding can also be affected by past exposure and also exists when there is feedback between the exposure and outcome, which is called **treatment-confounder feedback** (TCF) [59,77]. TCF is a key problem in causal inference for time-varying treatments. When treatment-confounder feedback exists, traditional adjustment methods may introduce bias in the effect estimates. That is, even if we had all the information required to validly estimate the average causal effect of any treatment on the outcome, we would generate a biased estimate using traditional methods [59].

Stratification and outcome regression are the commonly used methods to adjust for confounding, mainly with time-fix variables. Stratification means estimating the association between treatment and outcome in subsets–strata–of the study population defined by the confounders [59]. Outcome regression generally fits the model conditions on the treatment variable for effect estimation in causal inference [59,78]. Conditioning on confounders that are affected by previous treatment can create selection bias even if the confounder is not on a causal pathway between treatment and outcome [59]. Moreover, in time-varying variables, the confounder for subsequent treatment lies on a causal pathway from earlier treatment to an outcome that needs adjustment to prevent the back-door path. Thus, **traditional methods cannot be fixed** and essentially fail to control for TCF.

1.5.3 Causal effect estimation with time-varying variables

Conventional methods to perform longitudinal data analysis such as linear mixed-effects models and generalized estimating equations models are capable of handling time-varying covariates [79]. However, in the case of time-varying variables, the probability of treatment exposure differs across time, which requires the application of time-varying weights on the unit of analysis (e.g., individual subjects) and to eliminate TCF. In this context, Robins' generalized methods (G-methods) like G-computation, [80] g-estimation for Structural Nested Models [81], and inverse probability of treatment weighting (IPTW) for Marginal Structural Models (MSMs) [82] have been proposed as alternatives to standard regression adjustment for confounding, to overcome the aforementioned problem.

An alternative to regression adjustment is to utilize so-called IPTWs for MSMs to account for bias due to observed confounders as well as unobserved confounders [79]. The fundamental difference between regression adjustment and approaches using IPTW is that the former models the relationship between a covariate and the outcome, whereas the latter models the relationship between the covariate and the putative cause (i.e., treatment assignment) [63]. IPTWs are generated applying time-varying confounders and other covariates to a treatment/exposure, which stabilized the weights of exposures/treatments incorporating MSMs [79] holding the assumptions of causality [63].

G-computation or G-formula is a time-dependent and model-based generalization of standardization (i.e., can better handle heterogeneity involving time-varying exposure and confounding) [77,83]. This method, based on weighted averaging and a standard population, used to remove the effects of differences in age or other confounding variables (which yields the standardized mean outcome (e.g., risk)), underexposure regimens of interest (e.g., always treat and never treat) by averaging the confounder specific mean outcomes under that regimen over the joint distribution of confounders across the entire follow-up [83,84]. Moreover, these methods' computation is more complex with longitudinal exposure and may not fit with all types of data.[85]

G estimation is a two-step procedure that uses two models to estimate the causal effect [77]. The first is a causal model that includes the causal variable of interest and links the counterfactual outcome under no exposure during the follow-up (i.e., the outcome that would have been observed under no exposure during the follow-up) to the weighted sum of time spent in a given exposure status [44,77]. The second is a logistic regression for predicting exposure at each visit/timepoint based on the previous exposure and covariate histories and the counterfactual outcome[44,86]. However, it is rarely used in practice, due to a lack of off-the-shelf software [44].

On the contrary, IPTW for MSM estimation is by far the most popular method because, though complex, it appears much simpler than the other two alternative estimation methods [87] and the IPTW method relies on correct exposure model specification like binomial, continuous, Poisson distribution. IPTW for MSM is not widely used in health service research but is more frequently seen in epidemiological research. However, the IPTW estimator is applicable for capacity utilization exposure with binomial or continuous distribution and MSM is flexible for a binary outcome like in-hospital mortality for adjusting for both time-varying and time-fix covariates.

1.6 Research gaps and rationale of this dissertation

Patients admitted to the hospital desire safe care and treatment without harm. However, the hospital's workplace conditions, human errors, and unsafe acts alongside disease severity and comorbidities of patients serve as contributing factors for negative outcomes like mortality in hospitals. The structure, size, and capacity utilization pattern of the hospitals, which are hidden parts of the iceberg model sometimes play a role in in-hospital mortality and are less often measured. Moreover, several conditions of patient severity, capacity utilization of the hospitals, and individual factors collectively influence in-hospital mortality. Three research gaps have been identified in the literature and build the basis of this dissertation.

Firstly, some of the studies have been conducted in attempts to describe care demands in hospitals in terms of bed occupancy and patient turnover as headcount measures. These measures were concentrated on monthly or yearly aggregation for a single hospital or some hospital departments. Moreover, daily headcounts of patients in hospitals have seldom being investigated along with the clinical complexity (severity)

of patients in a longitudinal perspective. Describing care demand in both headcount and severity, longitudinally for a year utilizing all national general hospital data would reflect the variation of the healthcare system of Swiss hospitals. Performing such assessments can support healthcare hospital managers, healthcare planners, and policymakers in better observing daily care demand dynamics (e.g., variations by day of the week, weekend vs weekdays, seasons, and holidays) to identify and monitor likely work pressure within hospitals.

Secondly, as the notion of the complex patient, comorbidity plays a vital role in hospitals utilizing more resources and time for care and it's a strong predictor of in-hospital mortality. However, there is no uniformity in measuring comorbidity indices to predict in-hospital mortality in terms of methods and comorbidity weights of the inpatient population. Utilizing a nationwide inpatient multi-year dataset of Swiss hospitals provides the opportunity in deriving new comorbidity weights with improved predictive performance over the existing and most popular comorbidity weightings. The population-based comorbidity weights could be an optimized weighting system for considering comorbidities for researchers and health service planners for population-based analysis of health outcomes.

Thirdly, some of the studies have shown the association between capacity utilization and in-hospital mortality but were inadequate to demonstrate a causal link between them. Causal inference in observational studies highlights the importance of considering exposure of high capacity utilization to cause in-hospital mortality. A longitudinal study on care demand factors is pivotal to gain an understanding of the nature of time-varying exposures and confounders. Moreover, individual covariates like comorbidity weighting scores are important predictors influencing both exposure and outcome for risk adjustment causal models. The exposure (treatment) strategy of capacity utilization across different hospitals and its causal effect might provide possible safety tipping points in hospitals. To our knowledge, no observational study to investigate causal inference of capacity utilization and in-hospital mortality with time-varying variables has been conducted.

Finally, this study will contribute to the understanding of the hospital's system exploring care demand factors. Moreover, in Swiss hospitals, care demand with capacity utilization and disease severity is linked with staffing/resources to reduce

the cost and increase efficiency as a rationale to support Swiss DRG. Thus, hospital-level factors including patient's complexities could be the causes of patient's negative health outcomes and in-hospital mortality, which needs appropriate tools and methods like, causal inference to adjust and justify the relationships. Further, it will extend the knowledge for effective utilization of routine hospital data to identify work pressure in hospitals via tipping points analysis and causality in the observation study.

1.7 References

1. Carayon P, Wood KE. Patient safety. Information Knowledge Systems Management. 2009;8(1-4):23-46.
2. Runciman W, Hibbert P, Thomson R, Van Der Schaaf T, Sherman H, Lewalle P. Towards an International Classification for Patient Safety: key concepts and terms. International journal for quality in health care. 2009;21(1):18-26.
3. Duckett S, Jorm C. All complications should count: Using our data to make hospitals safer. Grattan Institute;2018.
4. Global Ministerial Summit on Patient Safety. Tokyo Declaration on Patient Safety. http://www.mhlw.go.jp/file/06-Seisakujouhou-10800000-Iseikyoku/0000204005.pdf. Published 2018. Accessed [2021.02.10].
5. Jha AK, Larizgoitia I, Audera-Lopez C, Prasopa-Plaizier N, Waters H, Bates DW. The global burden of unsafe medical care: analytic modelling of observational studies. BMJ quality & safety. 2013;22(10):809-815.
6. Leotsakos A, Zheng H, Croteau R, Loeb JM, Sherman H, Hoffman C, et al. Standardization in patient safety: the WHO High 5s project. International journal for quality in health care. 2014;26(2):109-116.
7. Kuntz L, Mennicken R, Scholtes S. Stress on the ward: Evidence of safety tipping points in hospitals. Management Science. 2014;61(4):754-771.
8. Madsen F, Ladelund S, Linneberg A. High levels of bed occupancy associated with increased inpatient and thirty-day hospital mortality in Denmark. Health Affairs. 2014;33(7):1236-1244.
9. AHRQ. AHRQ Common Formats - Hospital Version 1.2. https://www.psoppc.org/psoppc_web/publicpages/supportingDocsV1.2. Published 2013. Accessed [2021.02.09].
10. Collins S, Couture B, Dykes P, Schnipper J, Fagan M, Benneyan J, et al. Implementation, evaluation, and recommendations for extension of AHRQ Common Formats to capture patient-and carepartner-generated safety data. JAMIA open. 2018;1(1):20-25.
11. ADN. Patient Safety Analysis: "Other" Events Reveal 3X Higher Mortality Rate, Among Key Findings. https://blog.americandatanetwork.com/patient-safety-analysis-other-events-reveal-3x-higher-mortality-rate-among-key-findings. Published 2019. Accessed [2021.02.09].
12. Wright A, Bates DW. Distribution of problems, medications and lab results in electronic health records: the Pareto principle at work. Applied clinical informatics. 2010;1(1):32.
13. Shekelle PG, Wachter RM, Pronovost PJ, Schoelles K, McDonald K, Dy S, et al. Making health care safer II: an updated critical analysis of the evidence for patient safety practices. Evidence report/technology assessment. 2013(211):1.
14. Zegers M, De Bruijne M, Wagner C, Hoonhout L, Waaijman R, Smits M, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. Quality and safety in health care. 2009;18(4):297-302.
15. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. BMJ Qual Saf. 2012:BMJ's-2012-001159.
16. Halfon P, Staines A, Burnand B. Adverse events related to hospital care: a retrospective medical records review in a Swiss hospital. International Journal for Quality in Health Care. 2017;29(4):527-533.

17. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events Institute for Healthcare Improvement;2009.
18. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. 'Global trigger tool'shows that adverse events in hospitals may be ten times greater than previously measured. Health affairs. 2011;30(4):581-589.
19. Schwendimann R, Blatter C, Dhaini S, Simon M, Ausserhofer D. The occurrence, types, consequences and preventability of in-hospital adverse events–a scoping review. BMC Health Services Research. 2018;18(1):521.
20. Musy SN, Ausserhofer D, Schwendimann R, Rothen HU, Jeitziner M-M, Rutjes AW, et al. Trigger Tool–Based Automated Adverse Event Detection in Electronic Health Records: Systematic Review. Journal of medical internet research. 2018;20(5).
21. Johansen L. Production functions and the concept of capacity. Recherches récentes sur la fonction de production, Collection, Economie mathématique et économétrie. 1968;2:52.
22. Karagiannis R. A system-of-equations two-stage DEA approach for explaining capacity utilization and technical efficiency. Annals of Operations Research. 2015;227(1):25-43.
23. Rechel B, Wright S, Barlow J, McKee M. Hospital capacity planning: from measuring stocks to modelling flows. Bulletin of the World Health Organization. 2010;88(8):632-636.
24. Arfa C, Leleu H, Goaied M, van Mosseveld C. Measuring the capacity utilization of public district hospitals in Tunisia: using dual data envelopment analysis approach. International journal of health policy and management. 2017;6(1):9.
25. Adam T, Ebener S, Johns B, Evans DB. Capacity utilization and the cost of primary care visits: Implications for the costs of scaling up health interventions. Cost effectiveness and resource allocation. 2008;6(1):22.
26. Green LV, Liu N. A study of New York City obstetrics units demonstrates the potential for reducing hospital inpatient capacity. Medical Care Research and Review. 2015;72(2):168-186.
27. Boden D, Agarwal A, Hussain T, Martin S, Radford N, Riyat M, et al. Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital. Emergency Medicine Journal. 2016;33(2):85-90.
28. Jones R. Bed occupancy and hospital mortality. British Journal of Healthcare Management. 2016;22(7):380-381.
29. Agency for Healthcare Research and Quality. Making Health Care Safer. U.S. Department of Health & Human Services. https://archive.ahrq.gov/research/findings/evidence-based-reports/services/quality/er43/ptsafety/summary.html. Published 2001. Accessed May 24, 2018.
30. De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288.
31. Federal Statistical Office. Health Pocket Statistics 2017. Federal Statistical Office (FSO). https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.4342090.html. Published 2018. Accessed 2020-10-02.
32. Musy SN, Endrich O, Leichtle AB, Griffiths P, Nakas CT, Simon M. Longitudinal Study of the Variation in Patient Turnover and Patient-to-Nurse Ratio: Descriptive Analysis of a Swiss University Hospital. Journal of medical Internet research. 2020;22(4):e15554.

33. Park SH, Blegen MA, Spetz J, Chapman SA, De Groot H. Patient turnover and the relationship between nurse staffing and patient outcomes. Research in nursing & health. 2012;35(3):277-288.
34. Kim S, Choi B, Lee K, Lee S. An Evaluation of the Patient Clinical Complexity Level (PCCL) Method for the Complexity Adjustment in the Korean Diagnosis-Related Groups (KDRG). 2020.
35. Griffiths P, Maruotti A, Saucedo AR, Redfern OC, Ball JE, Briggs J, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ quality & safety. 2019;28(8):609-617.
36. Fusco D, Saitto C, Arcà M, Ancona C, Perucci CA. Cyclic fluctuations in hospital bed occupancy in Roma (Italy): supply or demand driven? Health services management research. 2003;16(4):268-275.
37. VanFosson CA, Yoder LH, Jones TL. Patient turnover: A concept analysis. Advances in Nursing Science. 2017;40(3):300-312.
38. Kaplan G, Bo-Linn G, Carayon P, Pronovost P, Rouse W, Reid P, et al. Bringing a systems approach to health. In. *NAM Perspectives, Discussion Paper*. Washington DC: National Academy of Medicine; 2013.
39. Wang F. Measurement, optimization, and impact of health care accessibility: a methodological review. Annals of the Association of American Geographers. 2012;102(5):1104-1112.
40. Monat JP, Gannon TF. What is systems thinking? A review of selected literature plus recommendations. American Journal of Systems Science. 2015;4(1):11-26.
41. Ecochallenge.org. The Iceberg Model. http://www.ascd.org/ASCD/pdf/journals/ed_lead/el200910_kohm_iceberg.pdf. Published 2009. Accessed [2021.02.09].
42. Goodman M. The iceberg model. Hopkinton, MA: Innovation Associates Organizational Learning. 2002.
43. Geulayov G, Casey D, McDonald KC, Foster P, Pritchard K, Wells C, et al. Incidence of suicide, hospital-presenting non-fatal self-harm, and community-occurring non-fatal self-harm in adolescents in England (the iceberg model of self-harm): a retrospective study. The Lancet Psychiatry. 2018;5(2):167-174.
44. Dukes O, Vansteelandt S. A note on G-estimation of causal risk ratios. American journal of epidemiology. 2018;187(5):1079-1084.
45. Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for understanding health and health services. The Annals of Family Medicine. 2009;7(4):357-363.
46. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, et al. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374.
47. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-383.
48. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998:8-27.
49. Quan H, Eastwood C, Cunningham CT, Liu M, Flemons W, De Coster C, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ open. 2013;3(10):e003716.
50. Menendez ME, Neuhaus V, Van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical Orthopaedics and Related Research®. 2014;472(9):2878-2886.

51. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009:626-633.
52. Chang H-J, Chen P-C, Yang C-C, Su Y-C, Lee C-C. Comparison of Elixhauser and Charlson methods for predicting oral cancer survival. Medicine. 2016;95(7).
53. Manning E, Gagnon M. The complex patient: a concept clarification. Nursing & health sciences. 2017;19(1):13-21.
54. Te Grotenhuis M, Pelzer B, Eisinga R, Nieuwenhuis R, Schmidt-Catran A, Konig R. When size matters: advantages of weighted effect coding in observational studies. International Journal of Public Health. 2017;62(1):163-167.
55. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC Health Serv Res. 2021;21(13).
56. Fortin Y, Crispo JA, Cohen D, McNair D, Mattison DR, Krewski D. Optimal look back period and summary method for Elixhauser comorbidity measures in a US population-based electronic health record database. 2017.
57. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. Annals of the American Thoracic Society. 2019;16(1):22-28.
58. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. American journal of public health. 2018;108(5):616-619.
59. Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.
60. Goldfield N. The evolution of diagnosis-related groups (DRGs): from its beginnings in case-mix and resource use theory, to its implementation for payment and now for its current utilization for quality within and outside the hospital. Quality Management in Healthcare. 2010;19(1):3-16.
61. Judea P. An introduction to causal inference. The International Journal of Biostatistics. 2010;6(2):1-62.
62. Zwahlen M, Salanti G. Causal inference from experiment and observation. Evidence-based mental health. 2018;21(1):34-38.
63. Thoemmes F, Ong AD. A primer on inverse probability of treatment weighting and marginal structural models. Emerging Adulthood. 2016;4(1):40-59.
64. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health information science and systems. 2014;2(1):1-10.
65. Lee C, Luo Z, Ngiam KY, Zhang M, Zheng K, Chen G, et al. Big healthcare data analytics: Challenges and applications. In: Handbook of large-scale distributed computing in smart healthcare. Springer; 2017:11-41.
66. Tennant P, Murray E, Arnold K, Berrie L, Fox M, Gadd S, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. International Journal of Epidemiology. 2020.
67. Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. International journal of epidemiology. 2020;49(1):322-329.
68. Pearl J. Causal diagrams for empirical research. Biometrika. 1995;82(4):669-688.
69. Bello NM, Ferreira VC, Gianola D, Rosa GJ. Conceptual framework for investigating causal effects from observational data in livestock. Journal of animal science. 2018;96(10):4045-4062.

70. Matthay EC, Glymour MM. A Graphical Catalog of Threats to Validity: Linking Social Science with Epidemiology. Epidemiology (Cambridge, Mass). 2020;31(3):376.
71. Robins JM. Discussion of Causal diagrams for empirical research by J. Pearl. Biometrika. 1995;82(4):695-698.
72. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. American journal of public health. 2005;95(S1):S144-S150.
73. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. International journal of epidemiology. 2018;47(5):1714-1722.
74. Sackett DL. Bias in analytic research. In: The case-control study consensus and controversy. Elsevier; 1979:51-63.
75. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CG. Time-varying covariates and coefficients in Cox regression models. Annals of translational medicine. 2018;6(7).
76. Williams JR, Crespi CM. Causal inference for multiple continuous exposures via the multivariate generalized propensity score. arXiv e-prints. 2020:arXiv: 2008.13767.
77. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. Bmj. 2017;359.
78. Edin M. Outcome regression methods in causal inference: The difference lasso and selection of effect modifiers. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1221802&dswid=890. Published 2018. Accessed [2021.02.09].
79. Bounthavong M. Using inverse probability of treatment weights & Marginal structural models to handle time-varying covariates. https://rpubs.com/mbounthavong/IPTW_MSM_Tutorial. Published 2018. Accessed [2021.01.08].
80. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Mathematical modelling. 1986;7(9-12):1393-1512.
81. Vansteelandt S, Joffe M. Structural nested models and G-estimation: the partially realized promise. Statistical Science. 2014;29(4):707-731.
82. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Statistical models in epidemiology, the environment, and clinical trials. Springer; 2000:95-133.
83. Daniel RM, Cousens S, De Stavola B, Kenward MG, Sterne J. Methods for dealing with time-dependent confounding. Statistics in medicine. 2013;32(9):1584-1618.
84. Wang A, Nianogo RA, Arah OA. G-computation of average treatment effects on the treated and the untreated. BMC medical research methodology. 2017;17(1):1-5.
85. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. American journal of epidemiology. 2011;173(7):731-738.
86. Picciotto S, Eisen EA, Chevrier J. 0351 G-estimation: why does it work and what does it offer? Occupational and Environmental Medicine. 2014;71(Suppl 1):A120-A121.
87. Vansteelandt S, Sjolander A. Revisiting g-estimation of the effect of a time-varying exposure subject to time-varying confounding. Epidemiologic Methods. 2016;5(1):37-56.

Chapter 2

Study aims

This dissertation is embedded in the INVEST (Investigating Safety tipping points in Swiss hospitals) study. The INVEST study was conducted with the routine data of all Swiss hospitals from 2012 to 2017 obtained from the Federal Statistics Office (FSO) of Switzerland in August 2019. FSO data consist of inpatients' clinical and administrative data, hospital resources, and staffing data from acute care hospitals and special care (e.g., Paediatric, gynecological, surgical, pediatrics, etc.) facilities. This dissertation used inpatient data from acute care (general) hospitals to maintain a homogenous study population. The data was obtained after the approval of our application, subject to a data protection contract (as stipulated by article 22 of the Swiss Federal Act on Data Protection) and the FSO provided anonymized administrative and clinical data from all Swiss hospital inpatient cases. Altogether there were 8.84 million inpatients cases for six years for our study. For the specific study, data were further selected by the years of discharge for the analysis and applied the exclusion criteria where necessary.

Given the research gaps, this dissertation has three aims:

1. To investigate the longitudinal variation of daily care demand in general hospitals for three measures: 1) capacity utilization; 2) patient turnover; and 3) patient clinical complexity level (PCCL) (Chapter 3)

2. To validate and compare a set of new Swiss Elixhauser comorbidity weightings against the Charlson and Elixhauser-based van Walraven weights in an adult in-patient population-based cohort of general hospitals (Chapter 4)

3. To explore time-varying covariates, i.e., capacity utilization, patient turnover and clinical complexities and individual variables through directed acyclic graphs (DAGs) and to estimate the causal effect of capacity utilization on in-hospital mortality (Chapter 5)

Chapter 3

Variation of daily care demand in Swiss general hospitals: a longitudinal study on capacity utilization, patient turnover and clinical complexity levels

Narayan Sharma[1], MSc; René Schwendimann[1,2], PhD; Olga Endrich[3], MD; Dietmar Ausserhofer[1,4], PhD; Michael Simon[1,5], PhD

[1]Institute of Nursing Science (INS), Department Public Health (DPH), Faculty of Medicine, University of Basel, Basel, Switzerland
[2]Patient Safety Office, University Hospital Basel, Basel, Switzerland
[3]Directorate of Medicine, Inselspital University Hospital Bern, Bern, Switzerland
[4]College of Health-Care Professions Claudiana, Bozen, Italy
[5]Nursing Research Unit, Inselspital University Hospital Bern, Bern, Switzerland

Published in the Journal of Medical Internet Research (JMIR)
DOI: https://doi.org/10.2196/27163

3.1 Abstract

Background: Variations in hospitals care demand relies not only on the patient's volume but also on the disease severity. Understanding both daily severity and patients' volume in hospitals could help to identify pressure zones in hospitals for better hospital capacity planning and policymaking.

Objective: This longitudinal study explores daily care demand dynamics in Swiss general hospitals for three measures: 1) capacity utilization; 2) patient turnover; and 3) patient clinical complexity level (PCCL).

Methods: A retrospective population-based analysis was conducted with one-year routine data of 1.2 million inpatients from 102 Swiss general hospitals. Capacity utilization was measured as a percentage of the daily maximum number of inpatients. Patient turnover was measured as a percentage of the daily sum of admissions and discharges per hospital. PCCL was measured as the average daily patient disease severity per hospital from the clinical complexity algorithm.

Results: There was a pronounced variability of care demand in Swiss general hospitals. The average daily capacity utilization ranged between 57.8% (95% CI 57.3–58.4) and 87.7% (95% CI 87.3–88), patient turnover ranged from 22.5% (95% CI 22.1–22.8) to 34.5% (95% CI 34.3–34.7) and mean PCCL from 1.26 (95% CI 1.25–1.27) to 2.06 (95% CI 2.05–2.07) among hospitals. Moreover, both within and between hospitals, all three measures varied distinctly between days, between days of the week, between weekdays and weekends, and between seasons.

Conclusions: While admissions and discharges drive capacity utilization and patient turnover variation, disease severity of each patient drives PCCL. Monitoring—and if possible, anticipating—daily care demand fluctuations is key to managing hospital pressure zones. This study provides a pathway for identifying patients' daily exposure to strained hospital systems for a time-varying causal model.

Keywords: inpatient population; routine data; general hospitals, capacity utilization; clinical complexity

3.2 Introduction

Hospitals are constantly challenged by changing patient care demands. If this outweighs available resources, it can affect the quality of care and patient safety [1]. Demand factors include daily patient volume, turnover and clinical complexity of patients requiring diagnoses and therapies [1,2]. Responding to variations in any of these factors hospitals adjust their resource supplies, e.g., by changing shift-level staffing or resources, for each day, between workdays and weekends, for different seasons and throughout the year [3,4].

Capacity utilization, which is based on the number of beds occupied vs those available [5], offers one perspective to view hospital care demand [2,6,7]. Over recent decades, most healthcare systems' capacity utilization has increased, while total numbers of available beds have decreased [3,8]. This trend mainly reflects policies to reduce health care costs and to increase efficiency, e.g., by the use of Diagnostic Related Groups (DRGs) [8,9].

If capacity utilization is too high (e.g., above 80% or 85%), it might overburden healthcare systems and their workforces [10,11], possibly leading to adverse patient outcomes such as infections or even death [5,12,13]. Capacity utilization is high—exceeding 90%—in Canada, Israel and Ireland, followed by the UK, Norway and Switzerland, all of which report figures above 80% [14]. Within a country, capacity utilization also varies between hospital types, geographic regions and populations served [6,15]. In Switzerland, the most recent annual capacity utilization figures for acute care hospitals were between 70% and 82% [16,17].

As noted above, care demand also relies on patient turnover and patient complexity [18]. Patient turnover refers to admission and discharge or transfer of patients between units or hospitals [19] requiring resources [20]. "Census variability," "churn" or "environmental turbulence" cover the same or similar concepts [21]. Patient complexity refers to the severity/complexity of each patient's clinical needs. For instance, patients admitted to the ICU generally require more resources than those on a general ward, representing a resource-intensive caseload.

Disease severity is commonly measured via the Charlson/Elixhauser comorbidity or case-mix index; however, this does not include all relevant clinical conditions or morbidities, and its interpretation is commonly influenced by reimbursement policies

or cost of medication/treatment [22-24]. One alternative measure is the patient clinical complexity level (PCCL). As part of the German (G-)DRG system, PCCL reflects upon not only complications and co-morbidities, but their levels of clinical severity on a 5-point scale (i.e., 0-4; no, mild, moderate, severe, and very severe clinical complexity) [25]. A complex algorithm, depending on primary and secondary diagnoses and estimated severities allows to determine their cumulative effect per treatment episode [22,26,27]. A higher PCCL indicates a more complex and resource-intensive caseload. As capacity utilization, patient turnover, and PCCL offer necessary perspectives on hospital care demand, all three are vital to optimal resource allocation [28]. All three are also connected, for instance complex patients usually stay longer in hospitals increasing capacity utilization. Furthermore, with each transfer or referring another hospital additional load is created as patients need to be assessed at admission or be prepared for discharge [21]. Understanding these factors' daily variation is a vital step towards optimizing healthcare structures and processes [29]. Regarding the daily variability of care demand, analysis of long-term data can help to anticipate when staffs or supplies will be depleted or strained, thereby indicating when and where to allocate resources. Particularly, the traditional measures of capacity utilization (e.g., midnight count [29,30]) and patient turnover (e.g., the inverse of the length of hospital stay [31,32]) may not convey the dynamic nature of actual daytime hospital care demands [33]; certainly, neither incorporates patient complexity/severity.

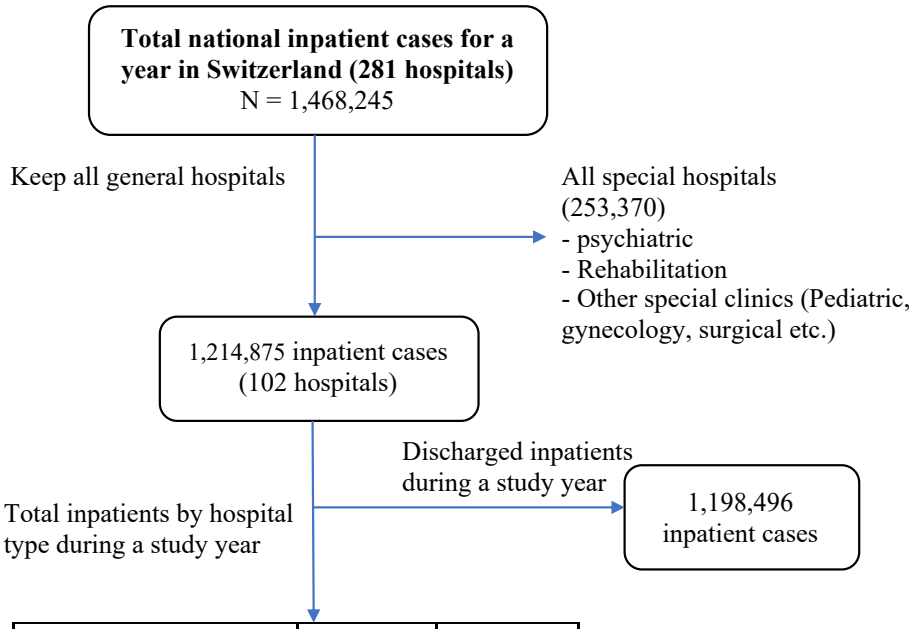
Also, the valid, highly granular longitudinal (daily/weekly) measurement of care demands offers a precise and in-depth view of how care needs fluctuate and evolve. As capacity utilization, patient turnover, and PCCL are time-sensitive variables, the three together offer great potential to accurately represent daily care demand dynamics. Such information should enable healthcare managers to anticipate capacity needs to accommodate patients during a typical weekday, weekend and seasonal peaks [3].

Therefore, this study aims to describe the daily care demand in general hospitals from a longitudinal perspective—specifically, the daily peaks and variations during weeks, i.e., weekends vs weekdays, as well as seasons. This study describes the daily variability of 1) capacity utilization; 2) patient turnover; and 3) patient clinical complexity levels of Swiss general hospitals' inpatient populations.

3.3 Methods

Study design, setting and population

This is a retrospective population-based analysis utilizing one-year hospital data extracted from the six years (2012–2017) dataset obtained from the Swiss Federal Statistics Office (FSO). Based on a data protection contract (as stipulated by article 22 of the Swiss Federal Act on Data Protection), the FSO provided anonymized data from all Swiss hospital inpatients hospitalized over the study period. The data covered general as well as specialized care facilities such as pediatric, gynecological, psychiatric and rehabilitative hospitals [34]. The FSO divides general hospitals into five classifications: university hospitals, tertiary care hospitals, large basic hospitals, medium basic hospitals and small basic hospitals. Classification is based on the number of cases treated per year and/or a weighted sum of service points (based on a combination of the number of hospital units and the levels of care delivered) assigned by the Swiss Medical Association (SMA) [17,34]. For instance, based on the SMA classification, a university hospital requires the weighted sum of service points (units) >100 and >30,000 cases per year [34]. For this study, we included a one-year patient population dataset from general hospitals to limit inter-hospital heterogeneity. Due to the Swiss Data Protection Act's stipulations regarding patient anonymity, we were unable to trace patients across years. An overview of the process of selecting inpatient cases for analysis is included in the flow diagram in Figure 1.



General hospitals (hospital types)	Number of hospitals	Inpatients
University Hospitals	5	236,649
Tertiary care hospitals	39	725,317
Large basic hospitals	15	110,638
Medium basic hospitals	27	124,579
Small basic hospitals	16	17,692

Figure 1. Flow diagram of inpatient cases for the analysis.

Dataset and study variables

We extracted one-year datasets including 1) all inpatients discharged during each study year, regardless of admission date, and 2) all inpatients admitted during each study year, regardless of discharge date. From the full FSO hospital dataset, we also extracted data for all relevant routine administrative and clinical variables for the study period (Multimedia Appendix 1).

Statistical analyses

All statistical analyses were conducted using R, version 3.6.3 for Mac OS [35]. The following statistical packages were used: 1) dplyr [36], tidyr [37] for data preparation; 2) lubridate [38] and stringr [39] for handling time and date; and 3) ggplot2 [40], patchwork [41] and scales [42] for plotting.

Descriptive overview

For each hospital type, in addition to the number of hospitals, the total numbers of patients, admissions and discharges were recorded by gender in frequencies and

percentages. Lengths of stay (LOSs) in days were calculated for all patients by subtracting each individual's admission date from their discharge date. The results were aggregated at the level of the hospital type and presented as means (with 95% confidence interval (CI)) and medians (with interquartile ranges (IQRs)) for each type.

Capacity utilization measure

For each hospital, each study day's capacity utilization was calculated as a percentage, using that hospital's highest recorded daily bed occupancy for that year as the denominator [5,9]. Each patient's admission and discharge dates were used to calculate the number of patients present each day in each hospital. The capacity utilization of a day in a given hospital included all patients admitted or discharged on that day and the number who were present the previous day and not discharged or deceased. As noted above patients admitted before the study year (e.g., in December of the previous year) were included in the study until their discharge. Patients not discharged during the study year were included until the end of the study year. Total capacity utilization was calculated for each hospital (n=102) for 365 days (with a total of 37,230 -time points).

Furthermore, daily percentages of capacity utilization were summarized by hospital type, along with a mean (95% CI), standard deviation (SD) and minimum-maximum (Min-Max). To visualize daily variations in capacity utilization, smoothed lines were plotted with the CI around the mean for each hospital type. For each hospital type, we also plotted graphs to show variation by day of the week. Finally, for each hospital type, weekday vs. weekend variations was plotted for each week of the year. There was no dot plotted for weekdays of the first week, as the first day of the study year was the weekend (Sunday).

Patient turnover measure

As done for capacity utilization, daily patient turnover was calculated for each hospital and aggregated by hospital type. The patient turnover rate was calculated as absolute counts of admissions, discharges and deceased patients for a particular day divided by the total of patients for that day per hospital [21]. As opposed to using the inverse of the average LOS method, this approach has the advantage of adequately representing the volume of activity either for entire days or short hospital stays as contributors to increasing patient throughput [19,21,43].

The percentage of patient turnover per day was calculated and further summarized by hospital type as mean (95% CI), SD and Min-Max figures. To display daily variations in patient turnover, similar displays were plotted for them as for capacity utilization.

Patient clinical complexity level (PCCL)

Our PCCL data covers only patients discharged during the study year (as the ICD-10 codes are not available for patients until they are discharged) and Swiss DRG version 6 was applied for that study year [44]. The PCCL calculation is based on a complex algorithm, providing clinical complexities and comorbidity levels (CCLs) values (0-4) for all possible primary and secondary diagnoses per patient case [26]. Developed as part of the CCL Refinement Project in Australia [26,27,45], this algorithm was applied to determine each patient's final PCCL. To facilitate this process, we used the grouping system provided by SwissDRG AG [44].

We began by organizing our data input into a readable format via the grouping system. We chose the "SwissDRG Batchgrouper Format 2017" short input format, which provides an anonymous case identifier, plus the patient's age, sex, admission and discharge date, LOS, primary and secondary diagnoses and treatment procedure codes. "DRG Output format for SwissDRG" results were then obtained, including PCCL values for each case. Individual PCCL values were further transformed to average daily PCCL value per hospital using each patient's admission and discharge dates (i.e., each case's PCCL value is applied to each day for that hospital until discharge). Daily PCCL values were further summarized by hospital type as means (95% CI), SDs and Max-Min figures. As for the other two measures, similar displays were plotted to display daily PCCL variation.

3.4 Results

Descriptive overview

During the study year, 1,214,875 inpatients stayed in the 102 Swiss general hospitals, of which 16,379 cases continued to the following year. The five university hospitals covered 20%, the 39 tertiary care hospitals covered 60% and the 58 basic hospitals covered 20% of the patient population. Overall, there were roughly 7% more female than male patients. The average patient LOS across all general hospitals was 6.43 (95%

CI 6.40–6.46) days; the median LOS was 3.7 (IQR: 2.0–7.0) days. The general characteristics of the study population by hospital type are presented in Multimedia Appendix 1.

Variation of daily capacity utilization

Average daily capacity utilization of the university hospitals ranged from 527–2,340 patients, of tertiary care hospitals from 87–1099 patients, of large basic hospitals from 16–304 patients, of medium basic hospitals from 7–179 patients, and of small basic hospitals from 1–93 patients. Notably, three small basic hospitals had average daily capacity utilization numbers below 10 patients.

Across the study period, the average daily capacity utilization was highest in university hospitals and the lowest in small basic hospitals (Table 1). However, the range, i.e., variation between the lowest- and highest-capacity utilizations for each year, was almost 98% (e.g., 1.7–100) for small basic hospitals, 92% for medium basic hospitals, 87% for large basic hospitals, 73% for tertiary care hospitals and 44% for university hospitals.

Table 1. Daily capacity utilization, patient turnover and patient clinical complexity level per hospital-by-hospital type from the one-year patient population

Hospital Type (Median[a])	Capacity utilization (%)			Patient turnover[e] (%)			Patient clinical complexity level (0-4[f])		
	Mean (95% CI[b])	SD[c]	Min–Max[d]	Mean (95% CI)	SD	Min–Max	Mean (95% CI)	SD	Min–Max
University Hospitals (988)	87.7 (87.3–88)	7.7	55.8–100	22.5 (22.1–22.8)	7.6	5.7–38.7	2.06 (2.05–2.07)	0.2	0.81–2.57
Tertiary care hospitals (298)	78.7 (78.5–78.9)	10.2	27.3–100	28.8 (28.6–28.9)	7.5	2.7–54.6	1.78 (1.78–1.79)	0.3	0.42–2.75
Large basic hospitals (120)	71.3 (70.9–71.6)	13.4	13.1–100	32.6 (32.4–32.9)	9.2	0–75.4	1.46 (1.45–1.47)	0.4	0.09–2.50
Medium basic hospitals (71)	65.3 (65–65.6)	15.2	5.9–100	34.5 (34.3–34.7)	10.7	0–109.1	1.26 (1.25–1.27)	0.6	0.00–2.93
Small basic hospitals (19)	57.8 (57.3–58.4)	22.2	1.7–100	24.3 (23.8–24.9)	22.0	0–200	1.65 (1.63–1.67)	0.8	0.00–4.00

[a]Median: Median number of beds utilized per day per hospital-by-hospital type.

[b]CI: Confidence Interval.
[c]SD: Standard deviation.
[d]Min-Max: Minimum-Maximum value on hospital level within hospital type.
[e]Patient turnover is the percentage of total patients admitted and discharged in a day.
[f]0–4: No clinical complexity, Mild clinical complexity, Severe clinical complexity, Very severe clinical complexity.

As indicated in the smooth curves and line charts by hospital categories, university hospitals' daily capacity utilization was high throughout the year. Among all hospital types, capacity utilization was lower through the summer months (June–August) (Figure 2).

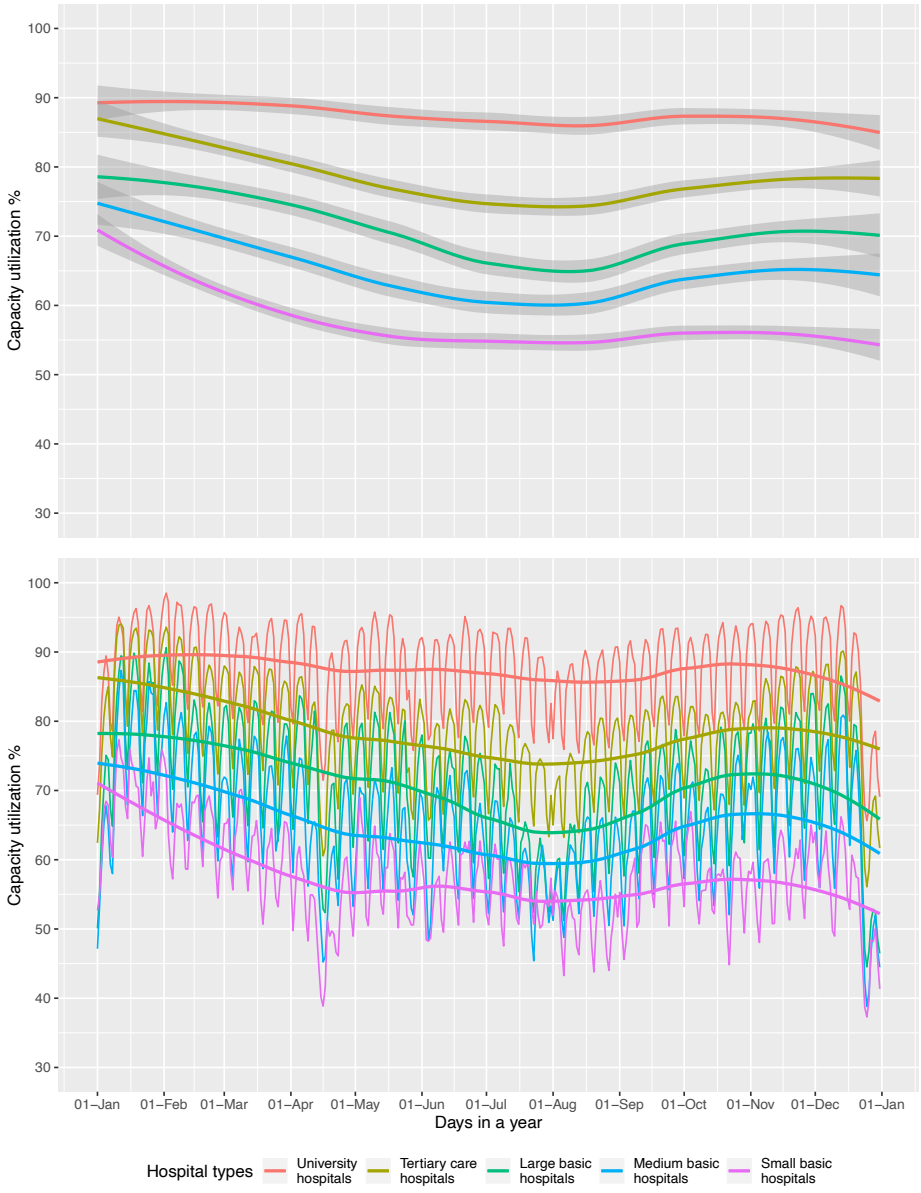


Figure 2. Daily capacity utilization of Swiss general hospital types for one year (smooth curve with mean between CIs and line chart)

There was a gradual increase in capacity utilization through the early days of the week (Mondays–Wednesdays, Figure 3) followed by the gradual decrease from Fridays to Sundays across all hospital types. There was a roughly 10% difference in capacity utilization during weekdays than at the weekend in all hospital types. (Table 2).

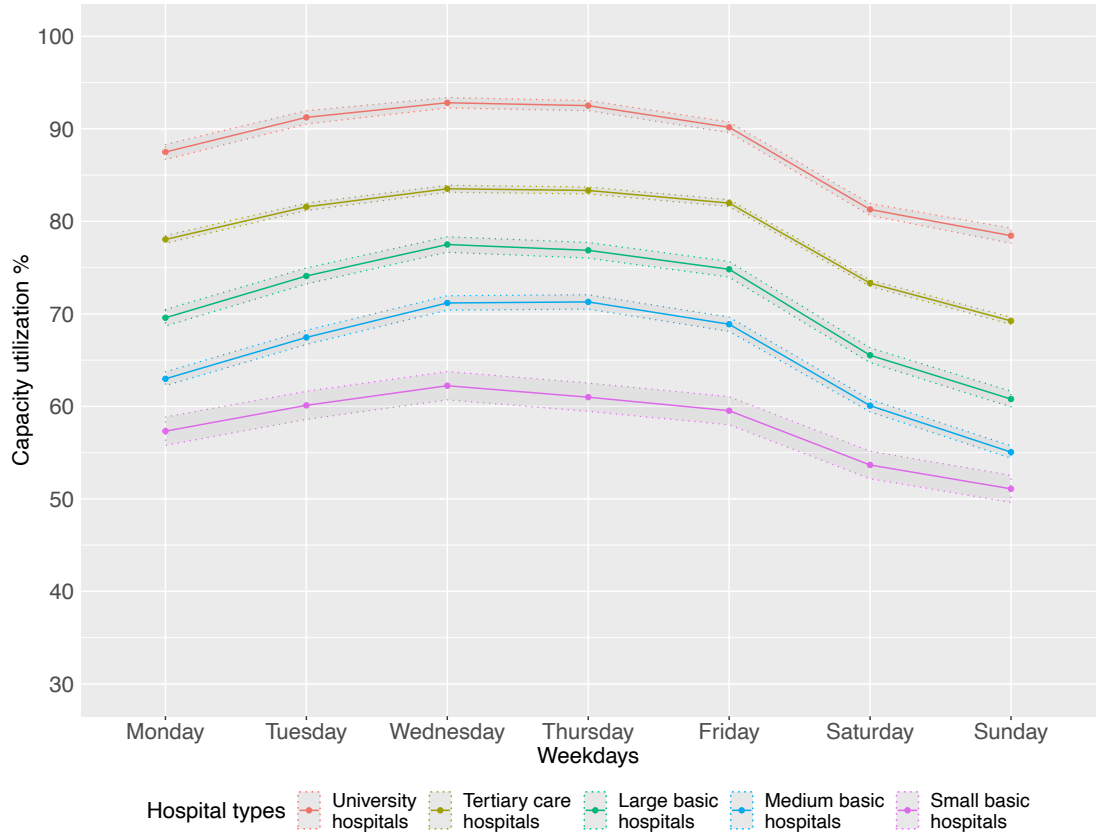


Figure 3. Capacity utilization of Swiss general hospital types with mid-CI mean by days of the week

Table 2. Daily capacity utilization, patient turnover and patient clinical complexity level by weekday vs weekend from the one-year patient population

Hospital Types	Capacity utilization (%)			Patient turnover (%)			Patient clinical complexity level (0-4[d])		
Days	Mean (95% CI[a])	SD[b]	Min–Max[c]	Mean (95% CI)	SD	Min–Max	Mean (95% CI)	SD	Min–Max
University Hospitals									
Weekday	90.8 (89.8–91.9)	3.8	72.8–96.2	25.5 (25.1–25.9)	1.4	16.6–26.6	2.04 (2.03–2.05)	0.21	1.22–2.57
Weekend	79.8 (78.9–80.6)	3.0	69.1–85.0	14.8 (14.6–15.1)	0.8	10.5–17.6	2.1 (2.08–2.13)	0.27	0.81–2.55
Tertiary care hospitals									
Weekday	81.7 (80.1–83.3)	5.6	63.5–92.1	31.9 (31.4–32.3)	1.6	22.2–33.2	1.76 (1.75–1.77)	0.28	0.84–2.69
Weekend	71.2 (69.9–72.4)	4.5	61.6–80.9	21.1 (20.8–21.4)	1.1	14.4–24.3	1.83 (1.82–1.84)	0.30	0.42–2.75
Large basic hospitals									
Weekday	74.6 (72.5–76.6)	7.4	49.6–86.8	35.9 (35.4–36.4)	1.9	25.2–37.9	1.44 (1.42–1.45)	0.38	0.14–2.25
Weekend	63 (61.4–64.6)	5.9	46.5–74.4	24.4 (23.9–24.8)	1.6	16.7–29.5	1.52 (1.5–1.54)	0.38	0.09–2.5
Medium basic hospitals									
Weekday	68.4 (66.4–70.3)	7.0	45.8–82.0	37.1 (36.8–37.5)	1.3	31.2–39.3	1.24 (1.22–1.25)	0.54	0.05–2.79
Weekend	57.6 (56.1–59.1)	5.5	44.5–70.0	27.8 (27.4–28.3)	1.7	24–34.6	1.31 (1.29–1.33)	0.57	0.00–2.93
Small basic hospitals									
Weekday	60 (58.3–61.6)	5.9	44.6–74.1	27 (26.3–27.6)	2.4	22.2–31.7	1.63 (1.6–1.65)	0.80	0.00–4.00
Weekend	52.2 (50.8–53.7)	5.3	41.4–65.7	18 (16.9–19.1)	4.0	11.3–28.8	1.69 (1.65–1.73)	0.81	0.00–4.00

[a]CI: Confidence Interval.
[b]SD: Standard deviation.
[c]Min-Max: Minimum-Maximum.
[d]0–4: No clinical complexity, Mild clinical complexity, Severe clinical complexity, Very severe clinical complexity.

Figure 4. Capacity utilization of Swiss general hospital types with mean between CI by weekdays and weekend for a year

Comparing weekdays with weekends, variations in capacity utilization for each week are shown in Figure 4. With very few exceptions (e.g., the 2nd, the 17th, and the final

week of the year), the weekly capacity utilization for weekdays was higher than for weekends. These weekly graphs also show lower capacity utilization during the summer months (weeks 20–35).

Variation of daily patient turnover

Throughout the year, daily patient turnover ranged from 72–468 patients for university hospitals, from 5–318 patients for tertiary care hospitals, from 0–91 patients for large basic hospitals, from 0–78 patients for medium basic hospitals, and from 0–33 patients for small basic hospitals. As the minimum value of zero indicates, some hospitals saw neither admissions nor discharges for some days.

During the study year, the mean daily patient turnover percentage was highest in medium-sized basic hospitals and lowest in university hospitals (Table 1). The difference in daily patient turnover range was highest for small basic hospitals, reducing with each increase in hospital size class. And as illustrated in the smoothed and line chart, daily turnover varied least in university hospitals and most in medium basic hospitals (Multimedia Appendix 2).

Exploring the frequency of change in patient movement by day of the week (Multimedia Appendix 2), we found that turnover was highest on Mondays and lowest during weekends. Differences in patient turnover during weekdays and weekends across five hospital types are shown in table 2. Across the five hospital types, daily patient turnover was almost 10% higher on weekdays than on weekends. The variation of patient turnover for each day of the week and by weekdays vs weekends is shown in Multimedia Appendix 2. In small basic hospitals, with few exceptions (e.g., the 31st and 52nd weeks), the mean daily patient turnover was higher for weekdays than for weekends in all hospital type. Overall patient turnover was lower during holiday seasons (weeks 17–18 and the final two weeks of the year).

Variation of daily patient clinical complexity level

Overall, while eleven percent of patients had severe clinical complexity (PCCL level 4), roughly 61% had none (PCCL level 0) (Table 3).

Table 3. Patient clinical complexity level by hospital types

Patient clinical complexity level (0–4) n (%)					
Hospital Type	No clinical complexity (0)	Mild clinical complexity (1)	Moderate clinical complexity (2)	Severe clinical complexity (3)	Very severe clinical complexity (4)
University Hospitals	122,567 (52.8)	3,281 (1.4)	28,812 (12.4)	41,159 (17.7)	36,308 (15.6)
Tertiary care hospitals	433,514 (60.6)	11,003 (1.5)	83,374 (11.6)	110,756 (15.5)	77,162 (10.8)
Large basic hospitals	73,800 (67.4)	1,397 (1.3)	10,801 (9.9)	14,590 (13.3)	8,984 (8.2)
Medium basic hospitals	90,284 (73.1)	1,679 (1.4)	11,237 (9.1)	12,865 (10.4)	7,465 (6.0)
Small basic hospitals	10,728 (61.5)	219 (1.3)	2,229 (12.8)	2,759 (15.8)	1,523 (8.7)
Total	730,893 (61.0)	17,579 (1.5)	136,453 (11.4)	182,129 (15.2)	131,442 (11.0)

Throughout the year, mean daily PCCL varied across the five-hospital types. It was highest in university hospitals (2.06 (95% CI 2.05–2.07)) and lowest in medium basic hospitals (1.26 (95% CI 1.25–1.27)) (Table 1). This is depicted in the smooth and line chart for five general hospital types (Multimedia Appendix 3). Mean PCCL gradually decreased from Monday until mid-week but remained highest during the weekend—the opposite of the usual pattern of capacity utilization and patient turnover explored by the days of the week (Multimedia Appendix 3).

Weekday/weekend differences in PCCL for the five general hospital types is shown in Table 2. Across all hospital types, PCCL was almost 0.07 points higher during the weekend than on weekdays. During weekdays, university hospitals' average daily PCCL was 2.04 (95% CI 2.03–2.05); during the weekend it was 2.1 (95% CI 2.08–2.13). Weekday vs weekend PCCL variation over one year is shown in Multimedia Appendix 3. Except for a small number of weeks (e.g., the 7th and the 31st weeks in small basic hospitals), across all hospital types, the PCCL for weekends was higher than for weekdays. Moreover, except for small basic hospitals, PCCL dropped in December. This was partly because PCCLs could only be calculated for patients discharged during the study year, i.e., ICD diagnostic codes were unavailable for patients not discharged during the year, and anonymity considerations made it impossible to track patients across years. However, mean values both for PCCL and

for LOS were also lower for patients discharged in November and December, with higher PCCL values assigned to patients who remained in hospital across the year-end (Multimedia Appendix 4).

3.5 Discussion

We examined one-year routine patient data from all 102 general hospitals in Switzerland. Average daily capacity utilization varied widely, from 57.8% in small basic hospitals to 87.7% in university hospitals. However, patient turnover was highest, at 34.5%, in medium basic hospitals and lowest, at 22.5%, in university hospitals. Moreover, average daily PCCL was highest in university hospitals, with 2.06, and lowest in medium basic hospitals, with 1.26. Surprisingly, in small basic hospitals, patient turnover was lower than in tertiary and either of the two other basic hospital types, both of which also had higher mean PCCLs throughout the year. Another interesting finding was that the average daily PCCL was highest on weekends. Additionally, all hospital types showed distinct weekday, weekend, and seasonal effects regarding capacity utilization, patient turnover and PCCL.

Concerns have been raised that capacity utilization alone does not explain hospitals' total care demand [6,18]. Therefore, we viewed this alongside daily volumes of admitted and discharged patients and their complexity [18]. This study explored all three measures, showing that capacity utilization and PCCL were highest but patient turnover lowest in university hospitals. Even with a large proportion of inpatients in tertiary care institutions, university hospitals generally operate at close to full capacity and with the most complex patient cases. Thus, more care resources need to be allocated to university hospitals [3]. On the other hand, in small basic hospitals, where capacity utilization and patient turnover were relatively low, PCCL was above those of the other basic hospitals. This indicates that complex cases are still treated in small basic hospitals, possibly due to geographic proximity, which may also relate to older patients' preference for them: across all hospital types, these hospitals have the highest mean patient age. In light of these small basic hospitals' continued relevance (as they still treat complex cases), they may warrant greater resource allocation.

We measured daily demand for Swiss general hospital care longitudinally for one year. This has also been observed in other studies regarding days of the week and for

Saturdays and Sundays, where both capacity utilization and patient turnover were lowest [3,46]. Moreover, we observed that PCCL was highest during weekends, possibly because more complex patients remain in the hospital through the weekend. Comparing weekly demand throughout the year, a clear distinction between weekends and weekdays was shown, with the highest variability occurring in small basic hospitals, possibly indicating suboptimal patient flow. Concurrently, seasonal variations were also seen. Capacity utilization was mainly highest in winter and relatively low during summer months, whereas patient turnover was constant throughout the year dropping towards the end of the year. However, PCCL remained quite constant, with a slight drop during the summer months and a marked reduction during December. These changes tended to correspond with holidays, possibly including higher patient discharge rates and fewer admissions before the holidays and at the end of the year. The capture of daily patient complexity during the end of the year was also reduced because ICD-10 codes were unavailable for patients who were admitted but not discharged during the study year. Furthermore, patients with higher PCCLs were more likely to have longer LOSs, particularly across the Christmas/year-end period.

Potential implications

Based on capacity utilization, patient turnover and PCCL, the variability of daily care demand in general hospitals directly impacts resource utilization. From the perspective of a single hospital, the extent of that impact depends on the degree of variation in care need, as well as on the hospital administrators' ability to adapt or otherwise respond to changes either in resource supply or demand. Our analysis on the demand dynamics of the Swiss healthcare system indicate that monitoring of care demand is useful to create surge capacity during disasters or COVID-19 [20] pandemic by alternative solutions like, smoothing workloads and coordinated early discharge. It also has the potential to help health system planners and hospital managers tailor their staffing and other resources to match care demand and the early planning of the admissions (e.g., surgeries or follow-up treatment) [47] to control patient's flow.

What the analysis cannot describe is the human and other resources needed to meet care demand i.e., balancing it or any other supply-demand chain. To do so would require a full exploration of the relevant human resources (e.g., physician and nurse

staffing) in the light of each hospital's care demand. Furthermore, application of time-driven activity-based costing (TDABC) methods could provide a framework to identify process improvements for healthcare delivery [48,49]. However, we were not able to consider TDABC as we don't have sufficient data regarding resources consumption (personnel, equipment and supplies) during the patients' journey along the clinical pathway [48].

Some studies have linked high capacity utilization and patient turnover with adverse patient outcomes [5,21]. In addition to these results, acknowledging the effect of clinical complexity alongside capacity utilization and patient turnover might bring us closer to understanding the factors that stress hospital systems, and the effects a stressed system has on patient outcomes such as in-hospital mortality. Describing daily care demand to identify meaningful variation will require further studies, e.g., examining patients' time-varying exposure to hospitals or units under pressure and the impact on the quality-of-care indicators/patient outcomes in causal models. Particularly, extending previous research on capacity utilization and in-hospital mortality [5,50], using daily capacity utilization as time-varying exposure (systemic stress factor); daily patient turnover and PCCL as time-varying confounders would be of interest. In a practical sense, this might also allow monitoring of pressure zones (e.g., to manage care demand where possible) in hospitals, which could reduce avoidable adverse events or death [5].

Strength and limitations

To our knowledge, this is the first study to explore hospital care demand dynamics via daily measurements of capacity utilization, patient turnover, and PCCL on a national health system level. Furthermore, applying the standard methodology, programming, and software for large datasets allowed computing and visualization of demand dynamics per day, day of the week, and weekdays versus weekends to gain a longitudinal perspective.

This study also has notable limitations. While we explored demand dynamics in detail, we could not do so with supply dynamics (e.g., staffing, resources)—an entire category of critical information in the demand-supply equation. Due to the large sample and FSO data composition (aggregated data), it was also not possible to explore demand dynamics at the unit level–the interface between patients' care

demand and health professionals' provision of care. Also, as we used codes assigned in routine data, the patient diagnoses and other variables could be biased by factors, such as the accuracy of physicians' and nurses' documentation, ICD-10 codes not available for patients who were not discharged, and intentional upcoding of diagnoses to more expensive Swiss DRG categories [51,52].

3.6 Conclusions

This study illustrates daily care demand based on capacity utilization, patient turnover and PCCL and its variability between the five classes of Swiss general hospitals. For all five types, our analyses indicated distinct differences in capacity utilization, patient turnover and PCCL between days of the week, weekdays versus weekends and seasons. This longitudinal study is a step towards detecting possible variables to be considered for time-varying exposure (e.g., capacity utilization) and confounders (e.g., PCCL) in developing a casual model of tipping points and their links with quality of care or patient outcomes. Essentially, the variability of care demand provides a new perspective for gauging when hospitals are under strain and could help avoid pressure zones with a combination of appropriate resource allocation and care demand planning in general hospitals.

3.7 References

1. Baker A. Book: Crossing the Quality Chasm: A New Health System for the 21st Century. BMJ: British Medical Journal. 2001;323(7322):1192. [PMID: 25057539]
2. Weber TA. Optimal control theory with applications in economics. London: The MIT Press; 2011.
3. Belciug S, Gorunescu F. Improving hospital bed occupancy and resource utilization through queuing modeling and evolutionary computation. Journal of biomedical informatics. 2015;53:261-269. [PMID: 25433363]
4. Griffiths P, Saville C, Ball J, Jones J, Pattison N, Monks T, et al. Nursing workload, nurse staffing methodologies and tools: A systematic scoping review and discussion. International journal of nursing studies. 2020;103:103487. [PMID: 31884330]
5. Kuntz L, Mennicken R, Scholtes S. Stress on the ward: Evidence of safety tipping points in hospitals. Management Science. 2014;61(4):754-771. [PMID:
6. Freedman S. Capacity and utilization in health care: the effect of empty beds on neonatal intensive care admission. American Economic Journal: Economic Policy. 2016;8(2):154-185. [PMID: 27942353]
7. Sloan FA, Hsieh C-R. Health economics. London: MIT Press; 2017.
8. Jones RP. Condition-specific growth in occupied beds in England following a sudden and unexpected increase in deaths. British Journal of Healthcare Management. 2019;25(6):1-8. [PMID:
9. Friebel R, Fisher R, Deeny SR, Gardner T, Molloy A, Steventon A. The implications of high bed occupancy rates on readmission rates in England: A longitudinal study. Health Policy. 2019. [PMID: 31262535]
10. Reader TW, Gillespie A. Patient neglect in healthcare institutions: a systematic review and conceptual model. BMC Health Services Research. 2013;13(1):156. [PMID: 23631468]
11. Ball JE, Murrells T, Rafferty AM, Morrow E, Griffiths P. 'Care left undone'during nursing shifts: associations with workload and perceived quality of care. BMJ Qual Saf. 2014;23(2):116-125. [PMID: 23898215]
12. Wagner J, Gabler NB, Ratcliffe SJ, Brown SE, Strom BL, Halpern SD. Outcomes among patients discharged from busy intensive care units. Annals of internal medicine. 2013;159(7):447-455. [PMID: 24081285]
13. Panagioti M, Khan K, Keers RN, Abuzour A, Phipps D, Kontopantelis E, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. bmj. 2019;366:l4185. [PMID: 31315828]
14. OECD. Health at a Glance 2017. OECD Publishing. OECD Indicators Web site. https://www.health.gov.il/publicationsfiles/healthataglance2017.pdf. Published 2017. Accessed 2020-07-08.
15. Fisher ES, Wennberg JE, Stukel TA, Skinner JS, Sharp SM, Freeman JL, et al. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. Health services research. 2000;34(6):1351. [PMID: 10654835]
16. WHO European Health Information Gateway. Bed occupancy rate (%), acute care hospitals only. https://gateway.euro.who.int/en/indicators/hfa_542-6210-bed-occupancy-rate-acute-care-hospitals-only/. Published 2018. Accessed 2020-05-08.
17. Federal Statistical Office. Health Pocket Statistics 2017. Federal Statistical Office (FSO). https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.4342090.html. Published 2018. Accessed 2020-10-02.

18. Bujold M, Pluye P, Légaré F, Haggerty J, Gore GC, El Sherif R, et al. Decisional needs assessment of patients with complex care needs in primary care: a participatory systematic mixed studies review protocol. BMJ open. 2017;7(11):e016400. [PMID: 29133314]
19. VanFosson CA, Yoder LH, Jones TL. Patient turnover: A concept analysis. Advances in Nursing Science. 2017;40(3):300-312. [PMID: 28266962]
20. Chopra V, Toner E, Waldhorn R, Washer L. How should US hospitals prepare for coronavirus disease 2019 (COVID-19)? Ann Intern Med. 2020;172(9):612-622. [PMID: 32160273]
21. Park SH, Blegen MA, Spetz J, Chapman SA, De Groot H. Patient turnover and the relationship between nurse staffing and patient outcomes. Research in nursing & health. 2012;35(3):277-288. [PMID: 22457013]
22. Mendez CM, Harrington DW, Christenson P, Spellberg B. Impact of hospital variables on case mix index as a marker of disease severity. Population health management. 2014;17(1):28-34. [PMID: 23965045]
23. Menendez ME, Neuhaus V, Van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical Orthopaedics and Related Research®. 2014;472(9):2878-2886. [PMID: 24867450]
24. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC Health Serv Res. 2021;21(13). [PMID:
25. Quentin W, Geissler A, Busse R. Measuring and comparing health system outputs: using patient classification systems for efficiency analyses. In: Health system efficiency: How to make measurement matter for policy and management [Internet]. European Observatory on Health Systems and Policies; 2016.
26. SwissDRG. Swiss Diagnosis Related Groups Version 8.0. In. Bern, Switzerland: SwissDRG AG; 2018.
27. Bauer M, Ostermann H. DRGs in transfusion medicine and hemotherapy in Germany. Transfusion medicine and hemotherapy. 2012;39(2):60-66. [PMID: 22670123]
28. Salisbury C, Boerma W. Balancing demand and supply in 24 Hour Primary Care. 2018.
29. Fieldston E, Ragavan M, Jayaraman B, Metlay J, Pati S. Traditional measures of hospital utilization may not accurately reflect dynamic patient demand: findings from a children's hospital. Hospital pediatrics. 2012;2(1):10-18. [PMID: 24319808]
30. Macy ML, Stanley RM, Lozon MM, Sasson C, Gebremariam A, Davis MM. Trends in high-turnover stays among children hospitalized in the United States, 1993–2003. Pediatrics. 2009;123(3):996-1002. [PMID: 19255031]
31. Unruh LY, Fottler MD. Patient turnover and nursing staff adequacy. Health services research. 2006;41(2):599-612. [PMID: 16584467]
32. Norrish BR, Rundall TG. Hospital restructuring and the work of registered nurses. The Milbank Quarterly. 2001;79(1):55-79. [PMID: 11286096]
33. DeLia D. Annual bed statistics give a misleading picture of hospital surge capacity. Annals of emergency medicine. 2006;48(4):384-388. e382. [PMID: 16997673]
34. De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288. [PMID: 26766626]
35. R: A Language and Environment for Statistical Computing [computer program]. R Foundation for Statistical Computing; 2019.

36. Wickham H, Francois R. dplyr: A Grammar of Data Manipulation. https://cran.r-project.org/web/packages/dplyr/index.html. Published 2020. Accessed 2020-08-28.
37. Wickham H. tidyr: Tidy Messy Data. https://cran.r-project.org/web/packages/tidyr/index.html. Published 2020. Accessed 2020-08-29.
38. Mailund T. Working with Dates: lubridate. In: R Data Science Quick Reference. Springer; 2019:195-203.
39. Wickham H. stringr: Simple, Consistent Wrappers for Common String Operations. https://cran.rstudio.com/web/packages/stringr/index.html. Published 2019. Accessed 2020-02-10.
40. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2016.
41. Pedersen T. patchwork: The Composer of Plots. https://cran.r-project.org/web/packages/patchwork/index.html. Published 2020. Accessed 2020-12-18.
42. Wickham H. scales: Scale Functions for Visualization. https://cran.r-project.org/web/packages/scales/index.html. Published 2020. Accessed 2020-05-25.
43. Park SH, Dunton N, Blegen MA. Comparison of Unit-Level Patient Turnover Measures in Acute Care Hospital Settings. Research in nursing & health. 2016;39(3):197-203. [PMID: 26998744]
44. SwissDRG. SwissDRG System 6.0/2017. SwissDRG AG. https://www.swissdrg.org/de/akutsomatik/archiv-swissdrg-system/swissdrg-system-602017. Published 2017. Accessed 2020-05-04.
45. Antioch K, Zhang X. Organ transplant AN-DRGs: modifying the exceptions hierarchy in casemix classification. Australian Health Review. 2000;23(1):137-152. [PMID: 10947598]
46. Pauls LA, Johnson-Paben R, McGready J, Murphy J, Pronovost PJ, Wu CL. The weekend effect in hospitalized patients: a meta-analysis. Journal of hospital medicine. 2017;12(9):760-766. [PMID: 28914284]
47. Shabanikiya H, Gorgi HA, Seyedin H, Jafari M. Assessment of hospital management and surge capacity in disasters. Trauma monthly. 2016;21(2). [PMID: 27626015]
48. Martin JA, Mayhew CR, Morris AJ, Bader AM, Tsai MH, Urman RD. Using time-driven activity-based costing as a key component of the value platform: a pilot analysis of colonoscopy, aortic valve replacement and carpal tunnel release procedures. Journal of clinical medicine research. 2018;10(4):314. [PMID: 29511420]
49. Keel G, Savage C, Rafiq M, Mazzocato P. Time-driven activity-based costing in health care: a systematic review of the literature. Health Policy. 2017;121(7):755-763. [PMID: 28535996]
50. Madsen F, Ladelund S, Linneberg A. High levels of bed occupancy associated with increased inpatient and thirty-day hospital mortality in Denmark. Health Affairs. 2014;33(7):1236-1244. [PMID: 25006151]
51. Quan H, Eastwood C, Cunningham CT, Liu M, Flemons W, De Coster C, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ open. 2013;3(10):e003716. [PMID: 24114372]
52. Wasserfallen J-B, Zufferey J. Financial impact of introducing the Swiss-DRG reimbursement system on potentially avoidable readmissions at a university hospital. Swiss medical weekly. 2015;145:w14097. [PMID: 25659119]

Chapter 4

Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data

Narayan Sharma[1], René Schwendimann[1,2], Olga Endrich[3], Dietmar Ausserhofer[1,4], Michael Simon[1,5]

[1]Institute of Nursing Science (INS), Department Public Health (DPH), Faculty of Medicine, University of Basel, Basel, Switzerland
[2]Patient Safety Office, University Hospital Basel, Basel, Switzerland
[3]Directorate of Medicine, Inselspital University Hospital Bern, Bern, Switzerland
[4]College of Health-Care Professions Claudiana, Bozen, Italy
[5]Nursing Research Unit, Inselspital University Hospital Bern, Bern, Switzerland

Published in the BMC Health Service Research: https://doi.org/10.1186/s12913-020-05999-5

4.1 Abstract

Background: Understanding how comorbidity measures contribute to patient mortality is essential both to describe patient health status and to adjust for risks and potential confounding. The Charlson and Elixhauser comorbidity indices are well-established for risk adjustment and mortality prediction. Still, a different set of comorbidity weights might improve the prediction of in-hospital mortality. The present study, therefore, aimed to derive a set of new Swiss Elixhauser comorbidity weightings, to validate and compare them against those of the Charlson and Elixhauser-based van Walraven weights in an adult in-patient population-based cohort of general hospitals.

Methods: Retrospective analysis was conducted with routine data of 102 Swiss general hospitals (2012–2017) for 6.09 million inpatient cases. To derive the Swiss weightings for the Elixhauser comorbidity index, we randomly halved the inpatient data and validated the results of part 1 alongside the established weighting systems in part 2, to predict in-hospital mortality. Charlson and van Walraven weights were applied to Charlson and Elixhauser comorbidity indices. Derivation and validation of weightings were conducted with generalized additive models adjusted for age, gender and hospital types.

Results: Overall, the Elixhauser indices, c-statistic with Swiss weights (0.867, 95% CI: 0.865–0.868) and van Walraven's weights (0.863, 95% CI: 0.862-0.864) had substantial advantage over Charlson's weights (0.850, 95% CI: 0.849–0.851) and in the derivation and validation groups. The net reclassification improvement of new Swiss weights improved the predictive performance by 1.6% on the Elixhauser-van Walraven and 4.9% on the Charlson weights.

Conclusions: All weightings confirmed previous results with the national dataset. The new Swiss weightings model improved slightly the prediction of in-hospital mortality in Swiss hospitals. The newly derive weights support patient population-based analysis of in-hospital mortality and seek country or specific cohort-based weightings.

Keywords: comorbidity indices; weights; risk adjustment; inpatient data; in-hospital mortality

4.2 Background

Critical health outcomes such as mortality often require effective risk adjustment based on patient characteristics to predict in-hospital mortality. This is also true for comorbidities [1,2], which function as major predictors of mortality [3]. Over one-third of hospitalized patients have at least one comorbidity; two-thirds of those over 65 [2,4] and three-quarters of those over 85 have at least two [5]. In addition to mortality, comorbidities are associated with lower health-related quality of life, increased disability and higher utilization of both health care services and prescribed medications [6-8].

Data on comorbidities are valuable both for comparison between patient populations and for risk adjustment regarding associated outcomes, especially mortality [9]. Two of the best-known measures are the Charlson Comorbidity Index and the Elixhauser Comorbidity Index [10,11]. When the Charlson Comorbidity Index was developed in 1987 it included 19 chronic conditions to predict one-year mortality, but has since been shortened to 17. The Elixhauser Comorbidity Index, which was developed in 1998, works on a similar system but includes 30 – or, for some variants, 31 – comorbidities. In addition to in-hospital mortality, it is also used to predict the length of stay, adverse events and hospital discharges [12,13]. Despite this additional versatility (covering acute and chronic conditions) and strong evidence that the Elixhauser Comorbidity Index is statistically superior to the Charlson Comorbidity Index [13,14], the Charlson Comorbidity Index continues to be used. Because of the fewer chronic conditions [15-17] and comparative ease of use in routine situations where time is limited.

Both indices work either via simple (unweighted) sum scores or as weighted scores assigning a risk weight to each comorbidity [6,18,19]. A weighted sum score/summary measure provides an attractive advantage over plain dummy variables [20,21], as it reduces the overfitting risk of more parameters, unjustifiable in small datasets [22] and limits computational requirements in large ones [21]. Additionally, evidence indicates that a weighted variable reduces type I errors compared to dummy variables while addressing multicollinearity concerns in regression analysis and organizing multiple highly correlated variables into more meaningful information [21,23]. The weight assigned to each comorbidity reflects a higher, lower or neutral risk of mortality [24]. Practically, mortality risk scores can

help to identify high-risk cases for special management and to assess provider services whose patients perform better or worse than expected from the summary measure of the morbidity burden.

To add to the value of early versions of the Elixhauser comorbidities, van Walraven et al. [25] used roughly 13 years' inpatient admission data from one Canadian hospital (1996–2008) to develop a set of weights (VW weights, i.e., the regression coefficient divided by the coefficient in the model with the smallest absolute value) for the 30 Elixhauser comorbidities associated with in-hospital mortality. Using the backward selection and an alpha inclusion criterion of 0.05 to identify independently associated comorbidities, van Walraven identified 21 comorbidities significantly associated with mortality. A VW weight was assigned to each of the 21 Elixhauser comorbidities. Ultimately, VW weights ranged from -7 to 12, with a weight of 0 assigned to the 9 non-significant comorbidities.

Since then, primarily in North America, studies have used VW weights to predict in-hospital mortality, especially in clearly defined patient groups such as surgical, orthopedic, or cancer patients and those in single hospitals or intensive care unit (ICU) [12,13,21,25,26]. Moreover, the comorbidity weighting system might differ between all hospitalization and a restricted cohort; mortality and other outcomes; and between the countries [27]. Few studies have applied comorbidity adjustments to national or regional inpatient datasets [21,28]. Therefore, an analysis of a large heterogeneous patient population from a national dataset (Switzerland) is justified both to provide an overview of Elixhauser comorbidities in a European sample and potentially to optimize the comorbidity weights. In addition to increasing the generalizability of these comorbidity weights, the use of such a dataset, representing all hospital inpatient cases (i.e., hospitalization episodes) from a large, heterogeneous patient population, would allow a very accurate comparison of weighting systems. Therefore, the aims of our study were 1) to derive a new Swiss comorbidity weighting on a national inpatient dataset to predict in-hospital mortality; 2) to validate Charlson, Elixhauser-van Walraven and new weights on a national inpatient dataset; and 3) to compare the predictive performance of in-hospital mortality of the three weighting systems.

4.3 Methods

Study design and population

This is a retrospective population-based analysis of six years' data (2012–2017) from the Swiss national inpatient dataset. Upon our application, subject to a data protection contract (as stipulated by article 22 of the Swiss Federal Act on Data Protection), the Swiss Federal Statistics Office (FSO) provided anonymized data from all Swiss hospital inpatient cases hospitalized between 2012 and 2017. This included not only general hospitals but also special care (e.g., pediatric, gynecological) facilities [29]. The FSO classifies general hospitals (University hospitals, Tertiary hospitals, and three Basic hospitals) into five different levels, based on the number of cases treated per year and/or a special hospital score assigned by Swiss Medical Association ("FMH-Kategorien"). For this study, special care hospitals and children were excluded because of the low levels of comorbidities and the relatively low risk of dying in the hospital [25]. For data protection reasons, age was grouped in five-year groups, and all inpatient cases below 20 years of age were excluded. The flowchart for the final adult population included 102 general hospitals (6,094,672 inpatient cases) for the analysis is reported in supplementary figure F1 (Additional file 1).

Dataset and classification of comorbidities indices

The dataset included patient characteristics including sex, age, hospital types, primary and secondary diagnoses based on International Classification of Diseases-10 (ICD-10) codes and hospital discharge information including in-hospital mortality. As condition coding in Switzerland is based on the ICD-10 German Modification (ICD-10 GM), reported in supplementary table S1 (Additional file 1), we used this to identify both Charlson and Elixhauser comorbidities. Specifically, we used Quan et al.'s ICD-10 codes [19] to determine each of the 17 Charlson and 31 Elixhauser comorbidities via the "Comorbidity" package in R [18]. This transforms ICD-10 codes into binary data the relevant comorbidities, their (unweighted) sum scores, and their Charlson and VW-weighted scores.

Descriptive analysis

The study population's general characteristics (hospital types, patient's sex, and age groups) were reported in the alive and mortality cohorts with percentages. The distributions of Charlson and Elixhauser comorbidities, unweighted and weighted scores were computed as percentages of index values of 0, 1–2, and ≥3 and <0, 0, 1–4, and ≥5; and as the Charlson weight do not use negative weightings, its weights were calculated for index values of 0, 1–4, and ≥5. For each characteristic and comorbidities, standardized mean differences (SMD) between the alive and mortality cohort were computed using "tableone" package in R. The SMD is identical to Cohen's D and provides an effect size estimate less sensitive to the sample size than p values between the cohorts. This is important in a dataset of the size used in this study. An SMD of zero means there is no difference in the characteristics (e.g., gender) between the alive and mortality cohort. SMDs greater than 0.1 indicate potentially relevant differences [30], i.e., showing unbalanced covariates and might have an association with mortality.

Derivation of Swiss comorbidity weights

The study population was randomly split into a derivation (50%) and a validation (50%) group. The derivation group was used to determine the adjusted association of all 31 Elixhauser comorbidities with death, treating the anonymous hospital identifier as a random effect [31]. Generalized additive regression models (GAM) can accommodate many predictors including random effect, able to handle large dataset easily and nonparametric in nature [32]. We fitted GAM to compute the odds ratios (OR) using the package "mgcv" [33] and R programming language, version 3.5.2 [34]. We utilized GAM with random effect components on the hospital level, as university and small hospitals are different in size and services in Switzerland. To identify Elixhauser comorbidities associated with in-hospital mortality, we retained variables based on an alpha inclusion criterion of 0.01.

To derive the Swiss weightings from the regression model's parameter estimates, we used the method described by Sullivan et al. [35]. Comorbidities not significantly associated with mortality were assigned a weight of zero. The number of (weighted) points assigned to each comorbidity equaled its regression coefficient divided by the coefficient in the model with the smallest absolute value [14,21,25,35] rounded to the

nearest whole number. Each person's new Elixhauser comorbidity weighting score was then calculated by summing up all points of all their coded comorbidities.

Validation and comparison of weighted comorbidity models

To validate and compare the performance of the three comorbidity weighting systems, we first created four multivariate in-hospital mortality prediction GAMs for the derivation group. The first model, 'base', contained no comorbidity data – only age group, sex, and hospital types. The other three models used the same variables as the base model, with the first, 'Charlson', using Charlson weights, the second, 'van Walraven', using the Elixhauser index with van Walraven weights, and the third, 'Swiss weights', using our newly-developed weights. We then validated all weights in validation groups by splitting the validation group into six groups by year of discharge. Altogether, 24 c statistics (including base models) were computed to validate the Charlson, van Walraven, and Swiss weights models in the validation sample. An additional four models were created using all cases (combining derivation and validation groups) to evaluate the performance of each model in the total patient population.

We assessed the various comorbidity weightings according to the model performance criteria. Discrimination, i.e., each model's ability to distinguish patients discharged alive from those who died in hospital, was compared using the concordance (c) statistic. The c-statistic quantified each model's ability to assign high probabilities of mortality to patients who died [36]. It's possible values range from 0.50 to 1.0, with 0.50 indicating no ability to discriminate, values less than 0.70 are considered poor, those between 0.70 and 0.80 acceptable, and those of 0.80 or above excellent [37]. Using bootstrap methods, we computed 95% confidence intervals for each c-statistic. Additionally, the observed value was also explored for each model from the predicted values to observe the model performances in the highest selected percentages (1%, 2%, 5% and 10%) in the derivation sample. We also graphed receiver-operating characteristic curves (ROC) for the visual presentation of the derivation group's c statistics. We compared the base model and existing comorbidity models with Swiss weight model using net reclassification improvement (NRI) for binary outcome [38-40] from the "nricens" package in R [41] using the Swiss derivation sample and

classification cut-off value of 0.023 (mortality proportion of the total study population). NRI measures the degree of improvement in predicted inpatient mortality probabilities when comorbidity weights are added to the base model [21,42]. Higher NRI values indicate more accurate reclassification.

Code validation and sensitivity analyses

We also evaluated the R comorbidity package's code handling accuracy in the Swiss setting. To do so we sampled 100 cases and manually reviewed the Swiss ICD-10 codes of the raw data, checking whether the "comorbidity" package had assigned each to the appropriate Charlson or Elixhauser comorbidity. We also performed sensitivity analyses to explore Switzerland's Major Diagnostic Categories' (MDCs') associations, which are based on ICD-10 GM (one way of expressing the reason for admission), if any, regarding the change in the predictability of in-hospital mortality in combination with the above models and to test whether the combined models' patterns differed from those of uncombined ones. MDCs are 24 mutually exclusive categories into which all primary diagnoses are assigned based on the Swiss diagnostic-related group (DRG) system for hospital reimbursement [43].

4.4 Results

Population characteristics

Overall, the adult inpatient population between 2012 and 2017 in all Swiss general hospitals (102) consisted of 6,094,672 cases. Among all hospitalized cases in our study population mortality was 2.3%. The characteristics of the adult inpatient cases are presented in Table 1. Inpatient cases had between 0 and 9 Charlson comorbidities (median 0, interquartile range (IQR): 0-1) and between 0 and 16 Elixhauser comorbidities (median 1, IQR: 0-2). The different categories of three comorbidity weightings are presented in supplementary table S2 (Additional file 1).

Table 1 General characteristics of the total study population

Parameters	Alive cohort (%)	Mortality cohort (%)	SMD
[a]Total population: N = 6,094,672	5,952,005 (97.7)	142,667 (2.3)	
Females	3,280,823 (55.1)	63,912 (44.8)	0.208
Age groups			1.006
20–24 years	215,672 (3.6)	292 (0.2)	
25–29 years	327,562 (5.5)	375 (0.3)	
30–34 years	415,022 (7.0)	526 (0.4)	
35–39 years	348,591 (5.9)	718 (0.5)	
40–44 years	299,985 (5.0)	1,368 (1.0)	
45–49 years	350,899 (5.9)	2,503 (1.8)	
50–54 years	408,028 (6.9)	4,312 (3.0)	
55–59 years	430,721 (7.2)	6,503 (4.6)	
60–64 years	466,543 (7.8)	9,068 (6.4)	
65–69 years	528,374 (8.9)	13,322 (9.3)	
70–74 years	554,612 (9.3)	16,899 (11.8)	
75–79 years	535,543 (9.0)	19,888 (13.9)	
80–84 years	509,225 (8.6)	24,853 (17.4)	
85–89 years	365,924 (6.1)	24,042 (16.9)	
90–94 years	161,236 (2.7)	14,156 (9.9)	
95+ years	34,068 (0.6)	3,842 (2.7)	
Hospital types			0.157
University (level 1)	1,078,612 (18.1)	29,379 (20.6)	
Tertiary care (level 2)	3,274,382 (55.0)	83,686 (58.7)	
Basic care (level 3)	736,465 (12.4)	14,863 (10.4)	
Basic care (level 4)	671,182 (11.3)	10,695 (7.5)	
Basic care (level 5)	191,364 (3.2)	4,044 (2.8)	
Number of Charlson comorbidities			1.234
0	3,642,650 (61.2)	17,465 (12.2)	
1–2	1,907,761 (32.1)	80,876 (56.7)	
>= 3	401,594 (6.7)	44,326 (31.1)	
Number of Elixhauser comorbidities			1.039
0	2,509,169 (42.2)	11,036 (7.7)	
1–2	2,106,780 (35.4)	43,494 (30.5)	
>= 3	1,336,056 (22.4)	88,137 (61.8)	

Abbreviations: SMD standardized mean difference between alive and mortality cohort

[a]Total population presented in row percentage

Prevalence of Charlson and Elixhauser comorbidity indices

The most common Charlson comorbidity was any malignancy (including lymphoma and leukemia, except malignant neoplasm of the skin) in both cohorts, alive (10.2 %) and mortality (37.6%), yet with marked differences between the two cohorts (SMD:

0.680). The prevalence for each Charlson comorbidity in the total population and the derivation is presented in supplementary table S3 (Additional file 1).

The most common Elixhauser comorbidities were uncomplicated hypertension (22.7 %) in the alive cohort, whereas in the mortality cohort, it was solid tumour without metastasis (33.7%). However, the most pronounced difference between both cohorts was observed for metastatic cancer (4.0% vs. 26.5%; SMD: 0.657). The prevalence for each Elixhauser comorbidity from the total population and derivation group is presented in the supplementary table S4 (Additional file 1).

Derivation of Swiss weights

In the derivation group, two of the 31 Elixhauser comorbidities showed no association with hospital mortality and were removed, leaving 29 in the final model with random effect on the hospital level. Sixteen were associated with increased mortality risk, with the strongest associations coming from metastatic cancer (OR: 4.09, 95% CI: 3.98–4.21) and liver disease (OR: 3.83, 95% CI: 3.70–3.97). At the other end of the spectrum, thirteen comorbidities were associated with a decreased risk of hospital mortality. The strongest of these were deficiency anaemia (OR: 0.54, 95% CI: 0.51–0.56) and obesity (OR: 0.59, 95% CI: 0.56–0.63). The adjusted coefficients were used to derive Swiss weights with a new maximum weight of 17, for metastatic cancer, and a new minimum of -7, for deficiency anaemia (Table 2).

Table 2 Prevalence, adjusted odds ratio and weights from the (new) Swiss derivation sample and the van Walraven (VW) derivation sample [25]

Elixhauser comorbidities	Alive cohort (%)	Mortality cohort (%)	SMD	Adjusted odds ratio (95% CI)		Weights	
	Swiss derivation sample			VW[a]	Swiss	VW	Swis
[b]Derivation	2,975,887 (97.7)	71,449 (2.3)					
Congestive heart failure	163,685 (5.5)	16,333 (22.9)	0.514	1.96 (1.85–2.07)	3.07 (3.00–3.14)	7	13
Cardiac	341,280 (11.5)	20,754 (29.0)	0.448	1.71 (1.62–1.80)	1.69 (1.66–1.73)	5	6
Valvular disease	117,450 (3.9)	6,568 (9.2)	0.213	0.91 (0.82–0.99)	0.92 (0.89–0.95)	-1	-1
Pulmonary circulation	53,292 (1.8)	4,813 (6.7)	0.247	1.48 (1.34–1.62)	1.62 (1.57–1.68)	4	6
Peripheral vascular	141,051 (4.7)	6,912 (9.7)	0.192	1.26 (1.17–1.36)	1.27 (1.24–1.31)	2	3
Hypertension (uncomplicated)	676,609 (22.7)	15,692 (22.0)	0.019	–	0.69 (0.68–0.70)	0	-4
Hypertension (complicated)	218,656 (7.3)	11,003 (15.4)	0.256	–	0.79 (0.77–0.81)	0	-3
Paralysis	61,546 (2.1)	5,153 (7.2)	0.246	1.93 (1.75–2.12)	2.60 (2.52–2.69)	7	11
Other neurological	120,045 (4.0)	8,011 (11.2)	0.273	1.83 (1.70–1.96)	2.45 (2.39–2.52)	6	10
Chronic pulmonary	170,770 (5.7)	8,269 (11.6)	0.209	1.36 (1.29–1.44)	1.31 (1.27–1.34)	3	3
Diabetes, uncomplicated	245,817 (8.3)	9,059 (12.7)	0.145	–	1.09 (1.06–1.11)	0	1
Diabetes, complicated	66,161 (2.2)	2,763 (3.9)	0.096	–	0.89 (0.86–0.93)	0	-1
Hypothyroidism	126,062 (4.2)	3,454 (4.8)	0.029	–	0.76 (0.74–0.79)	0	-3
Renal failure	289,047 (9.7)	20,526 (28.7)	0.497	1.63 (1.54–1.73)	2.06 (2.02–2.11)	5	8
Liver disease	49,916 (1.7)	5,822 (8.1)	0.303	2.97 (2.73–3.22)	3.83 (3.7–3.97)	11	16
Peptic ulcer disease, excluding bleeding	5,808 (0.2)	258 (0.4)	0.032	–	–	0	0
AIDS/HIV	2,300 (0.1)	85 (0.1)	0.013	–	–	0	0
Lymphoma	25,049 (0.8)	1,759 (2.5)	0.127	2.55 (2.31–2.81)	2.19 (2.07–2.31)	9	9
Metastatic cancer	119,667 (4.0)	18,907 (26.5)	0.657	3.30 (3.10–3.52)	4.09 (3.98–4.21)	12	17
Solid tumour without metastasis	268,298 (9.0)	24,046 (33.7)	0.631	1.47 (1.39–1.56)	2.36 (2.3–2.42)	4	10

Rheumatoid arthritis/collagen vascular diseases	47,305 (1.6)	1,254 (1.8)	0.013	–	0.91 (0.86–0.97)	0	-1
Coagulopathy	90,551 (3.0)	9,528 (13.3)	0.382	1.30 (1.22–1.40)	2.12 (2.07–2.18)	3	9
Obesity	68,155 (2.3)	1,011 (1.4)	0.065	0.64 (0.53–0.77)	0.59 (0.56–0.63)	-4	-6
Weight loss	98,545 (3.3)	9,527 (13.3)	0.369	1.85 (1.67–2.04)	1.67 (1.63–1.71)	6	6
Fluid and electrolyte disorders	257,618 (8.7)	17,440 (24.4)	0.434	1.61 (1.53–1.69)	1.58 (1.55–1.61)	5	5
Blood loss	19,759 (0.7)	685 (1.0)	0.033	0.81 (0.70–0.93)	0.66 (0.60–0.71)	-2	-5
Deficiency	72,290 (2.4)	1,886 (2.6)	0.013	0.80 (0.71–0.90)	0.54 (0.51–0.56)	-2	-7
Alcohol abuse	96,708 (3.2)	3,086 (4.3)	0.056	–	0.75 (0.72–0.78)	0	-3
Drug abuse	38,044 (1.3)	583 (0.8)	0.045	0.50 (0.42–0.60)	0.67 (0.61–0.73)	-7	-5
Psychoses	29,598 (1.0)	404 (0.6)	0.049	–	0.72 (0.65–0.79)	0	-4
Depression	173,898 (5.8)	3,715 (5.2)	0.028	0.73 (0.67–0.80)	0.73 (0.70–0.75)	-3	-3

Abbreviations: SMD standardized mean difference between alive and mortality cohort, VW[a] van Walraven, `–` excluded in the final model, [b]Row percentage
Note: The total cohort percentages can exceed 100%, as each admission contributes to one or more comorbidities. Swiss weights are calculated by dividing the coefficient of each comorbidity by the coefficient in the model with the smallest absolute value (which is 'diabetes uncomplicated' with a coefficient of 0.084) and rounding to the nearest whole number.

Validation and comparison of weighted comorbidity models

All three comorbidity weighting systems (Charlson, Elixhauser van Walraven and Swiss) indicated higher in-hospital mortality risk than the base model, showing the conditional interpretation of weights for each of the weighted models. Each model performed similarly across all years in validation groups as in the derivation groups. Overall, the c-statistic for the 6-year cohort were: 0.757 (95% CI: 0.755–0.759) for the base model, 0.850 (95% CI: 0.849–0.851) for Charlson, 0.863 (95% CI: 0.862-0.864) for VW Elixhauser and 0.867 (95% CI: 0.865–0.868) with Swiss Elixhauser. These c-statistics were similar in the development and validation cohorts. All differences and the rankings they established among models were statistically significant. (Additional file 1, Table S5). In comparison, the model with Swiss weights discrimination was slightly better with some c-statistic variability across the six years' data.

Additionally, 1% highest predicted value, showed the same order of the model's performance from the observed mortality (base: 10.7%, Charlson: 18.5%, VW Elixhauser: 20.4%, Swiss Elixhauser: 20.9% (Table S6, Additional file 1). As shown in receiver-operating characteristic (ROC) curves (Fig. 1) the Swiss weights model's discrimination was better than the Charlson's or base model's, and only slightly better than the van Walraven's. The NRI confirm this picture (Table 3). Comparing the Swiss weights with VW weights showed an NRI of 1.6% (95%-CI: 1.3–2.0) with differences in predicted probabilities of mortality (among those who died) of 1.4% and differences in predicted probabilities of alive (among those who lived) by 0.02%.

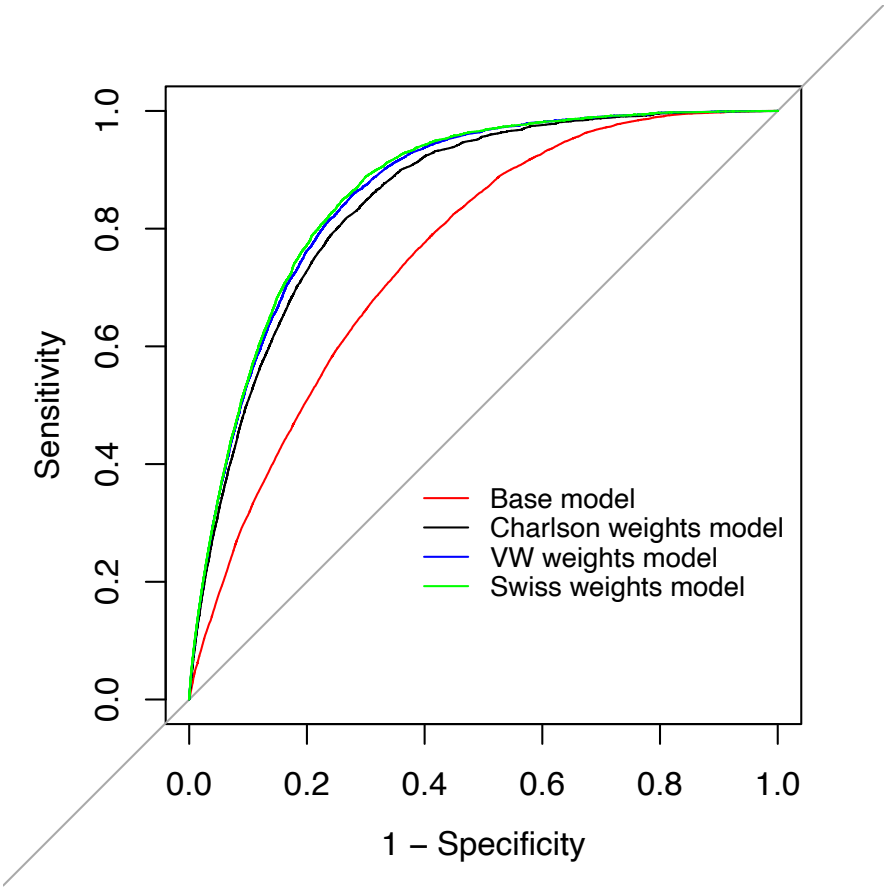


Fig. 1 Receiver-operating characteristic (ROC) curves for generalized additive models predicting in-hospital mortality. Base model (AUC 0.757): age group, sex, hospital types; Charlson weights model (AUC 0.850): base and Charlson weights; VW weights model (AUC 0.863): base and Elixhauser/ van Walraven weights; Swiss weights model (AUC 0.867): base and Elixhauser/ Swiss weights. Straight diagonal line in the middle showing null model (AUC 0.500).

Table 3 Comparison of Swiss weights model with Base, Charlson and VW weights models based on the Net Reclassification Improvement (NRI)

	Derivation group				
Comparison models	NRI (95% CI)	Mortality increased Pr(Up | Case)	Alive increased Pr(Up | Ctrl)	Mortality decreased Pr(Down | Case)	Alive decreased Pr(Down | Ctrl)
Swiss weights vs. Base model	0.355 (0.352–0.357)	0.448 (0.445–0.450)	0.074 (0.074–0.074)	0.134 (0.133–0.136)	0.115 (0.115–0.116)
Swiss weights vs. Charlson weights model	0.049 (0.044–0.052)	0.297 (0.294–0.299)	0.058 (0.058–0.059)	0.251 (0.250–0.253)	0.062 (0.061–0.062)
Swiss weights vs. VW weights model	0.016 (0.013–0.020)	0.157 (0.155– 0.159)	0.021 (0.021–0.022)	0.143 (0.140–0.145)	0.023 (0.023–0.024)

Abbreviations: NRI Net Reclassification Improvement with classification cut-off 0.023, CI confidence interval
Pr(Up, Down) | (Case, Ctrl) represents the proportion of patients whose predicted probabilities increased or decreased for in-hospital mortality and alive cohorts respectively
NRI= (Pr(Up | Case) - Pr(Down | Case)) + (Pr(Down | Ctrl) - Pr(Up | Ctrl))
Base model: age group, sex, hospital types
Charlson weights model: base and Charlson weights
VW weights model: base and Elixhauser/ van Walraven weights
Swiss weights model: base and Elixhauser/ Swiss weights

Finally, the sensitivity analysis using MDCs did not offer any improvements in the models' performance.

4.5 Discussion

This study used a six-year dataset of a multi-million-patient population to explore Charlson and Elixhauser comorbidities with different weightings to predict in-hospital mortality. We first derived a set of Swiss weightings for the 31 Elixhauser comorbidities using the national inpatient dataset. The analysis confirmed Charlson and Elixhauser comorbidities with van Walraven's weights performance for mortality prediction, while the newly derived Swiss weightings slightly improved the mortality prediction for the 31 Elixhauser comorbidities.

Although, the optimized Swiss weightings performed only slightly better than the Charlson and Elixhauser-van Walraven sets they also supplied weights for eight Elixhauser comorbidities (e.g. diabetes, hypertension, and psychosis) eliminated by van Walraven et al. (2009) [25]. Of the risk-associated comorbidities retained in both the van Walraven and the Swiss weights, several comorbidities showed similar

results, e.g., the highest odds ratios to metastatic cancer and liver disease. And regarding the comorbidities with negative associations, only small differences were observed between the van Walraven and Swiss weights (e.g., hypothyroidism or obesity were likely to be healthier).

From an epidemiological perspective, overall hospitalization mortality was only 2.3%, but in-hospital mortality is higher in patients with chronic diseases. Chronic diseases such as cancer, heart and liver diseases increase the risk of dying in hospitals, while certain other less severe diseases, (e.g., hypertension, anaemia and hypothyroidism) have a lower risk. This might be due to the relatively higher frequency of less severe diseases and some reported along with other acute conditions for the same patients. Furthermore, the interpretation of the algebraic sign of a single coefficient from such a joint model is mainly for the derivation of the weights, especially negative weights do not support the survival of the patients. These results are in line with those of Zellweger et al.'s [44] study using the Swiss national death registry of hospital inpatient data from 2010–2012. Furthermore, van Walraven et al.'s [25] study based on a single Canadian hospital's records and Thompson et al., [21] using Maryland State inpatient data, showed similar results. These relations could insight the global burden of in-hospital mortality is due to rising chronic diseases.

The existing weighting systems [11,13,21,25] represent data from a specific geographical region, patient group, or even limited numbers of hospitals or settings, matching the generalizability of these weighting systems remained difficult. As this study addresses such issues, with a large dataset representing the Swiss inpatient population, it provides Swiss comorbidity adjustments for the prediction of mortality or other health outcomes. The c statistics reported in our study (weighted models) are around 10% higher than those reported in van Walraven's study [25]. Several reasons might explain this increase: the GAM modelling approach (with binomial family) including random effects contributing around 2% improvement in c statistic without random effect, the study cohort and hospital types included might raise the base model and largely the conditional interpretation of weightings effect raised c statistic of weighted models. With the new eight derivations, the additional eight significant variables might have played a role too. However, a slightly improved performance of

the Swiss weights system suggests that it might be worthwhile to derive country- or region-specific comorbidity weights from representative patient populations.

C-statistics and ROCs are widely used to assess predictive performance. Nonetheless, one downside of comparing c-statistic and ROCs is that differences between c-statistics are often small, [45] as it was the case when we compared our new weights and van Walraven's. Over the past decade, it has become common to use NRIs to compare different models' performance, even though it might differ with the cut-offs taken for analysis [39,46]. In our study, taking the same cut-offs for all models, NRI calculations confirmed the three weighting systems' rankings i.e., Swiss, van Walraven and Charlson weights.

The primary strength of this study was the large sample size and the heterogeneity of the Swiss inpatient population across all general hospitals over six years, which made it representative of the entire country. To our knowledge, this study is the first to derive and validate Elixhauser weightings in Swiss hospital inpatient data. We used standard regression methodology for large datasets, including random effects at the hospital level, and internally validated our models. We also used accepted methods to modify our adjusted model into a Swiss weightings system that re-includes the association of several comorbidities (e.g., diabetes, hypertension, Psychoses) formerly excluded from the Elixhauser index in the VW study [35]. Despite differences in individual comorbidities' prevalence and weightings, Charlson, Elixhauser/VW, and the Swiss weights performed well across the derivation, validation, and all-cases groups. We also used NRIs, allowing a robust comparison of model performance. Finally, the methods we applied were explicit and can be replicated by other researchers, who can adjust or control for patient comorbidity via their hospital and national databases. Moreover, the managerial utility could be done using this method by identifying high-risk patients for safe care and by evaluating hospitals performance based on the patient's outcome.

Our study also has certain notable limitations. We derived our weights using statistical criteria, while clinical knowledge might be needed to determine each comorbidity's value. Since we used codes assigned in routine data, the capture of the comorbidities could be influenced by other factors, such as physician and nurse documentation, code assignment accuracy, and the possibility that capture of

comorbidities is biased towards those for which the Swiss DRG / MDC pays more [43,47]. The negative coefficients/weights might be artefacts, as they are computed using routine data and coding of these is influenced by the main diagnose (e.g., deficiency anaemia, diabetes or hypertension are far more likely to be recorded when a patient had few other serious or acute problems). The direction of the coefficients is also driven by the joint adjusted model, which makes the interpretation of a single coefficient not meaningful. Moreover, some researchers believe current comorbidity indices are not suitable for use as predictors of patient-centered outcomes like rehabilitation, readmission, fee-for-services while weightings might differ in restricted cohorts, other outcomes and countries [27]. Additionally, Swiss data protection regulations prevented us from obtaining the inpatients' exact age, we could not differentiate children exactly under 18 years and could not specify each year. This also might have influenced the predictive accuracy of the tested models.

4.6 Conclusions

We found that Elixhauser/van Walraven weightings performed well in a large Swiss dataset and could derive Swiss weightings with statistically significant, yet with a small improvement in mortality prediction. Although the Swiss weightings showed slightly improved mortality predictions, we confirmed the validity of the Elixhauser/van Walraven weightings. The results provide evidence that Elixhauser/van Walraven weightings continue to be the preferred choice for weighting. In the Swiss context and possibly in countries with ICD-10 GM (German Modification) the derived weights are an option and to identify high-risk patients for safe care/treatment. Given access to similar data, researchers could use the methods described here to validate existing weightings such as van Walraven or derive their own country- or region-specific morbidity weights, although improvements might be small.

4.7 References

1. Harrison C, Britt H, Miller G, Henderson J. Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice. BMJ open. 2014;4(7):e004694.
2. Pefoyo AJK, Bronskill SE, Gruneir A, Calzavara A, Thavorn K, Petrosyan Y, et al. The increasing burden and complexity of multimorbidity. BMC public health. 2015;15(1):415.
3. Metcalfe D, Masters J, Delmestri A, Judge A, Perry D, Zogg C, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. BMC medical research methodology. 2019;19(1):115.
4. Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. Bmj. 2015;350:h176.
5. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012;380(9836):37-43.
6. Fraccaro P, Kontopantelis E, Sperrin M, Peek N, Mallen C, Urban P, et al. Predicting mortality from change-over-time in the Charlson Comorbidity Index: A retrospective cohort study in a data-intensive UK health system. Medicine. 2016;95(43).
7. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. Medical care. 2017;55(7):698-705.
8. Divo M, Cote C, de Torres JP, Casanova C, Marin JM, Pinto-Plata V, et al. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine. 2012;186(2):155-161.
9. Toson B, Harvey LA, Close JC. The ICD-10 Charlson Comorbidity Index predicted mortality but not resource utilization following hip fracture. Journal of clinical epidemiology. 2015;68(1):44-51.
10. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998:8-27.
11. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-383.
12. Chang H-J, Chen P-C, Yang C-C, Su Y-C, Lee C-C. Comparison of Elixhauser and Charlson methods for predicting oral cancer survival. Medicine. 2016;95(7).
13. Menendez ME, Neuhaus V, Van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical Orthopaedics and Related Research®. 2014;472(9):2878-2886.
14. Fortin Y, Crispo JA, Cohen D, McNair DS, Mattison DR, Krewski D. External validation and comparison of two variants of the Elixhauser comorbidity measures for all-cause mortality. PloS one. 2017;12(3):e0174379.
15. Dailiana Z, Papakostidou I, Varitimidis S, Michalitsis S, Veloni A, Malizos K. Surgical treatment of hip fractures: factors influencing mortality. Hippokratia. 2013;17(3):252.
16. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. Annals of Internal Medicine. 2007;147(2):73-80.
17. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Medical care. 2004;42(4):355-360.
18. Gasparini A. comorbidity: An R package for computing comorbidity scores. Journal of Open Source Software. 2018;3(23):648.
19. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005:1130-1139.

20. Te Grotenhuis M, Pelzer B, Eisinga R, Nieuwenhuis R, Schmidt-Catran A, Konig R. When size matters: advantages of weighted effect coding in observational studies. International Journal of Public Health. 2017;62(1):163-167.
21. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, et al. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374.
22. Pavlou M, Ambler G, Seaman SR, Guttmann O, Elliott P, King M, et al. How to develop a more accurate risk prediction model when there are few events. Bmj. 2015;351:h3868.
23. Song M-K, Lin F-C, Ward SE, Fine JP. Composite variables: when and how. Nursing research. 2013;62(1):45.
24. Pule ML, Buckley E, Niyonsenga T, Roder D. The effects of comorbidity on colorectal cancer mortality in an Australian cancer population. Scientific reports. 2019;9(1):8580.
25. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009:626-633.
26. Austin SR, Wong Y-N, Uzzo RG, Beck JR, Egleston BL. Why summary comorbidity measures such as the Charlson comorbidity index and Elixhauser score work. Medical care. 2015;53(9):e65.
27. Kumar A, Karmarkar AM, Graham JE, Resnik L, Tan A, Deutsch A, et al. Comorbidity indices versus function as potential predictors of 30-day readmission in older patients following postacute rehabilitation. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences. 2017;72(2):223-228.
28. Stanley J, Sarfati D. The new measuring multimorbidity index predicted mortality better than Charlson and Elixhauser indices among the general population. Journal of clinical epidemiology. 2017;92:99-110.
29. De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288.
30. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research. 2011;46(3):399-424.
31. Wilson M, Welch J, Schuur J, O'Laughlin K, Cutler D. Hospital and emergency department factors associated with variations in missed diagnosis and costs for patients age 65 years and older with acute myocardial infarction who present to emergency departments. Academic Emergency Medicine. 2014;21(10):1101-1108.
32. Li Z, Wood SN. Faster model matrix crossproducts for large generalized linear models with discretized covariates. Statistics and Computing. 2020;30(1):19-25.
33. Li Z, Wood SN. Faster model matrix crossproducts for large generalized linear models with discretized covariates. Statistics and Computing. 2019:1-7.
34. R: A Language and Environment for Statistical Computing [computer program]. R Foundation for Statistical Computing; 2019.
35. Sullivan LM, Massaro JM, D'Agostino Sr RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Statistics in medicine. 2004;23(10):1631-1660.
36. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. American journal of epidemiology. 2011;173(6):676-682.
37. Pencina MJ, D'Agostino RB. Evaluating discrimination of risk prediction models: the C statistic. Jama. 2015;314(10):1063-1064.
38. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine. 2008;27(2):157-172.

39. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Comments on 'Integrated discrimination and net reclassification improvements—Practical advice'. Statistics in Medicine. 2008;27(2):207-212.
40. Pencina MJ, D'Agostino Sr RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statistics in medicine. 2011;30(1):11-21.
41. Inoue E, Inoue ME. Net Reclassification Improvement (NRI) for Risk Prediction Models with Time to Event and Binary Response Data. 2018.
42. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. Journal of clinical epidemiology. 2011;64(7):749-759.
43. Wasserfallen J-B, Zufferey J. Financial impact of introducing the Swiss-DRG reimbursement system on potentially avoidable readmissions at a university hospital. Swiss medical weekly. 2015;145:w14097.
44. Zellweger U, Junker C, Bopp M, Group SNCS. Cause of death coding in Switzerland: evaluation based on a nationwide individual linkage of mortality and hospital in-patient records. Population health metrics. 2019;17(1):2.
45. Loging WT. Bioinformatics and Computational Biology in Drug Discovery and Development. Cambridge University Press; 2016.
46. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Annals of internal medicine. 2014;160(2):122-131.
47. Quan H, Eastwood C, Cunningham CT, Liu M, Flemons W, De Coster C, et al. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). BMJ open. 2013;3(10):e003716.

Chapter 5

Causal effect of capacity utilization on in-hospital mortality: evidence of safety tipping points in Swiss general hospitals

Narayan Sharma[1], MSc; Giusi Moffa[2], PhD; René Schwendimann[1,3], PhD; Olga Endrich[4], MD; Dietmar Ausserhofer[1,5], PhD; Michael Simon[1,6], PhD

[1]Institute of Nursing Science (INS), Department Public Health (DPH), Faculty of Medicine, University of Basel, Basel, Switzerland
[2]Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland
[3]Patient Safety Office, University Hospital Basel, Basel, Switzerland
[4]Directorate of Medicine, Inselspital University Hospital Bern, Bern, Switzerland
[5]College of Health Care-Professions Claudiana, Bozen, Italy
[6]Nursing Research Unit, Inselspital University Hospital Bern, Bern, Switzerland

Submitted to peer review journal

5.1 Abstract

Background: High bed-occupancy (capacity utilization) rates are commonly thought to increase in-hospital mortality; however, little evidence supports a causal relationship between the two. This observational study aimed to assess three time-varying covariates—capacity utilization, patient turnover and clinical complexity level—, and to estimate any potential causal effect of high capacity utilization on 14-day in-hospital mortality.

Methods: This retrospective population-based analysis was based on routine administrative data (n=1,152,506 inpatient cases) taken from one unspecified calendar year between 2012 and 2017 for 102 Swiss general hospitals. We investigated potential effects of capacity utilization on in-hospital mortality. Considering the longitudinal nature of the problem, along with available literature and expert knowledge, we represented the underlying data generating mechanism as a directed acyclic graph. To adjust for patient turnover and patient clinical complexity levels as time-varying confounders, we fitted a marginal structure model (MSM) that used inverse probability of treatment weights (IPTWs). We also adjusted for patient age and sex, weekday-vs-weekend, comorbidity weight, and hospital type.

Results: For each participating hospital, our analyses indicated the 85th percentile of highest recorded capacity utilization as a safety tipping point. This occurred at 42.1%–95.9% of each institution's bed-occupancy. For each additional day of exposure to ≥85th percentile (high) capacity utilization, our marginal structure model incorporating IPTWs showed a 2% increase in the odds of 14-day in-hospital mortality (OR 1.02, 95% CI: 1.01 to 1.03).

Conclusions: Causal modelling of data from Swiss general hospitals indicates that patient exposure to high capacity utilization has a causal effect on in-hospital mortality. Each participating hospital had its own tipping point, that is, an occupancy level that consistently leads to increased mortality and requires appropriate human resources to reduce its effect.

Keywords: causal effect, time-varying covariates, capacity utilization, in-hospital mortality

5.2 Introduction

Several observational studies have linked high bed-occupancy (capacity utilization) rates in hospitals with in-hospital mortality [1-3]. While those studies account for numerous factors, they also acknowledge that the associations they show do not indicate causality [1,2]. Logically, though, an unexpected rise in care demand (high patient volume, turnover, and case severity) could exceed a hospital's human resource adequacy on certain days (e.g., on weekends). Such situations would delay treatment for some patients and prevent early recognition of deterioration in others. Both cases would contribute to adverse patient outcomes [3]. The link between capacity utilization and in-hospital mortality warrants further research due to the time-varying exposure[4] of care demand (e.g., capacity utilization) and supply (e.g., staffing). Thus, the possible causal link between time-varying predictors and the outcome (e.g., in-hospital mortality) might require flexible care staffing for safer hospitals in general.

The effect of a time-varying exposure is often overwhelmed by time-varying confounding [5], e.g., the effect of time-varying confounders on exposure-outcome relationships. One example is daily patient flow. This affects our outcome of interest—mortality—by influencing patient exposure (e.g., capacity utilization) at each measurement point. Over time, if the analysis does not adequately adjust for this influence, it will distort the apparent association between exposure and outcome. That is, the initial exposure at a given measurement point affects the value of certain variables (confounders) at subsequent time points. In discussions of causality, this is called exposure- or treatment-confounder feedback (TCF) [6,7].

In a hospital setting, the extent of daily capacity utilization may be influenced both by daily patient flow and by mean disease severity; further, today's capacity utilization might influence tomorrow's patient flow. TCF induces a type of bias not generally correctible via standard regression methods, as they employ the same model to infer the effects of early and late exposures; as a result, they cannot control for time-varying factors that arise along the causal pathway between early exposure situations and later outcomes [5].

Correcting for TCF starts with checking for its presence. This requires a rigorous logical exploration of study variables and *a priori* assumptions via expert knowledge and literature so that directed acyclic graphs (DAGs) can be formulated [8]. DAGs

facilitate the understanding, representation and communication of key concepts, making them particularly helpful in identifying confounders and potential sources of bias in exposure-outcome relationships [9,10]. By representing variables as nodes linked by arrows, they show paths with particular characteristics (e.g., chains, forks) [11]. Relying on the concept of directional separation (d-separation) [12,13], Pearl's (1995) graphical rules allow us first to identify sets of variables usable for risk adjustment, then to eliminate bias for a given exposure [14,15].

Causal effects of time-varying exposures are beyond the scope of simpler regression models. For example, neither the logistic regression used by Kuntz et al. nor the Poisson regression used by Madsen et al. was used to identify TCF to assess the association between bed-occupancy and in-hospital mortality [2,3]. Instead, where TCF is present, estimating causal effects with time-varying variables requires Robins' generalized methods (G-methods) [16,17]. The most popular of these is the inverse probability of treatment weighting (IPTW) for marginal structural models (MSMs) [18,19].

Inference of an exposure's causality regarding an outcome relies on a set of *identifiability assumptions* [6]. The three most common of these are 1) ignorability/exchangeability, 2) positivity and 3) correct specification of the IPTW model [6,19,20].

To our knowledge, causal effects of capacity utilization on in-hospital mortality have not yet been specifically investigated. Moreover, previous studies aggregated time-varying exposures (e.g., bed-occupancy rates) into monthly or annual estimates at the hospital level to define each institution's safety tipping point (i.e., the critical point at which adverse events or in-hospital mortality increase significantly) [2,3]. For instance, a study in German hospitals showed a safety tipping point of 92.5% capacity utilization, after which the risk of in-hospital mortality increased significantly [3]. However, these findings do not support a causal interpretation, as any potential TCF is ignored. We also need to consider the level of care each patient requires depending either on disease severity or comorbidity score or on staffing level and individual characteristics such as age or sex, as these could influence a causal relationship between capacity utilization and in-hospital mortality.

To address these limitations, this study aims to investigate the potential causal effect of capacity utilization on in-hospital mortality. To do so it will use routine data from the Swiss national inpatient population. It has two specific aims: 1) to use DAGs to describe a potential mechanism via which the time-varying exposure (i.e., daily capacity utilization) and time-varying confounders (i.e., daily patient turnover and patient clinical complexity levels (PCCL)) affect in-hospital mortality; and 2) to estimate the causal effect of capacity utilization on 14-day in-hospital mortality consistent with such a DAG and using MSM/IPTW-based estimates.

5.3 Methods

Design, context and participants

This is a retrospective longitudinal observational study using patient data routinely submitted to the Swiss Federal Statistics Office (FSO). As stipulated by Article 22 of the Swiss Federal Act on Data Protection, the FSO provided anonymized annual data on all Swiss hospital inpatients from 2012 to 2017. The FSO classifies general hospitals into five types: university hospitals, tertiary care hospitals, large basic hospitals, medium basic hospitals and small basic hospitals. Each institution's classification is based on the number of cases treated per year and/or a special hospital score assigned by the Swiss Medical Association [21,22].

To comply with Swiss data protection regulations, the exact year of the six-year period is anonymized and we included only one annual patient population dataset; to reduce between-hospital heterogeneity, we included only general (acute care) hospitals. Further, as it was impossible to link observations across the calendar year, we excluded patients admitted over the last two weeks of the year. Online supplementary figure S1 provides a flow diagram depicting our inpatient case selection process.

Dataset and variables

The dataset included variables from the FSO's routine administrative data. Additional variables that change over time with fluctuations in care demand (e.g., daily capacity utilization, daily patient turnover and average daily PCCL value) were computed at the hospital level (see online supplementary table S1). Daily patient turnover and daily capacity utilization were computed as percentages per hospital. Average daily

disease severity was computed as PCCL value per hospital per day. PCCL is a measure of the cumulative effect of a patient's comorbidities and/or complications (CC) for each episode of care. Values range from 0 (no CC) to 4 (very severe CC) [23]. The outcome of interest was 14-day in-hospital mortality, i.e., all deaths occurring during inpatients' first 14 days in the hospital, where more than 90% of all inpatients were discharged within the time frame. Additionally, we included individual-level variables (age, sex, Elixhauser index/Swiss comorbidity weights [24]) and variables linked to days of the week, e.g., weekends.

Treatment/exposure strategy

For exposure we considered a binary variable indicating the level of capacity utilization as above or below a critical threshold. We derived each threshold from the distribution of capacity utilization (daily inpatient loads as percentages of the highest daily inpatient load) per hospital. Identifying a critical cut-off relevant to a particular outcome (mortality) is challenging. Some studies have found that a capacity utilization above 80-90% will lead to increases in infection risk, serious medical errors and mortality [2,25,26]. E.g., Kuntz et al.[3] identified a safety tipping point at 92.5% bed occupancy.

However, our study included a heterogeneous group of hospitals whose capacity utilization varies throughout the year [27]. To explore the distribution of capacity utilization for all hospitals we used violin plots, with the 85th percentile as a breakpoint for each hospital, that is, as a safety tipping point cutoff. To evaluate the robustness of this approach we implemented sensitivity analyses using cutoffs bracketing the 85th percentile and the effect estimation was computed. Models were fitted using the generalized estimating equation (GEE) approach in R's "geepack" package [3,28,29].

Exploring relationships between time-varying variables through DAGs

Causal DAGs are mathematically grounded [30] graphical representations of data-generating mechanisms. They provide a visual understanding of the qualitative relationships between different variables. Additionally, thanks to graphical criteria based on the rules of d-separation, the DAG's structure helps researchers identify and evaluate each variable's role, i.e., as a mediator, collider or confounder [6]. Thus, when

we wish to assess an exposure's causal effect (if any) on an outcome, we can depict the available evidence as a DAG [6], which explicitly describes both the model and its underlying assumptions [8,31].

We started development of a causal DAG immediately after this study's conception. Mainly, we were interested in the relationship between capacity utilization and in-hospital mortality, while accounting for hospital and individual-level factors that may influence those relationships. We used two levels of capacity utilization as exposure, 14-day in-hospital mortality as the outcome and patient turnover and PCCL as the main time-varying confounders. No measurement error and no other confounding factors were assumed.

By blocking all back-door (i.e., non-causal) paths [32] between exposures and outcomes, adjusting for possible confounders allows estimation of exposures' causal effect on target outcomes. To construct a DAG that adequately captures current knowledge we used DAGitty (http://www.dagitty.net/), iterating through several steps to synthesize evidence by adding time-varying covariates, individual covariates and unmeasured variables (see online supplementary figures S2A-D).

One underlying assumption is that a particular graphical structure will repeat across various time points. Therefore, to simplify the visual representation, the final DAG in Figure 1 shows only the connection between variables for the first two of the analysis's fourteen-time points. Including the unmeasured variable not captured in our dataset—staffing—in the DAG allowed it to show how that variable might affect both exposure and outcomes, e.g., staffing could be linked with patient turnover [33] and in-hospital mortality [34]. Day-one exposure (to capacity utilization) affects the time-varying confounder (patient turnover) on day two. We also need to acknowledge an unmeasured variable that affects both patient turnover and 14-day mortality forming TCF (Figure 1). Therefore, traditional methods (e.g., stratification, outcome regression) cannot reliably analyses this situation, instead of producing biased estimates of the effect of capacity utilization on mortality [6,7,35].

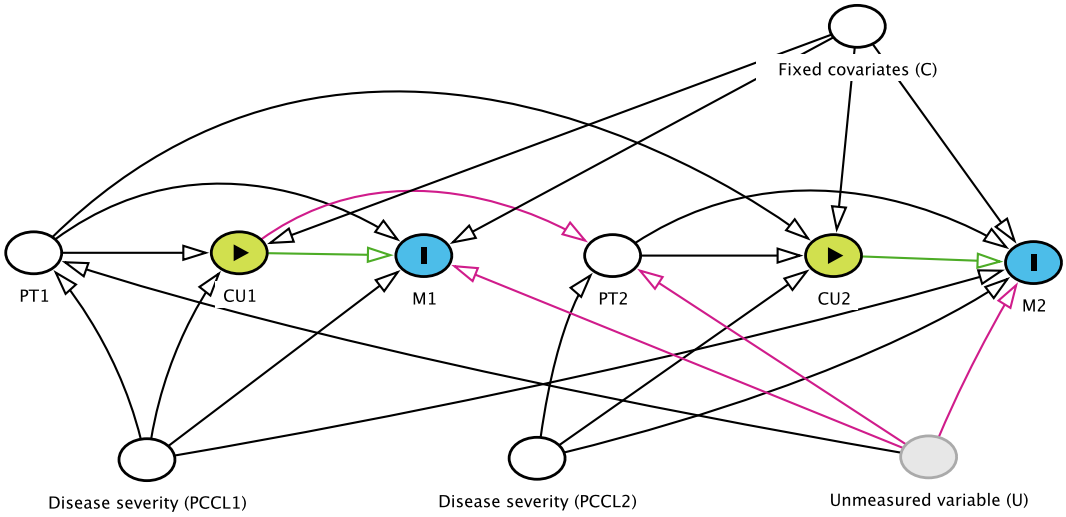


Figure 1: Causal DAG for time-varying exposures (capacity utilization: CU1, CU2), time-varying confounders (patient turnover: PT1, PT2 and average patient clinical complexity level: PCCL1, PCCL2) and outcomes (mortality: M1, M2) at day one and day two with other fixed covariates (C) (e.g., age, sex, comorbidity weights) and unmeasured variables (U) (e.g., staffing).
Note: the arrows are limited to prevent overcrowding. The green path represents a causal path (open), the black path an adjusted (blocked confounder) path and the red path the biasing path as per (http://www.dagitty.net/)

Statistical analysis

We reported the study population's descriptive statistics (including sex, age group, hospital type, Elixhauser index-Swiss comorbidity weighting categories, PCCL categories, patients, day/weekday details) overall and separately for all patients with 14-day mortality (Table 1). Additionally, we described the daily distribution of each time-varying variable (capacity utilization, patient turnover and PCCL) via medians, interquartile ranges (IQRs), and minimum-maximum (Min-Max) per hospital by hospital type

In preparation for statistical evaluation, we used the raw data to derive the time-varying and outcome variables in a format suitable for the intended longitudinal analysis. Each case was followed up for a maximum of 14 days. Where patients left or died within 14 days, their time-varying data were included respectively until discharge or death. Distributions of total exposure days for all study samples were explored for each class of general hospitals [36].

We evaluated the suspected causative role of capacity utilization on in-hospital mortality based on a DAG describing the hypothesized mechanism over time (Figure

1). As our safety tipping point for Swiss hospitals, we decided on utilization equal to or above the 85th percentile of each hospital's patient load. Additionally, we adjusted for time-fixed covariates (age, sex, hospital type, Elixhauser index/Swiss comorbidity weight, and weekdays) that affected both exposure and outcomes.

To estimate a causal effect of exposure to high capacity utilization on in-hospital mortality we fit our MSMs using IPTW,[18] as the weighting eliminates any treatment-confounder feedback regarding the time-varying covariates depicted in the final DAG. At any given time point, IPTW for MSM allows estimates without either (1) adjusting away part of the effect [37], or (2) introducing Berkson's bias (e.g., false observation of a negative correlation between two positive traits) [6,38,39].

For longitudinal analysis, IPTW can be derived for each observation by multiplying weights evaluated at each time point; the resulting weighting is usually standardized to improve precision [6,19]. In this expression of stabilized weights, capacity utilization is the dependent variable. The numerator of the IPTW at each time point includes probability of observed exposure at each time point conditioned with exposure history of previous time point and time-fixed covariates. The denominator includes probability of the observed exposure history of previous time point, time-varying covariates history (patient turnover and PCCL) and time-fixed covariates [20]. To calculate the weights we used the R software's "ipw" package [38].

Marginal structural modelling is flexible enough to handle diverse types of data. In our case, we used multivariable logistic regression, with mortality as the binary outcome. To estimate the effect of increasing exposure by one or more days before the endpoint, our model uses the cumulative number of days with high capacity utilization exposure. To fit the model we used the GEE approach as implemented in R's "geepack" package [40]. We did this first without, then with IPTW [41]. Finally, we derived odds ratios (with 95% CIs) for fourteen-day in-hospital mortality with cumulative exposure to high capacity utilization.

Additionally, the model was adjusted for five fixed covariates: weekend, comorbidity weights, hospital types, sex and age. To complement our primary analysis, we conducted a daily alternative analysis of the total/short-term effect of exposure to high capacity utilization, with observations repeated until the fourteen-day point [42].

Finally, to assess the extent of potential model misspecifications we also considered an analysis with stabilized weights truncated at their 1st and 99th percentiles.

5.4 Results

We analyzed annual data collected over one calendar year on 1,152,506 inpatient cases in 102 Swiss general hospitals, excluding admissions from the study year's final 14 days. Of these, 53.4% were female. One-fifth of admissions were to university hospitals; 36.6% of patients had positive Elixhauser-Swiss comorbidity weighting scores. Ten percent were very severe clinically complex. The overall fourteen-day in-hospital mortality rate was 1.5% (16,998); the death rate was highest (2.3%) in small basic hospitals. Detailed characteristics of the study population, including totals and fourteen-day mortality rates, are shown in Table 1.

Table 1. General characteristics of the study population

	Total study population	14-days mortality (%)
Total population	1,152,506	16,998 (1.5)
Male	536,763	9,662 (1.8)
Female	615,743	7,567 (1.2)
Age groups		
- 0–19 years	152,887	558 (0.4)
- 20–29 years	85,498	79 (0.1)
- 30–39 years	125,141	156 (0.1)
- 40–49 years	99,401	373 (0.4)
- 50–59 years	139,695	1,159 (0.8)
- 60–69 years	163,193	2,270 (1.4)
- 70–79 years	192,614	4,121 (2.1)
- 80–89 years	156,026	5,803 (3.7)
- 90+ years	38,051	2,479 (6.5)
Hospital types		
- University (level 1)	222,552	3638 (1.6)
- Tertiary care (level 2)	688,637	10,591 (1.5)
- Large basic (level 3)	105,809	1,194 (1.1)
- Medium basic (level 4)	118,705	1,191 (1.0)
- Small basic (level 5)	16,803	384 (2.3)
Elixhauser index (mean (SD))	1.41 (1.81)	3.14 (2.06)
Elixhauser index (Swiss comorbidity weighting score)		
- <0	197,262	630 (0.3)
- =0	532,629	1,694 (0.3)
- >0 to <5	76,495	754 (1.0)
- ≥5	345,501	13,641 (3.9)

Individual (PCCL)		
- No clinical complexity (0)	705,437	3,759 (0.5)
- Mild clinical complexity (1)	16,933	65 (0.4)
- Moderate clinical complexity (2)	130,866	1,346 (1.0)
- Severe clinical complexity (3)	174,576	3,915 (2.2)
- Very severe clinical complexity (4)	124,694	7,913 (6.3)
Days of admission		
- Mondays	216,990	2,890 (1.3)
- Tuesdays	204,425	2,700 (1.3)
- Wednesdays	199,349	2,561 (1.3)
- Thursdays	184,595	2,538 (1.4)
- Fridays	161,107	2,591 (1.6)
- Saturdays	86,680	1,844 (2.1)
- Sundays	99,360	1,874 (1.9)
Admission during weekdays		
- Weekdays	966,466	13,530 (1.4)
- Weekends	186,040	3,718 (2.0)

SD: Standard Deviation; PCCL: Patient Clinical Complexity Level

Daily distribution of time-varying variables

Each of the time-varying covariates fluctuated daily across all general hospitals. In University hospitals, while daily capacity utilization and PCCL value were highest, daily patient turnover was lowest. The daily distributions of time-varying variables—capacity utilization, patient turnover and patient clinical complexity level—are shown in Table 2.

Table 2. Daily distribution of time-varying variables capacity utilization, patient turnover and patient clinical complexity level per hospital by hospital type

Hospital Types	Number of hospitals	Capacity utilization (%)		Patient turnover (%)		Patient clinical complexity level (0-4)	
	N	Median (IQR)	Min–Max	Median (IQR)	Min–Max	Median (IQR)	Min–Max
University Hospitals	5	89.3 (83.3–93.5)	55.8–100	21.8 (17.1–29.4)	5.7–38.7	2.06 (2.05–2.07)	0.81–2.57
Tertiary care hospitals	39	79.5 (72.5–86)	27.3–100	29.3 (23.9–34.1)	2.7–54.6	1.78 (1.78–1.79)	0.42–2.75
Large basic hospitals	15	72.3 (62.9–81.1)	13.1–100	33.3 (25.9–39)	0–75.4	1.46 (1.45–1.47)	0.09–2.50
Medium basic hospitals	27	66.4 (55.1–76.5)	5.9–100	34.5 (27.5–41.5)	0–109.1	1.26 (1.25–1.27)	0.00–2.93
Small basic hospitals	16	59.1 (41–76.7)	1.7–100	21.4 (5.9–37)	0–200	1.65 (1.63–1.67)	0.00–4.00

IQR: Inter Quartile Range; Min–Max: Minimum–Maximum; 0–4: (respectively) No clinical complexity, Mild clinical complexity, Moderate clinical complexity, Severe clinical complexity, Very severe clinical complexity; Maximum utilization on hospital level was 100% on the day with the most patients admitted during the year

Treatment-exposure strategy and distribution

Across general hospitals, the range of capacity utilization was distinct but broad—and much broader in small hospitals. On the hospital level, then, 85th-percentile capacity utilization, which we considered the safety tipping point for high exposure,[3] was unique for each institution. Details of exposure to high capacity utilization—reaching the cutoffs evaluated per hospital type—are explored in Figure 2. For university hospitals tipping points ranged from 92.8% to 95.9%; for tertiary care hospitals they ranged from 81.4% to 93.5%; for large basic hospitals from 77.4% to 90.9%; for medium basic hospitals from 64.9% to 86.8%; and for small basic hospitals from 42.1% to 91.1%. The distribution of safety tipping points across hospital types is synthesized in online supplementary figure S2. The total study population was exposed to 6,867,658 hospital days, of which 1,279,021 (18.6%) included high capacity utilization. The distribution of patient treatment-exposure days by hospital type is shown in the online supplementary table S2. The robustness of the safety tipping point analysis, using cutoffs different than the 85th percentile of capacity utilization, is demonstrated in online supplementary table S3. These results appear consistent with one another.

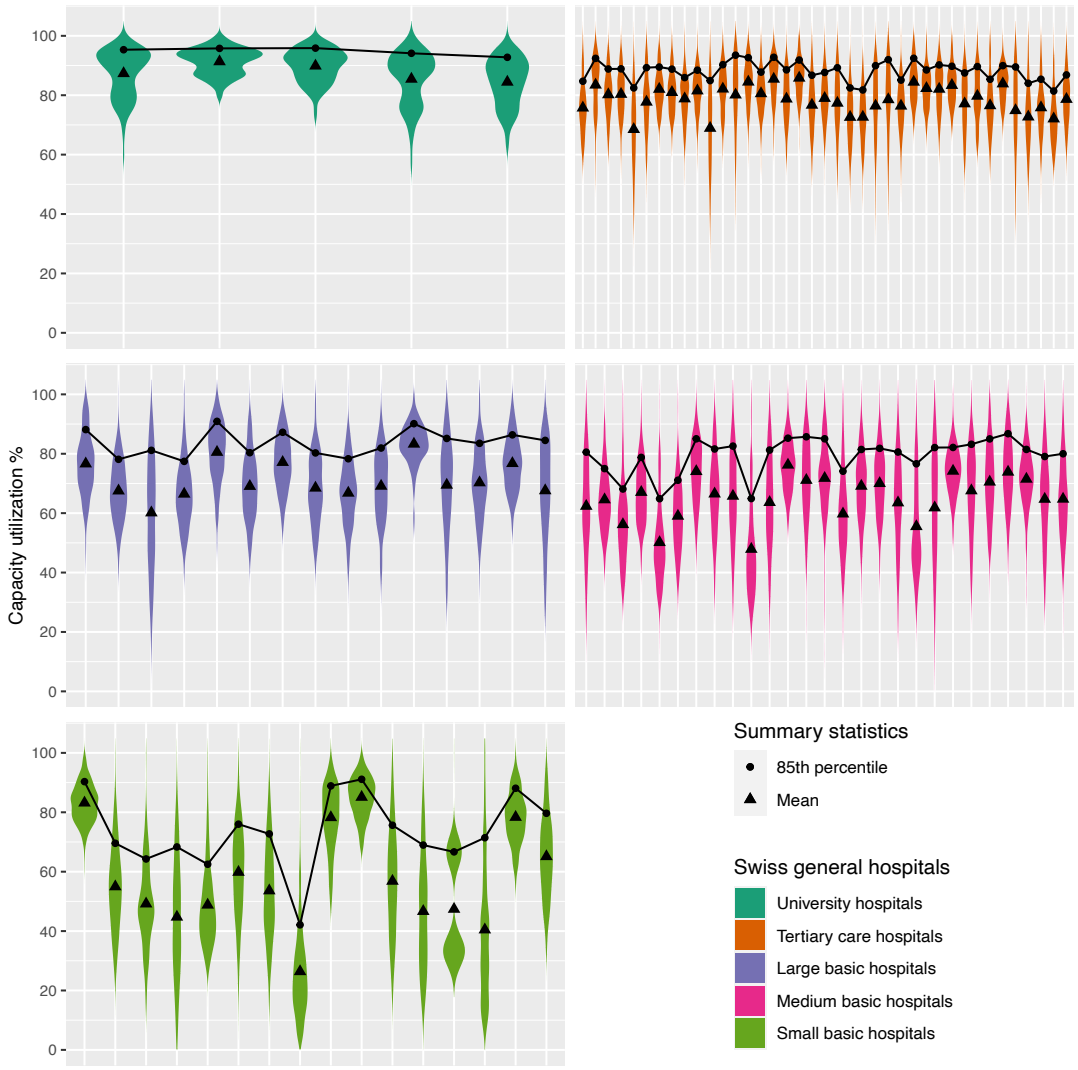


Figure 2. Violin plots showing the density distribution of daily capacity utilization (exposure) across 102 Swiss general hospitals for a study year with 85th percentile (safety tipping point) and mean capacity utilization per hospital type

Causal effect of capacity utilization on in-hospital mortality

The MSMs fitted via GEE (and including stabilized IPTW) allow weighted estimates of the causal effect of exposure to high capacity utilization on mortality (Table 3). One additional day of exposure to high capacity utilization increases the odds of 14-day in-hospital mortality by 2% (OR 1.02, 95% CI: 1.01 to 1.03). The distribution of computed stabilized weights was characterized by a median (IQR) of 0.99 (0.93 to 1.05), Min–Max equal to 0.17–18.8 and a mean of 1.00. The density distribution of the stabilized IPTW is shown in the online supplementary figure S3.

For comparison, we also report odds ratios from the same multivariable logistic model as for the MSM—still using GEEs but without IPTW; therefore, this model is not

adjusted for time-varying confounders. Using this model, an additional day of exposure to high capacity utilization was associated with only a 1% increase in the odds of 14-day in-hospital mortality (OR 1.01, 95% CI: 1.00 to 1.02).

Our analysis also highlights other associations. E.g., the odds of dying are 9% higher during weekends than on weekdays; and the odds of dying are higher for higher Swiss comorbidity weighting scores. Moreover, the odds of dying are considerably higher in small basic hospitals than in university hospitals. And the alternative analysis of the total/short term effect of high capacity utilization yielded 10% higher odds of 14-day mortality. As expected, truncating the lowest and uppermost 1% of the IPTW data resulted in a slightly reduced effect but greater precision (online supplementary table S4 and table S5).

Table 3. The adjusted effects of cumulative daily exposure to high capacity utilization on 14-day in-hospital mortality without and with IPTW (MSM)

	Without IPTW			**With IPTW (MSM)**		
	Estimate	**p-value**	**Odds Ratio (95% CI)**	**Estimate**	**p-value**	**Causal Odds Ratio (95% CI)**
Daily exposure to Capacity Utilization						
- Below 85th percentile	Reference		1	Reference		1
- 85th percentile and above	0.012	<0.05**	1.01 (1.00 to 1.02)	0.016	<0.001***	1.02 (1.01 to 1.03)
Other adjusted variables						
Weekdays						
- Weekday	Reference			Reference		
- Weekend	0.091	<0.001***	1.10 (1.06 to 1.13)	0.088	<0.001***	1.09 (1.05 to 1.13)
Hospital types						
- University (level 1)	Reference		1	Reference		1
- Tertiary care (level 2)	0.020	0.312	1.02 (0.98 to 1.06)	0.022	0.25	1.02 (0.98 to 1.06)
- Large basic (level 3)	-0.148	<0.001***	0.86 (0.80 to 0.92)	-0.141	<0.001***	0.87 (0.81 to 0.93)
- Medium basic (level 4)	-0.115	<0.001***	0.89 (0.83 to 0.95)	-0.121	<0.001***	0.89 (0.83 to 0.95)
- Small basic (level 5)	0.219	<0.001***	1.24 (1.12 to 1.38)	0.512	<0.001***	1.67 (1.46 to 1.90)
Elixhauser index (Swiss Comorbidity weights)						
- < 0	Reference		1	Reference		1
- =0	1.090	<0.001***	2.97 (2.71 to 3.26)	1.090	<0.001***	2.97 (2.71 to 3.26)
- >0 to <5	0.847	<0.001***	2.33 (2.10 to 2.60)	0.865	<0.001***	2.38 (2.13 to 2.65)
- ≥5	2.060	<0.001***	7.82 (7.22 to 8.49)	2.080	<0.001***	8.01 (7.38 to 8.70)

Age (5-year groups)	0.028	<0.001***	1.03 (1.03 to 1.03)	0.028	<0.001***	1.03 (1.02 to 1.03)
Sex						
- Male	Reference		1	Reference		1
- Female	-0.266	<0.001***	0.76 (0.74 to 0.79)	-0.270	<0.001***	0.76 (0.74 to 0.79)

Significance codes: ***p <0.001, **p <0.01, *p <0.05
IPTW: Inverse Probability of Treatment/exposure Weight of capacity utilization ≥85th percentile for daily-varying confounders, patient turnover and PCCL (Patient Clinical Complexity Level).

5.5 Discussion

This observational study examined the causal effect of capacity utilization on in-hospital mortality using one-year patient population data from all 102 Swiss general hospitals. To represent the collective knowledge from the literature we drafted causal DAGs. These captured the influence of time-varying confounders and time-fixed covariates on capacity utilization and in-hospital mortality. The final one identified TCF when unmeasured variables (e.g., staffing) influenced both mortality and patient turnover on consecutive days. With an increase of one day in the cumulative number of days for which capacity utilization was high, there was a 2% increase in the odds of 14-day mortality. Increasing the cumulative exposure to high capacity utilization by three days yielded a 4.9% increase in the odds of 14-day in-hospital mortality. Based on our findings, using IPTW for MSM produced an estimate closer to the reality, i.e., we determined by adjusting all time-varying and time-fixed confounders we could think of via our DAG.

The distribution of the study hospitals' capacity utilization showed that each hospital has its safety tipping point. This depends on each hospital's size, available services and resources, with smaller hospitals showing lower and larger hospitals higher tipping points. Madsen et al.'s 2014 study in 72 Danish hospitals used 80–85% for high bed occupancy, correlating this with a 9% increase in mortality [2]. Another study in 83 German hospitals placed the tipping point at 92.5% [3], showing one in seven deaths was possibly related to high occupancy. As these studies took bed occupancy at the time of admission and did not trace cumulative exposure over each hospital stay, their effects appear larger than they are. Moreover, they considered only one specific cut-off for all studied hospitals; therefore, they may not have adequately captured inter-hospital variations in safety tipping points.

This is a large-scale study examining multiple factors that influence capacity utilization and the potential causal effect of capacity utilization on mortality. For instance, we also observed the weekend effect in Swiss hospitals. Similar effects were also seen in a study in Danish hospitals [2] and in others conducted in UK and US hospitals [43,44], all of which may also have been influenced by staffing patterns [44]. Furthermore, the odds of dying were also higher for patients with higher comorbidity weighting scores [24]. These scores showed results similar to those of earlier studies in Canada and the US [45,46]. Overall, though, the relationships those studies identified between capacity utilization and in-hospital mortality also showed the relationship was confounded by care demand factors including comorbidities and staff composition.

As noted, to eliminate the effect of TCF, we computed IPTW for each case. This required including and evaluating time-varying confounders, e.g., daily patient turnover, average daily PCCL value, alongside time-fixed variables. To our knowledge, this is the first use of a *G-method* [6] (e.g., fitting MSMs with IPTW[20]) to assess the causal effect of capacity utilization on in-hospital mortality.

Methodologically, then, this study differs in one major way from others that have used traditional methods of risk adjustment:[26] it is the first to adjust for the bias of TCF [2,47]. Further, DAGs allowed us to explore the qualitative relationships that link unmeasured variables such as staffing levels [34,48] with both confounders and outcomes. A study in English general hospitals [34] showed a 3% increase in in-hospital mortality among patients cared for by multiple nursing staff (RNs and nursing assistants). While our data did not include the daily staffing variable but we included it as an unmeasured variable in our DAG to visualize the TCF. Afterwards, we believe that IPTW-informed MSMs corrected our estimates.

From a hospital-managerial perspective, uncovering substantial changes in capacity utilization over time calls for accurate monitoring of capacity utilization and its distribution. Shifts in daily capacity utilization change resource/staff needs, e.g., weekend workloads demand more staff. Volatile capacity utilization might lower the safety tipping points in small hospitals. It could also have negative impacts during periods of increased patient flow (e.g., during the COVID-19 crisis). In small basic

hospitals, dynamic capacity utilization might even partly explain why their odds of inpatient mortality are so much higher than in university hospitals.

This study had certain notable limitations. Firstly, Swiss data protection regulations prevented us from linking patient data from one year to another. Therefore, we were unable to construct full datasets for patients admitted over the entire study year. Secondly, assumptions [19,20] of causal inferences also applied to our study. For example, although we tried to include unmeasured variable (e.g., staffing) in our DAG, we considered applying the ignorability assumption [19,49]. However, we realized there might be a chain of confounders affecting the relationship, e.g., patients' previous experiences with hospital services. Another was positivity [19,49], which we tried to address by using the 85th percentile (of each hospital's highest occupancy) as a tipping point of high exposure across all hospitals, rather than 85% of full capacity utilization. However, for some cases, a very short length of stay could either have high exposure or low exposure. The third important assumption was that our chosen IPTW model is correctly specified. As models to estimate IPTW for large datasets can specify widespread weights, misspecification can occur [49]. We verified that the average IPTW was 1.0 and observed that min-max values were not particularly extreme. To gauge the extent of possible misspecifications, then, we also calculated estimates with 1% truncations of each specified weighting [49].

5.6 Conclusion

This observational study aimed to evaluate the causal effect of capacity utilization on 14-day in-hospital mortality in Swiss general hospitals. To adjust for time-varying confounding and to estimate causal effect via MSM, we employed IPTW. Using literature and expert knowledge, directed acyclic graphs (DAGs) allowed us to determine time-varying exposure, confounders and fixed covariates. Our analyses indicated that a one-day increase in cumulative exposure to high capacity utilization caused a 2% increase in the odds of 14-day in-hospital mortality. Safety tipping points varied across hospitals, depending on each institution's distribution of capacity utilization throughout the year. Some hospitals' widely distributed capacity utilization might have impaired their responsiveness to changing demands, possibly resulting in adverse outcomes including mortality. Finally, hospitals need both to

understand the risk of high capacity utilization and to know when they reach their safety tipping points: administrators must ensure that resources—especially staffing—safely meet daily care demands.

5.7 References

1. Wise J. Cutting medical bed occupancy is linked to reduced patient mortality in UK hospital, study shows. Bmj. 2015;351:h4977.
2. Madsen F, Ladelund S, Linneberg A. High levels of bed occupancy associated with increased inpatient and thirty-day hospital mortality in Denmark. Health Affairs. 2014;33(7):1236-1244.
3. Kuntz L, Mennicken R, Scholtes S. Stress on the ward: Evidence of safety tipping points in hospitals. Management Science. 2014;61(4):754-771.
4. Wiens J, Guttag J, Horvitz E. Patient risk stratification with time-varying parameters: a multitask learning approach. The Journal of Machine Learning Research. 2016;17(1):2797-2819.
5. Vansteelandt S, Sjolander A. Revisiting g-estimation of the effect of a time-varying exposure subject to time-varying confounding. Epidemiologic Methods. 2016;5(1):37-56.
6. Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.
7. McGrath S, Lin V, Zhang Z, Petito LC, Logan RW, Hernán MA, et al. gfoRmula: An R Package for Estimating the Effects of Sustained Treatment Strategies via the Parametric g-formula. Patterns. 2020:100008.
8. Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. International journal of epidemiology. 2020;49(1):322-329.
9. Williams TC, Bach CC, Matthiesen NB, Henriksen TB, Gagliardi L. Directed acyclic graphs: a tool for causal studies in paediatrics. Pediatric research. 2018;84(4):487-493.
10. Textor J, van der Zander B, Gilthorpe MS, Liśkiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. International journal of epidemiology. 2016;45(6):1887-1894.
11. Bello NM, Ferreira VC, Gianola D, Rosa GJ. Conceptual framework for investigating causal effects from observational data in livestock. Journal of animal science. 2018;96(10):4045-4062.
12. Robins JM. Discussion of Causal diagrams for empirical research by J. Pearl. Biometrika. 1995;82(4):695-698.
13. Pearl J. Causal diagrams for empirical research. Biometrika. 1995;82(4):669-688.
14. Matthay EC, Glymour MM. A Graphical Catalog of Threats to Validity: Linking Social Science with Epidemiology. Epidemiology (Cambridge, Mass). 2020;31(3):376.
15. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. International journal of epidemiology. 2018;47(5):1714-1722.
16. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Mathematical modelling. 1986;7(9-12):1393-1512.
17. Vansteelandt S, Joffe M. Structural nested models and G-estimation: the partially realized promise. Statistical Science. 2014;29(4):707-731.
18. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Statistical models in epidemiology, the environment, and clinical trials. Springer; 2000:95-133.
19. Thoemmes F, Ong AD. A primer on inverse probability of treatment weighting and marginal structural models. Emerging Adulthood. 2016;4(1):40-59.
20. Bounthavong M. Using inverse probability of treatment weights & Marginal structural models to handle time-varying covariates.

https://rpubs.com/mbounthavong/IPTW_MSM_Tutorial. Published 2018. Accessed [2021.01.08].
21. De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288.
22. Federal Statistical Office. Health Pocket Statistics 2017. Federal Statistical Office (FSO). https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.4342090.html. Published 2018. Accessed 2020-10-02.
23. Kim S, Choi B, Lee K, Lee S. An Evaluation of the Patient Clinical Complexity Level (PCCL) Method for the Complexity Adjustment in the Korean Diagnosis-Related Groups (KDRG). 2020.
24. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC Health Serv Res. 2021;21(13).
25. Sun Y, Heng BH, Tay SY, Tan KB. Unplanned 3-day re-attendance rate at Emergency Department (ED) and hospital's bed occupancy rate (BOR). International journal of emergency medicine. 2015;8(1):32.
26. Jones R. Bed occupancy and hospital mortality. British Journal of Healthcare Management. 2016;22(7):380-381.
27. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Variation of daily care demand in Swiss general hospitals: a longitudinal study on capacity utilization, patient turnover and clinical complexity levels. In:2021.
28. Royston P, Sauerbrei W. Multivariable modeling with cubic regression splines: a principled approach. The Stata Journal. 2007;7(1):45-70.
29. Højsgaard S, Halekoh U, Yan J, Højsgaard MS. Package 'geepack'. R package version. 2016:1.2-0.2015.
30. Barrett M. An introduction to directed acyclic graphs. https://cran.r-project.org/web/packages/ggdag/vignettes/intro-to-dags.html. Published 2018. Accessed [2020.02.10].
31. Dawid AP. Beware of the DAG! Paper presented at: Causality: objectives and assessment2010.
32. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. Annals of the American Thoracic Society. 2019;16(1):22-28.
33. Park SH, Blegen MA, Spetz J, Chapman SA, De Groot H. Patient turnover and the relationship between nurse staffing and patient outcomes. Research in nursing & health. 2012;35(3):277-288.
34. Griffiths P, Maruotti A, Saucedo AR, Redfern OC, Ball JE, Briggs J, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ quality & safety. 2019;28(8):609-617.
35. Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. bmj. 2018;360.
36. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research. 2011;46(3):399-424.
37. Robins JM, Greenland S, Hu F-C. Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. Journal of the American Statistical Association. 1999;94(447):687-700.
38. van der Wal WM, Geskus RB. ipw: an R package for inverse probability weighting. J Stat Softw. 2011;43(13):1-23.

39. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004:615-625.
40. Halekoh U, Højsgaard S, Yan J. The R package geepack for generalized estimating equations. Journal of Statistical Software. 2006;15(2):1-11.
41. Gerhard T, Delaney JA, Cooper-DeHoff RM, Shuster J, Brumback BA, Johnson JA, et al. Comparing marginal structural models to standard methods for estimating treatment effects of antihypertensive combination therapy. BMC medical research methodology. 2012;12(1):1-8.
42. Keogh RH, Daniel RM, VanderWeele TJ, Vansteelandt S. Analysis of longitudinal studies with repeated outcome measures: adjusting for time-dependent confounding using conventional methods. American journal of epidemiology. 2018;187(5):1085-1092.
43. Sharp AL, Choi H, Hayward RA. Don't get sick on the weekend: an evaluation of the weekend effect on mortality for patients visiting US EDs. The American journal of emergency medicine. 2013;31(5):835-837.
44. Marang-van de Mheen P, Vincent C. Moving beyond the weekend effect: how can we best target interventions to improve patient care? BMJ Quality & Safety. 2021.
45. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, et al. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374.
46. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009:626-633.
47. Boden D, Agarwal A, Hussain T, Martin S, Radford N, Riyat M, et al. Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital. Emergency Medicine Journal. 2016;33(2):85-90.
48. Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. New England Journal of Medicine. 2002;346(22):1715-1722.
49. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. American journal of epidemiology. 2008;168(6):656-664.

Chapter 6

Synthesis and Discussion

In this final chapter, the results of chapters three to five of this dissertation are synthesized and key findings are discussed. Moreover, the methodological strength and limitations of the dissertation are described. The last section of this chapter provides implications to explore time-varying and time-fix factors and to apply causal inference in health service research and its application in hospital settings to improve patient safety and quality of care.

6.1 Key findings

In the first study phase, we conducted a longitudinal analysis of the variation of care demands in Swiss general hospitals. Our findings indicated the daily variation of care demand through capacity utilization, patient turnover, and patient clinical complexity level (PCCL) across 102 hospitals. Average daily capacity utilization and daily PCCL were highest in university hospitals while patient turnover was lowest as compared with the other hospital types. Similarly, patient turnover was highest in medium basic hospitals while showing the lowest PCCL levels, and capacity utilization were lowest in small basic hospitals. Overall, there was a substantial variation of all three care demand measures per day, days of the week, weekend vs weekdays, and seasons throughout the year across the hospitals. This study provided a better understanding of the variation of patients' daily exposure to possible pressure zones in hospitals through three care demand measures **(Chapter 3)**.

In **Chapter 4** we reported on the investigation of comorbidities of Swiss hospital patients as an indicator for complex patients, predicting in-hospital mortality. A new Swiss comorbidity weight was developed for the Elixhauser comorbidity index from a large inpatient sample. The new comorbidity weights were then validated in a different patient sample by comparing the new weights to two internationally used comorbidity weights to predict in-hospital mortality. The predictive performance of in-hospital mortality for the new Swiss weights was slightly higher than the predictive performance of the previously used Elixhauser-based van Walraven weights and Charlson weights. These findings support that patient population-based analysis of in-hospital mortality and seeking country or specific cohort-based weightings. Moreover, the comorbidity weighting score of each patient is the important individual

level confounder and the predictor of in-hospital mortality and for risk-adjustment models.

Finally, in **Chapter 5**, we described and assessed time-varying variables (daily capacity utilization, daily patient turnover, and daily PCCL value) and time-fixed covariates including comorbidity weighting score through literature and expert knowledge building directed acyclic graphs (DAGs), to estimate the causal effect of capacity utilization on in-hospital mortality. Grounded on the DAG and an exposure strategy of capacity utilization based on 85-percentile of safety tipping points in Swiss hospitals, causal effects of capacity utilization on in-hospital mortality were estimated with inverse probability of treatment weights (IPTW) for the marginal structure model (MSM). Safety tipping points of high capacity utilization ranged between 42.1% to 95.9% across the Swiss general hospitals. There was a 2% increase in odds of 14-day in-hospital mortality, for each additional day of patients exposed to high capacity utilization.

6.2 System approach for patient safety in hospitals

Hospital is a system, that applies scientific insights to understand the major inputs like care demand (patient volume and complexity) and care supply (staff level) to obtain outputs (e.g., quality of care). A mismatch between care demand and care supply consequently leads to patient safety issues in hospitals [1,2]. Thus, hospitals try to balance the variability of daily care demand and supply, to prevent strain in the hospital system and to reduce adverse patient outcomes [2]. Previous studies have shown that the higher workload of hospital care providers decreases care service rates [3], which might be due to the relatively lower flexibility of supply (e.g., staffing) than care demand (e.g., patients' volume) [3,4]. This dissertation mainly targeted the care demand as a major component of the hospital system and provided new insights on its causal link with patient safety and in-hospital mortality.

6.2.1 Care demand as a leading component of safety tipping points

As described in **Chapter 1**, in the iceberg model of system thinking, incidents (e.g., all-cause mortality) occurring in hospitals might be the results of underlying hospital structures and service processes (e.g., teamwork, communication), and the

patterns/trends of the causes, like variation of daily capacity utilization [5]. The structures of hospitals, patient volume, and disease severity cumulatively represent the care demand in the hospital and therefore refer to total workload. For instance, larger hospitals accommodate higher patient volumes and higher complexity of diseases [6] (care demand) than smaller hospitals. Yet, the higher care demand requires higher resources (e.g., more staff or higher cost) to accommodate the total workload [7]. The foremost part of the dissertation analyzed care demands in Swiss hospitals with measures of capacity utilization, patient turnover, and PCCL **(Chapter 3)**. The variation on care demand represents the major component of the system approach depending on the types of Swiss hospitals, trend/patterns of the causes throughout the year.

The findings of our study showed **capacity utilization** in larger hospitals was higher than in small hospitals. Since we measured daily care demand, we could describe demand variation by days of the week, weekends, seasons, and during holidays. Moreover, other studies conducted in US hospitals revealed that the average bed occupancies increase with the increasing number of beds in the hospitals [8,9]. Studies have shown that capacity utilization is lower by 20% on the weekends [10,11], which is due to the limited staff and services at these times [12]. Similarly, studies have also shown the hospital admissions were significantly higher in the winter than in other seasons [13] and during holidays (e.g., Christmas) hospital discharges were relatively high [13,14]. These changes result, high capacity utilization during some days of the year which crossed safety tipping points with the higher patient risk of death or adverse events.

Daily **patient turnover** another component of care demand which was lower in university hospitals compared to smaller hospitals, since larger hospitals have not only high occupancy rates but also longer length of stay compared to smaller hospitals [8]. Patients' turnover is generally measured with the percentage of daily patient admissions, discharges, and transfers within the units of the hospital from the total patient count of that day [15,16]. Some studies used the *inverse of the length of stay (1/LOS)* [17,18] as an average measure of turnover from the admission and discharged patients in hospitals or units. However, we used daily patients' admissions and discharges over the total patient count of that day in the hospital as a bottom-up

approach. Patient turnover and its variability were higher during weekdays and lower at the weekends, with similar findings in another study done in a Swiss university hospital [4]. This shift of patient turnover entails a lower workload during the weekend [19,20] and the differences in care provision between weekdays and weekends [21]. Similarly, in our study lower patient turnover was also observed during holidays (e.g., Summer, or Easter) and at the end of the year (e.g., Christmas), which was influenced by the higher number of patient discharges and lower admissions. Discharged patients during holidays and at the end of the year might have prompt outpatient follow-up and a higher risk of death or readmission [14].

Another measure of care demand was the patient's **severity or complexity of diseases** in hospitals. Even though some studies used comorbidity or multi-morbidity or case-mix index for measuring patient's severity [22,23], we used PCCL (a cumulative measure of comorbidities and complexities) [24] with levels of severity ranging from 0-4: no clinical complexity to very severe clinical complexity. Unlike the case-mix index, PCCL depicts a patient's health conditions (ICD 10 codes) and is less influenced by treatment costs and DRG algorithms [25] representing one component of care demand in hospitals. Unsurprisingly our study showed the highest daily PCCL values in university hospitals and lowest value in smaller hospitals, which was in line with similar observations of the highest disease complexity in university hospitals [8,25]. Interestingly, PCCL was higher in small basic hospitals than in large and medium basic hospitals in Switzerland. This difference might be due to older patients admitted to smaller hospitals due to the geographic proximity [26] and more often, general medical problems are treated in small hospitals. Moreover, daily PCCL was highest at the weekends, showing more complex patients staying over the weekend and less severe cases being discharged before the weekend. This difference might be one of the reasons for the weekend effect in hospitals [16,27], with more complex patients being exposed to low-staffed hospitals providing fewer hospital services during the weekend [16]. However, there are few studies considering disease complexity or severity as a care demand perspective [28,29] and included not yet for causal analysis. Yet, our study **(Chapter 3)** accounted for both components of headcounts and severity in a longitudinal analysis.

6.2.2 Care supply and individual covariates of patient safety

We explored mainly three measures of care demand in hospitals in **Chapter 3**: capacity utilization, patient turnover, and PCCL. **The care supply** (e.g., **staffing**, types of hospitals with different services) also influences patients' safety in hospitals [1,30,31]. An English longitudinal hospital study showed an additional 3% increase in in-hospital mortality for each day a patient was exposed to RN staffing below the unit mean [41]. Thus, staffing was explored as an unmeasured variable via DAG being an important dimension of a hospital system, which enabled us to use an alternative method for estimating the causal effect of capacity utilization on in-hospital mortality. Therefore, staffing is an important dimension of the system approach, which depends on the policy and care models in place in the healthcare system [32]. Activity-based funding policies like the Swiss Diagnosis-Related Groups (DRG) [33] try to enhance efficiency by reducing inputs (e.g., staffing or cost) and maximizing outputs [34,35]. Studies have shown patient outcomes like failure to rescue, pressure ulcers, falls, infections, and postoperative sepsis are associated with lower staffing of registered nurses (RN) in hospitals [36-38]. Furthermore, as described in **Chapter 1,** different healthcare professionals provide care for patients, but differences in skills between healthcare professionals can also result in differences in patient safety outcomes. For instance, studies have shown the low RN staffing is associated with poorer healthcare quality [39,40]. Staffing was considered one of the dimensions influencing the risk adjustment model in healthcare.

Care supply also differs between **hospital types** depending on the available services and geography. Some hospitals have high resources and some are operating with low resources, mainly because of the population coverage (e.g., Cities vs countryside) and specialized services [42]. Alongside the resources, the service provided by them differs, meaning university hospitals have high capacity and patient's severity **(Chapter 3)** and consuming more resources while providing a variety of (specialized) services [43] (e.g., to operate as university hospitals in Switzerland must have >100 cumulative service points (units) as per Swiss Medical Association) [44]. Relatively, smaller hospitals have lower capacity and less severe patient cases with lower service points [42]. Moreover, the Swiss hospitals experienced the diversity, by ecology, languages, and safety culture of German, French and Italian speaking regions [45].

These differences were included in our study with hospital-level having an anonymous hospital ID, hospital types representing the difference in services, hospital size, and hospital coverage.

Moreover, **other individual covariates** [46] such as age, sex, comorbidities [47,48] are important predictors of in-patient mortality as described in **Chapter 1**. For instance, hospital-acquired infections by age groups, i.e., children are less likely to be infected than the aging population [49]. With the aging patient population, comorbidities are an important individual factor to link non-clinical and clinical issues (e.g., substance abuse and mental illness) through comorbidity score [46]. Thus, individual comorbidities were considered a strong predictor of adverse events and in-hospital mortality, and it's used for risk adjustment of exposure and outcomes [50-52].

6.3 Comorbidity weighting systems to predict in-hospital mortality

Patients' clinical conditions or combined burden of diseases can be grouped in comorbidity indices as described in **Chapter 1**. Applying the weights of each comorbidity with the comorbidity indices (i.e., Charlson and Elixhauser) helped to derive a weighted summary score of hospital patients **(Chapter 4)**. Depending on the patient population or specific group of patients (e.g., cancer or surgical) new weightings for each Elixhauser comorbidity were developed [50-53] in the past. Among them, Charlson [53] and van-Walraven [52] weights are the most popular weighting systems to predict in-hospital mortality. However, these weights used one hospital for the derivation of comorbidity weights with limiting evidence of the comorbidity weights representing the national wide patient population. Population-based comorbidity weights from the different hospital types might improve the prediction of in-hospital mortality. Moreover, the comorbidity weighting system differs between all hospitalizations and a restricted cohort; mortality and other outcomes; and between the countries [46,54]. Therefore, an analysis of a large heterogeneous patient population from a Swiss national dataset justified an overview of Elixhauser comorbidities in a European sample and optimized the comorbidity weights [46].

Justifying the need for new weights from a heterogenous Swiss patient population was further acknowledged from the validation process applied in **Chapter 4**. Firstly,

the weighting system via a cumulative score for each patient provides an individual representation of each patient's case diagnosed with one or more comorbidities from the 31 comorbidities list of the Elixhauser index [51]. Secondly, the cumulative score is easily handled for the risk adjustment modeling reducing the overfitting bias [55] of many variables in smaller datasets and reducing the computation complexities in large datasets. Furthermore, our study clarifies the process, using c-statistics to validate the new comorbidity weights using validation samples in six years which was persuaded to k-fold cross-validation [56] in machine learning. A similar approach for validating derived algorithms was used in studies related to predictive modeling of health outcomes [52,56,57]. The small differences in c-statistic in our study between Swiss and van-Walraven models were further validated through the visualization of receiver-operating characteristic (ROC) curves and net reclassification improvement (NRI) as novel methods [58,59]. Both the methods justified the differences between the models ranking Swiss, van-Walraven, and Charlson comorbidity weights models. Thus, given access to similar data, researchers could derive country- or region-specific comorbidity weights, and hospital managers could identify high-risk patient groups [60] and evaluate the performance of the hospitals via benchmarking or provider profiling [61] using comorbidity weighting scores.

Finally, **for risk adjustment models** of health outcomes, individual comorbidities and the comorbidities weighting score of each patient accurately justify exposure and outcome [62,63]. For the casual models, comorbidities are the confounders [11,64] which could influence both the hospital exposures and outcomes. Moreover, it could interact with other individual patient and system-level variables (e.g., age, sex, hospital types) and need careful adjustment for outcome prediction, describing patient populations, and causal effect estimation.

6.4 Causal inference using routine data in hospitals

As described in **Chapter 1,** causal inference is a core task of science, regardless of whether the study is randomized or observational [65,66]. Many journal editors request to avoid causal language [67], without being able to make explicit references to causal effects and many observational studies can only be expressed in a meandering way of "association" or "impact" or "benefit"[66]. Another common

phrase in this context is "association is not causation" [68,69] because the methods used to estimate causal effects are not the same as those used to estimate association. We agree, in fact, confounding is always present, and therefore association is not necessarily causation for both randomized and observational studies [66]. However, we feel more confident for the estimate of randomized trials, giving a causal interpretation than the observation study. However conducting RCTs with a large sample to quantify the long-term causal effect is considered unethical when we are interested in potentially unhealthy exposures [70] (e.g., smoking and lung cancer) [66]. Thus, observational data provide a means towards the ends of a natural experiment [71], which needs causal language to describe research aims and evaluate them using scientific methods.

Avoiding the causal language in an observational study makes it impossible to express the research aims unambiguously [66]. Subsequently, the question of causation is appropriate in the title and in the introduction section to describe causal phenomenon towards the aims, in methods explaining the possible ways of estimating causal effects and in discussion providing arguments for and against the causal interpretation of a causal model in an observational study [66,70]. This applies also to capacity utilization and in-hospital mortality, and where it is not feasible to conduct an RCT. In health service research, it's important to define variables as exposures, confounders, and outcomes, from individual and hospital- levels and time-dependent variables, as done in earlier chapters. Nevertheless, one aim of this dissertation was to assess our causal question through a DAG in an observational study and to estimate the causal effect of capacity utilization on in-hospital mortality using routine hospital data from the Swiss patient population **(Chapter 5)**.

6.4.1 Investigating safety tipping points via capacity utilization

It's difficult to identify safety tipping points for a group of hospitals, since the service, patient's volume, and complexity differ with the types and context of hospitals. The study conducted by Kuntz et al. entailed safety tipping points when higher in-hospital mortality was seen above 92.5% of capacity utilization in German hospitals [72]. However, daily variability of capacity utilization fluctuates, on some days patients are exposed to high capacity utilization and on some days to lower capacity utilization.

After a thorough analysis, the single tipping point of capacity utilization across all Swiss hospitals didn't reflect the variability of capacity utilization between hospitals throughout the year. Studies from Denmark and England have suggested between 80-90% of capacity utilization as a single tipping point in hospitals [11,73,74]. However, we explored safety tipping points based on the distribution of capacity utilization in each Swiss hospital at the 85th percentile **(Chapter 5)**. The robustness of choosing the 85th percentile was also evaluated choosing the other possible points around the 85th percentile [72]. As the distribution of capacity utilization varied widely from the university to the small basic hospitals resulting in tipping points for each of the 102 hospitals. Thus, the average tipping point for Swiss university hospitals was 94.8%, which was close to German hospitals in the study done by Kuntz et al. in 2014. However, for Swiss small basic hospitals, the average daily tipping point was 73.5% with a wide range from 42.1% to 91.1%. Each of the Swiss general hospitals has its own safety tipping point and it was further synthesized as the distribution of tipping points by hospital types (Figure 6.1), which is a different approach from a single tipping point across all hospitals used in the past [72]. The tipping points for each of the hospitals were used as a cut-off for exposure to high and low-capacity utilization (exposure/treatment strategy) for estimating the causal effect of capacity utilization on in-hospital mortality.

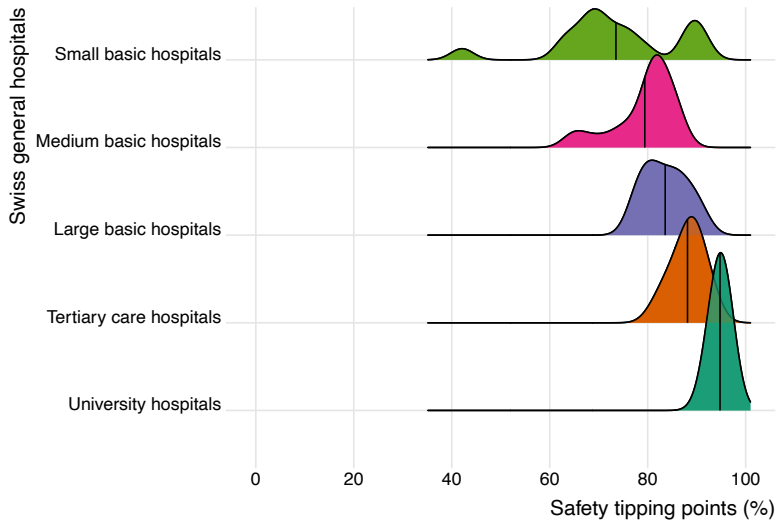


Figure 6.1. Distribution of safety tipping points and vertical line is an average tipping point across Swiss hospital types

From the managerial point of view, when safety tipping points are reached, the hospital will experience a quality problem, which may threaten the survival of the patients and also the quality of care. The variability of capacity utilization in small hospitals is more extreme having some hospitals with relatively low tipping points (e.g., 42.1%), which also indicates that during some days these hospitals experienced more than twice as many patients as on average approaching their maximum capacity. For instance, in small hospitals during the COVID-19 pandemic, capacity utilization tipping points [75] changed as the flow of patients was higher compared to the condition before COVID-19 [76]. Thus, a study like INVEST is required for the assessment of new tipping points. Moreover, small Swiss hospitals having widely distributed capacity utilization, and lower tipping points showed very few patient admissions, or on some days there were no inpatients at all. This raised the question of the need for those hospitals or to reduce the number of hospitals in some regions of Switzerland. Having the wide distribution of capacity utilization in a Swiss hospital at a policy level, planning the number of hospitals [77], capacity pooling [72], and workload smoothing during high patient flow could be alternative measures to control the negative health outcomes after reaching/crossing safety tipping points.

6.4.2 Causal effect of capacity utilization on in-hospital mortality

Investigating causal effects of capacity utilization on in-hospital mortality in hospitals with a patient population is challenging. Since the patient population in hospitals are heterogeneous alongside the different grade of clinical complexity and diseases. With the patient's complexity, the time-varying nature of capacity utilization varies daily throughout the year. Additionally, time-varying confounders like patient turnover and PCCL are linked with exposure, i.e., daily capacity utilization and outcome, i.e., in-hospital mortality. Nevertheless, one aim of this dissertation was to investigate daily capacity utilization, daily patient turnover, and average daily PCCL value via DAG to estimate the causal effect of capacity utilization on in-hospital mortality.

Covariates of capacity utilization and in-hospital mortality were **explored using DAGs**, where evidence from literature and expert knowledge was used. We followed the DAG building steps of Hernán and Robins for time-varying variables [65] by adding time-fixed covariates and unmeasured variables which is somehow similar to

the formal process of evidence synthesis-DAG (mapping, translation, and integration) introduced by Ferguson et al. [78]. With Time-varying confounders (daily patient turnover and daily PCCL) of capacity utilization and mortality, individual variables (age, sex, comorbidity score) were added as fixed covariates in the DAG linking with both, exposures and outcomes. Furthermore, an unmeasured variable (e.g., staffing) was also added in the DAG, as there was evidence that staffing patterns in hospitals might influence patient turnover and mortality [41] showing treatment-confounder feedback (TCF) [65].

TCF was identified through DAG with the unmeasured variable (staffing) in our dataset which is called causal insufficient or latent confounding [78,79]. This condition of having TCF in a longitudinal study is not identified in traditional observational studies e.g., outcome regression or generalized estimating equation (GEE). Thus, the traditional methods with time-varying variables hold the bias which is not eliminated [65,79,80]. Including unmeasured variable in a DAG identify TCF with time-varying variables which are visually identified during the DAG building process [80], which is an additional advantage of using DAG in causal inference and enabled us to choose an alternative method of causal effect estimation e.g., MSM.

Based on the causal DAG and treatment (exposure) strategy of safety tipping points we took longitudinal exposure of patients for 14-days from the day of admission, since it covers 90% of discharged patients. Though complex, **IPTW for MSM** was used for time-varying variables to estimate the causal effect of capacity utilization on 14-day in-hospital mortality. MSM yielded a 2% increase in the odds of 14-day in-hospital mortality when patients were exposed to an additional day of high capacity utilization. The difference in estimate was identified while comparing to the GEE model without IPTW [81]. Thus, IPTW balanced the exposures of patients between low-capacity utilization and high capacity utilization, as we do randomization of patients in the control and treated group in RCT [82]. With this method we believe, IPTW for MSM eliminates biases (TCF with unmeasured variable (e.g., staffing) and non-randomization of exposure groups) [66] compared to traditional regression methods (e.g., GEE) for time-varying variables or outcome regressions in observational studies [65,81,83]. This shows IPTW for MSM in causal inference estimated the effect close to reality.

Comparing results, a study conducted in Danish hospitals reported high capacity utilization was associated with a significant 9% increase in-hospital mortality [11]. Another study in an English general hospital, by reducing bed occupancy from 93.7% to 90.2% decreased all markers of mortality (4.5-4.8%) [73]. However, both studies did not incorporate the natural change in occupancy rates during the hospital stay from the admission day and did not figure out TCF. Additionally, our study included daily cumulative exposure, when the patients exposed to additional three days of high capacity utilization in Swiss hospitals could yield a 4.9% increase in odds of 14-days of in-hospital mortality. This showed a relative increase in the odds of mortality with many days of exposure to high capacity utilization.

The **adjusted variables in MSM** also hold meaningful information while drawing the causal interpretation of capacity utilization and in-hospital mortality. The cumulative comorbidity score of each patient was an important individual confounder, showing the risk of dying increased with an increasing score. Studies conducted in the US and Canada [50,52] also showed comorbidities are important confounders for risk adjustment. Another important indication was the weekend effect in Swiss hospitals. Similar results of weekend effects were observed in UK hospitals, ranging from 2 to 10% [84-86]. Finally, types of hospitals hold an interesting result in our study, that the odds of dying in small basic hospitals was higher than in university hospitals, showing the variability of Swiss hospitals or possibly systemic threat/risk for patients treated in small hospitals.

Finally, with IPTW for MSM, we tried to adjust possible time-varying covariates, confounders, and unmeasured variables in the model to draw causal interpretation, applying identifiability assumptions [65], even though we have large routine data. This **suggests future research** could include staffing/skill-mix or process factors like leadership, safety climate dimensions into the causal model to reflect reality in broader perspectives and to meet ignorability assumptions. Other G-methods proposed by Robin [87], parametric g-formula, and g-computation are the alternatives to MSM [88] which need further exploration with longitudinal/continuous exposures or with different data structures [87]. Even though these methods are often used in epidemiology or RCT, it was our opportunity to explore with hospital data and to increase its scope in observational health service research.

6.5 Strength and limitation of the dissertation

This dissertation ("INVEST study") is the first study in Switzerland utilizing multi-million routine hospital data guided by the iceberg model of system thinking to examine safety tipping points in Swiss hospitals from a causal inference perspective. Moreover, causal modeling contributed significantly to understand the link between capacity utilization and in-hospital mortality from the safety tipping phenomenon in hospitals. The knowledge gained through this dissertation can be translated to healthcare services to monitor/plan care demand and supply longitudinally, prediction of health outcomes via comorbidities adjustment, and causal inference in observational studies.

Care demand variation in all Swiss hospitals was explored longitudinally for daily capacity utilization, patient turnover, and PCCL which are among the main drivers in the hospital system guiding care supply and resource utilization. The opportunity to observe care demand in all general hospitals and hospital types enabled us to consider the context of hospitals or hospital characteristics for further analyses. This showed when hospitals are strained, indicating the time-varying causal elements of the system thinking and where hospital managers/clinicians need attention for providing safe care.

Comorbidity being an individual covariate, it was also our opportunity to utilized all six-year data for the derivation of new Swiss comorbidity weights using the Elixhauser index. Swiss comorbidity weights performed slightly better than Elixhauser van-Walveren and Charlson weights in the Swiss patient population. New weights could be generalized for the prediction of health outcomes including in-hospital mortality locally and internationally. The new weighting enabled us to customize the cumulative comorbidity score as an important confounder representing the complexity of patients in a causal model between capacity utilization and in-hospital mortality.

Causal inference in an observational study provided an insight exploring variables through DAGs, as a novel approach in the field of health service research. DAG helped us to explore exposure, confounders, and outcomes for a causal model from the longitudinal perspective and TCF. Another to identify traditional methods of

stratification or regression that couldn't handle time-varying covariates with TCF and proposed the alternative Robin's G-methods [87]. Causal inference in the observational study provided an opportunity to measure similar causal effects from the routinely collected data without consuming time/resources and without conducting randomized experiments.

Besides these strengths, there are also some limitations of this dissertation, namely regarding the availability of data, and methodological concerns of routine hospital data and assumptions for the causal question:

1) Alongside the variation care demand in Swiss hospitals, we were not able to explore daily care supply (e.g., staffing, skill mix) in a longitudinal perspective. The main reason was the staffing data availability annually, which therefore we couldn't use on daily basis to fit with daily care demand. For care demand, PCCL was explored from discharged patients, since ICD-10 codes were not available for patients who were not discharged. Similarly, we were not allowed to link data between patients across the years, due to the Swiss data protection regulations;
2) In the second part of the dissertation, though we used a very large national set of patient discharge data for the derivation of Swiss comorbidity weights and validate them internally, we were not able to conduct further validation from the external data. This was due to the constraints of available time and resources and the unavailability of external data. In-hospital mortality prediction was based on limited variables, including comorbidity weights, which were extracted through ICD-10 codes recorded by nurses and physicians in the hospitals, where documentation bias might be an issue.
3) In the third part of the dissertation, the major limitations were assumptions on the causal interpretation which are called identifiability assumptions (assumptions that we would have to make even if we had an infinite amount of data) [65]. They include mainly ignorability, positivity, and correct specification of standard weight used in IPTW for MSM. In the DAG building process and IPTW, there might be other confounders that were not included in the study, like staff composition or process factors like leadership or safety culture in hospitals. Another limitation could be dichotomous exposure using

safety tipping points across hospitals, which practically rules out the consideration of any continuous factor [65]. However, it was difficult to handle continuous exposure with bigger heterogeneous data. Apart from these assumptions, having routine data we couldn't trace measurement errors of variables and couldn't avoid the bias, as patients admitted from the last two weeks of the year were excluded.

6.6 Implication of causal inference and its application in hospitals

INVEST was the first study in the field of capacity utilization and in-hospital mortality in Swiss hospitals from a causal inference perspective. Observational studies dealing with routine hospital data can provide data-driven evidence for decision or policy-making and application of those evidence in hospital settings for better patient care. Policymakers and hospital managers could control treatments (exposures) either via planning of hospital admissions and effective utilization of staffing resources based on daily decisions of local health care leaders with the potential of influencing health outcomes and costs in the long run. Based on the experiences of INVEST study, further research considering both demand (patient volume) and supply (staffing levels) dynamics of healthcare might estimate the accurate causal effect of exposures on outcomes in health service research.

Firstly, the use of routine data or big data in healthcare provides information that could be utilized in healthcare systems under strain in a short time. Unlike RCTs, routine data models can provide prompt evidence for decision-making. For instance, the current COVID-19 pandemic is a notable example of hospitals leveraging big data analytics. As in our study, the increasing rise of COVID-19 cases increased capacity utilization and disease severity and has been utilized to forecast short-term needs of ICU beds and resources in Chilian hospitals using autoregressive, machine learning, and epidemiological models [89]. Moreover, hospital care demand analysis was tailored for predictive modeling of the COVID-19 pandemic based on machine learning risk prioritization tools to predict transfers from general wards to ICUs [90]. Likewise, regular studies on care demand measures could identify, weekend and seasonal effects of patient's volume [4,13,28,38] and complex patients for planning staffing levels and resources accordingly.

Secondly, there is potential to use electronic health records for the prediction of health outcomes (like mortality or chronic disease) for better health services and research. This process of developing models, investigating to evaluate specific clinical and non-clinical factors that may contribute to a person's risk of developing a chronic disease or in-hospital mortality, and the evaluation of quality in hospitals or health services [91,92]. In our study, we tried to utilized comorbidities as clinical factors to predict in-hospital mortality and a fair comparison was done with the existing weighting systems. Furthermore, comorbidity weighting system/comorbidity is one of the important predictors of other health outcomes like cancer [93], increase length of hospital stay [94], and predict hospital discharge.

Thirdly, applying a causal inference perspective in health service research helps to identify more appropriate estimates of the causal relationship between capacity utilization and in-hospital mortality. Conducting RCTs in this context would not only be time and resources consuming, but even not possible as assigning patients randomly to high-occupancy and low-occupancy sites seem not feasible. Applying the causal inference methodology using routine data saves resources and time and helps to get to more realistic estimates closer to randomized experiments on a large scale [95]. As explained by Rubin in 2008, many randomized experiments should be duplicated in a similar sample, when designing non-randomized studies whose purpose is to obtain, as closely as possible the same answer of the research question from both ways [96]. Thus, the traditional methods of risk adjustment hold bias with one or another confounding due to the TCF created by time-varying covariates and unmeasured variables, which could be corrected using Robin's G-methods [87]. Therefore, further research should look into methodological approaches to account for a causal relation between hospital exposures and outcomes.

Clinical and management **practice in hospitals or health care systems** was another aspect of this dissertation. Firstly, evidence from research and experiences from clinical practice should be known or health care staff aware of patient's safety issues are the result not only from clinical conditions or severity but also from hospital's structures, trends, and patients and staff's beliefs or values. Secondly, considering healthcare professional's readiness for change, when existing models and organizational practices are not adequate for reducing safety issues [97]. Finally, the

actual change recommended should be acted in hospitals or hospital units to observe an impact. This process of "data, information, knowledge and practice" continuum "learning healthcare system cycle" [98] should be practiced in hospital settings, for instance, evidence from real-world data often used for drug production, pricing, and treatment [99,100].

From the experiences in the INVEST study, we learned that the system approach of care demand and care supply is balancing the health system. In some hospitals, it might be beneficial to invest more resources or staff to cope with daily demand dynamics with some e.g., might need lower staffing and resources. For instance, studies have shown managing unnecessary variability of care demand (e.g., optimizing patient flow by limiting overcrowding [101]) reduced nursing stress and improved patient safety [2]. Furthermore, unit-level analysis (e.g., surgical, medical) of care demand are required for a more accurate estimate, particularly in large hospitals [4]. The variability of care demand measures across Swiss hospitals entails the differences in services in hospitals and units, e.g., in some days admissions coming from emergency surgeries might be double than planned surgeries and medical transfers/referral [101].

Additionally, the insights from the INVEST study and other studies showed how "complex patients" are an important predictor of health outcomes like mortality. The complex patient was represented as a cumulative comorbidity score since every single patient could have multiple clinical conditions whose weighting differs [46,102]. So, the comorbidity count could not address the burden of several diseases in a patient. Comparing comorbidity weighting systems could identify the true preference of the health care system of the nationals or regions e.g. in the US and Europe [46]. Moreover, these methods could be used to identify high-risk patients for safer care and treatment and evaluation of hospitals' performance based on patient's predicted outcomes [51,52].

The causal effect of high capacity utilization depends on the safety tipping points of Swiss hospitals. Safety tipping points in Swiss hospitals are not fixed particularly at 80% or 90% and vary with hospitals. Understanding the variability of hospitals' tipping points, hospital managers and policymakers should initiate policies to lower the capacity utilization as soon as the tipping point is reached [103]. This could be

done by either increasing number of staff per bed occupancy or early discharge planning of elective/non-urgency patients or capacity pooling [72,103] from one hospital to a nearby hospital, which is one of the best examples in COVID-19 pandemics [104]. Lastly, estimating the causal effect of system-related factor-like capacity utilization with mortality could be an example to investigate the effect of other factors like medical treatment, language preference, staffing pattern, cultural differences on health outcomes, using routine data.

6.7 Conclusions

Investigating patient safety in hospitals remains a challenge for healthcare system managers and researchers. Assessing safety tipping points via capacity utilization in hospitals with the causal inference perspective adds promising encounters in an observational study. This dissertation filled knowledge gaps at the level of patients, healthcare providers, policymakers, and researchers. Exploring variation of care demand through capacity utilization, patient turnover, and PCCL longitudinally provided the variability of time-varying covariates as a care demand in hospitals. At the same time, individual comorbidity weighting scores are based on Swiss population-based weights and important predictors of in-hospital mortality. Moreover, adjusting possible confounders with an alternative causal model via DAGs showed patient's exposure to high capacity utilization has a causal effect on in-hospital mortality. Enhancing the useability of routine data in causal inference will ultimately lead to better patients' outcomes via monitoring care demand, provider profiling, and optimal management of healthcare resources/cost based on care demand.

6.8 References

1. Clarke SP, Donaldson NE. Nurse staffing and patient care quality and safety. Patient safety and quality: An evidence-based handbook for nurses. 2008.
2. Litvak E, Buerhaus PI, Davidoff F, Long MC, McManus ML, Berwick DM. Managing unnecessary variability in patient demand to reduce nursing stress and improve patient safety. The Joint Commission Journal on Quality and Patient Safety. 2005;31(6):330-338.
3. Kc DS, Terwiesch C. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. Management science. 2009;55(9):1486-1498.
4. Musy SN, Endrich O, Leichtle AB, Griffiths P, Nakas CT, Simon M. Longitudinal Study of the Variation in Patient Turnover and Patient-to-Nurse Ratio: Descriptive Analysis of a Swiss University Hospital. Journal of medical Internet research. 2020;22(4):e15554.
5. Goodman M. The iceberg model. Hopkinton, MA: Innovation Associates Organizational Learning. 2002.
6. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. New England Journal of Medicine. 2011;364(22):2128-2137.
7. Bohmer R, Pisano G, Sadun R, Tsai T. How hospitals can manage supply shortages as demand surges. Harvard Business Review. 2020;3.
8. Phillip PJ, Mullner R, Andes S. Toward a better understanding of hospital occupancy rates. Health care financing review. 1984;5(4):53.
9. Arfa C, Leleu H, Goaied M, van Mosseveld C. Measuring the capacity utilization of public district hospitals in Tunisia: using dual data envelopment analysis approach. International journal of health policy and management. 2017;6(1):9.
10. Harrison GW, Shafer A, Mackay M. Modelling variability in hospital bed occupancy. Health Care Management Science. 2005;8(4):325-334.
11. Madsen F, Ladelund S, Linneberg A. High levels of bed occupancy associated with increased inpatient and thirty-day hospital mortality in Denmark. Health Affairs. 2014;33(7):1236-1244.
12. Duvald I. Exploring reasons for the weekend effect in a hospital emergency department: an information processing perspective. Journal of Organization Design. 2019;8(1):1-27.
13. Khan RC, Halder D. Effect of seasonal variation on hospital admission due to cardiovascular disease-findings from an observational study in a divisional hospital in Bangladesh. BMC cardiovascular disorders. 2014;14(1):1-6.
14. Lapointe-Shaw L, Austin PC, Ivers NM, Luo J, Redelmeier DA, Bell CM. Death and readmissions after hospital discharge during the December holiday period: cohort study. bmj. 2018;363.
15. Park SH, Blegen MA, Spetz J, Chapman SA, De Groot H. Patient turnover and the relationship between nurse staffing and patient outcomes. Research in nursing & health. 2012;35(3):277-288.
16. Needleman J, Liu J, Shang J, Larson EL, Stone PW. Association of registered nurse and nursing support staffing with inpatient hospital mortality. BMJ quality & safety. 2020;29(1):10-18.
17. Unruh LY, Fottler MD. Patient turnover and nursing staff adequacy. Health services research. 2006;41(2):599-612.
18. Park SH, Dunton N, Blegen MA. Comparison of Unit-Level Patient Turnover Measures in Acute Care Hospital Settings. Research in nursing & health. 2016;39(3):197-203.

19. Swiger PA, Vance DE, Patrician PA. Nursing workload in the acute-care setting: A concept analysis of nursing workload. Nursing outlook. 2016;64(3):244-254.
20. Myny D, Van Hecke A, De Bacquer D, Verhaeghe S, Gobert M, Defloor T, et al. Determining a set of measurable and relevant factors affecting nursing workload in the acute care hospital setting: a cross-sectional study. International journal of nursing studies. 2012;49(4):427-436.
21. Bell CM, Redelmeier DA. Mortality among patients admitted to hospitals on weekends as compared with weekdays. New England Journal of Medicine. 2001;345(9):663-668.
22. Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. The Annals of Family Medicine. 2012;10(2):134-141.
23. Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. bmj. 2020;368.
24. Kim S, Jung C, Yon J, Park H, Yang H, Kang H, et al. A review of the complexity adjustment in the Korean Diagnosis-Related Group (KDRG). Health Information Management Journal. 2020;49(1):62-68.
25. Lerch MM, Rathmayer M, Siegmund B, Wilke M, Wedemeyer H, Stallmach A, et al. Limits of the G-DRG system to reflect Complexity in German University Hospitals. Zeitschrift fur Gastroenterologie. 2020;58(8):747-753.
26. Kim K, Burnette D, An S, Lee M, Cho S. Geographic proximity to neighborhood resources and depression among older adults in South Korea. Aging & Mental Health. 2020:1-7.
27. Sharp AL, Choi H, Hayward RA. Don't get sick on the weekend: an evaluation of the weekend effect on mortality for patients visiting US EDs. The American journal of emergency medicine. 2013;31(5):835-837.
28. Ravaghi H, Alidoost S, Mannion R, Bélorgeot VD. Models and methods for determining the optimal number of beds in hospitals and regions: a systematic scoping review. BMC health services research. 2020;20(1):1-13.
29. Shippee ND, Shah ND, May CR, Mair FS, Montori VM. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. Journal of clinical epidemiology. 2012;65(10):1041-1051.
30. Dunton N, Gajewski B, Taunton RL, Moore J. Nurse staffing and patient falls on acute care hospital units. Nursing outlook. 2004;52(1):53-59.
31. Griffiths P, Saville C, Ball JE, Jones J, Monks T, team SNCTs. Beyond ratios-flexible and resilient nurse staffing options to deliver cost-effective hospital care and address staff shortages: a simulation and economic modelling study. International Journal of Nursing Studies. 2021:103901.
32. Fisher ES, Staiger DO, Bynum JP, Gottlieb DJ. Creating Accountable Care Organizations: The Extended Hospital Medical Staff: A new approach to organizing care and ensuring accountability. Health affairs. 2006;25(Suppl1):W44-W57.
33. Schreyögg J, Stargardt T, Tiemann O, Busse R. Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine European countries. Health care management science. 2006;9(3):215-223.
34. SwissDRG. Swiss Diagnosis Related Groups Version 8.0. In. Bern, Switzerland: SwissDRG AG; 2018.
35. Wasserfallen J-B, Zufferey J. Financial impact of introducing the Swiss-DRG reimbursement system on potentially avoidable readmissions at a university hospital. Swiss medical weekly. 2015;145:w14097.

36. Unruh LY, Zhang NJ. Nurse staffing and patient safety in hospitals: new variable and longitudinal approaches. Nursing research. 2012;61(1):3-12.
37. Patrician PA, Loan L, McCarthy M, Fridman M, Donaldson N, Bingham M, et al. The association of shift-level nurse staffing with adverse patient events. JONA: The Journal of Nursing Administration. 2011;41(2):64-70.
38. He J, Staggs VS, Bergquist-Beringer S, Dunton N. Nurse staffing and patient outcomes: a longitudinal study on trend and seasonality. BMC nursing. 2016;15(1):1-10.
39. Bridges J, Griffiths P, Oliver E, Pickering RM. Hospital nurse staffing and staff–patient interactions: an observational study. BMJ quality & safety. 2019;28(9):706-713.
40. Saville CE, Griffiths P, Ball JE, Monks T. How many nurses do we need? A review and discussion of operational research techniques applied to nurse staffing. International journal of nursing studies. 2019;97:7-13.
41. Griffiths P, Maruotti A, Saucedo AR, Redfern OC, Ball JE, Briggs J, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ quality & safety. 2019;28(8):609-617.
42. Saraceno B, van Ommeren M, Batniji R, Cohen A, Gureje O, Mahoney J, et al. Barriers to improvement of mental health services in low-income and middle-income countries. The Lancet. 2007;370(9593):1164-1174.
43. Aiken LH, Sermeus W, Van den Heede K, Sloane DM, Busse R, McKee M, et al. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. Bmj. 2012;344.
44. De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health systems in transition. 2015;17(4):1-288.
45. Schwendimann R, Zimmermann N, Küng K, Ausserhofer D, Sexton B. Variation in safety culture dimensions within and between US and Swiss Hospital Units: an exploratory study. BMJ quality & safety. 2013;22(1):32-41.
46. Sharma N, Schwendimann R, Endrich O, Ausserhofer D, Simon M. Comparing Charlson and Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data. BMC Health Serv Res. 2021;21(13).
47. Stavem K, Hoel H, Skjaker SA, Haagensen R. Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients. Clinical epidemiology. 2017;9:311.
48. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. Medical care. 2017;55(7):698-705.
49. Ghisolfi S, Almås I, Sandefur JC, von Carnap T, Heitner J, Bold T. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. BMJ global health. 2020;5(9):e003094.
50. Menendez ME, Neuhaus V, Van Dijk CN, Ring D. The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. Clinical Orthopaedics and Related Research®. 2014;472(9):2878-2886.
51. Thompson NR, Fan Y, Dalton JE, Jehi L, Rosenbaum BP, Vadera S, et al. A new Elixhauser-based comorbidity summary measure to predict in-hospital mortality. Medical care. 2015;53(4):374.
52. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009:626-633.

53. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-383.
54. Cho SI, Yoon S, Lee H-J. Impact of comorbidity burden on mortality in patients with COVID-19 using the Korean health insurance database. Scientific reports. 2021;11(1):1-9.
55. Pavlou M, Ambler G, Seaman SR, Guttmann O, Elliott P, King M, et al. How to develop a more accurate risk prediction model when there are few events. Bmj. 2015;351:h3868.
56. Kuhn M, Johnson K. Applied predictive modeling. Vol 26: Springer; 2013.
57. Shalev-Shwartz S, Ben-David S. Understanding machine learning: From theory to algorithms. Cambridge university press; 2014.
58. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology (Cambridge, Mass). 2010;21(1):128.
59. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Annals of internal medicine. 2014;160(2):122-131.
60. Pencina KM, Pencina MJ, D'Agostino Sr RB. What to expect from net reclassification improvement with three categories. Statistics in medicine. 2014;33(28):4975-4987.
61. Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. Bmj. 2015;350:h176.
62. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. American journal of epidemiology. 2011;173(6):676-682.
63. Li B, Evans D, Faris P, Dean S, Quan H. Risk adjustment performance of Charlson and Elixhauser comorbidities in ICD-9 and ICD-10 administrative databases. BMC health services research. 2008;8(1):1-7.
64. Keezer MR, Sisodiya SM, Sander JW. Comorbidities of epilepsy: current concepts and future perspectives. The Lancet Neurology. 2016;15(1):106-115.
65. Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.
66. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. American journal of public health. 2018;108(5):616-619.
67. Ruich P. The use of cause-and-effect language in the JAMA Network journals. AMA Style Insider. 2017. In:2018.
68. Altman N, Krzywinski M. Association, correlation and causation. Nature Methods. 2015;12(10):899-900.
69. Pearl J. Causal diagrams for empirical research. Biometrika. 1995;82(4):669-688.
70. Robins JM, Weissman MB. Commentary: Counterfactual causation and streetlamps: what is to be done? International journal of epidemiology. 2016;45(6):1830-1835.
71. Rosenbaum PR. Observational study. Encyclopedia of statistics in behavioral science. 2005.
72. Kuntz L, Mennicken R, Scholtes S. Stress on the ward: Evidence of safety tipping points in hospitals. Management Science. 2014;61(4):754-771.
73. Boden D, Agarwal A, Hussain T, Martin S, Radford N, Riyat M, et al. Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital. Emergency Medicine Journal. 2016;33(2):85-90.

74. Jones R. Bed occupancy and hospital mortality. British Journal of Healthcare Management. 2016;22(7):380-381.
75. Roxby P, Butcher B, England R. Pressure on hospitals 'at a really dangerous point'. BBC. https://www.bbc.com/news/health-55362681. Published 2020. Accessed [2021.03.11].
76. Thiabaud A, Iten A, Balmelli C, Senn L, Troillet N, Widmer A, et al. Cohort profile: SARS-CoV-2/COVID-19 hospitalised patients in Switzerland. Swiss medical weekly. 2021;151(0708).
77. Klauss G, Staub L, Widmer M, Busato A. Hospital service areas–a new tool for health care planning in Switzerland. BMC health services research. 2005;5(1):1-15.
78. Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. International journal of epidemiology. 2020;49(1):322-329.
79. Malinsky D, Danks D. Causal discovery algorithms: A practical guide. Philosophy Compass. 2018;13(1):e12470.
80. VanderWeele TJ, Hernán MA, Robins JM. Causal directed acyclic graphs and the direction of unmeasured confounding bias. Epidemiology (Cambridge, Mass). 2008;19(5):720.
81. Bounthavong M. Using inverse probability of treatment weights & Marginal structural models to handle time-varying covariates. https://rpubs.com/mbounthavong/IPTW_MSM_Tutorial. Published 2018. Accessed [2021.01.08].
82. Zwahlen M, Salanti G. Causal inference from experiment and observation. Evidence-based mental health. 2018;21(1):34-38.
83. Thoemmes F, Ong AD. A primer on inverse probability of treatment weighting and marginal structural models. Emerging Adulthood. 2016;4(1):40-59.
84. Marang-van de Mheen P, Vincent C. Moving beyond the weekend effect: how can we best target interventions to improve patient care? BMJ Quality & Safety. 2021.
85. Pauls LA, Johnson-Paben R, McGready J, Murphy J, Pronovost PJ, Wu CL. The weekend effect in hospitalized patients: a meta-analysis. Journal of hospital medicine. 2017;12(9):760-766.
86. Honeyford K, Cecil E, Lo M, Bottle A, Aylin P. The weekend effect: does hospital mortality differ by day of the week? A systematic review and meta-analysis. BMC health services research. 2018;18(1):1-13.
87. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. American journal of epidemiology. 2011;173(7):731-738.
88. Wang A, Nianogo RA, Arah OA. G-computation of average treatment effects on the treated and the untreated. BMC medical research methodology. 2017;17(1):1-5.
89. Goic M, Bozanic-Leal MS, Badal M, Basso LJ. COVID-19: Short-term forecast of ICU beds in times of crisis. PloS one. 2021;16(1):e0245272.
90. Cheng F-Y, Joshi H, Tandon P, Freeman R, Reich DL, Mazumdar M, et al. Using machine learning to predict ICU transfer in hospitalized COVID-19 patients. Journal of clinical medicine. 2020;9(6):1668.
91. Betihavas V, Davidson PM, Newton PJ, Frost SA, Macdonald PS, Stewart S. What are the factors in risk prediction models for rehospitalisation for adults with chronic heart failure? Australian Critical Care. 2012;25(1):31-40.
92. Griffiths P, Maben J, Murrells T. Organisational quality, nurse staffing and the quality of chronic disease management in primary care: observational study using routinely collected data. International journal of nursing studies. 2011;48(10):1199-1210.

93. Sarfati D, Koczwara B, Jackson C. The impact of comorbidity on cancer and its treatment. CA: a cancer journal for clinicians. 2016;66(4):337-350.
94. Sarfati D, Tan L, Blakely T, Pearce N. Comorbidity among patients with colon cancer in New Zealand. The New Zealand medical journal. 2011;124:1338-.
95. Rutkowski D, Delandshere G. Causal inferences with large scale assessment data: using a validity framework. Large-scale Assessments in Education. 2016;4(1):1-18.
96. Rubin DB. For objective causal inference, design trumps analysis. Annals of Applied Statistics. 2008;2(3):808-840.
97. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. International journal of medical informatics. 2013;82(5):e73-e86.
98. Dagliati A, Tibollo V, Sacchi L, Malovini A, Limongelli I, Gabetta M, et al. Big data as a driver for clinical decision support systems: a learning health systems perspective. Frontiers in Digital Humanities. 2018;5:8.
99. Eichler HG, Bloechl-Daum B, Broich K, Kyrle PA, Oderkirk J, Rasi G, et al. Data rich, information poor: can we use electronic health records to create a learning healthcare system for pharmaceuticals? Clinical Pharmacology & Therapeutics. 2019;105(4):912-922.
100. Garrison Jr LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: The ISPOR real-world data task force report. Value in health. 2007;10(5):326-335.
101. Litvak E. Optimizing patient flow by managing its variability. Front Office to Front Line: Essential Issues for Health Care Leaders Oakbrook Terrace, IL: Joint Commission Resources. 2005;91:111.
102. Grant RW, Ashburner JM, Hong CS, Chang Y, Barry MJ, Atlas SJ. Defining patient complexity from the primary care physician's perspective: a cohort study. Annals of Internal Medicine. 2011;155(12):797-804.
103. Auger KA, Kenyon CC, Feudtner C, Davis MM. Pediatric hospital discharge interventions to reduce subsequent utilization: a systematic review. Journal of hospital medicine. 2014;9(4):251-260.
104. Alban A, Chick SE, Dongelmans DA, Vlaar AP, Sent D. ICU capacity management during the COVID-19 pandemic using a process simulation. Intensive care medicine. 2020:1-3.

Appendices

Variation of daily care demand in Swiss general hospitals: a longitudinal study on capacity utilization, patient turnover and clinical complexity levels

Multimedia Appendix 1

Description of variables and general characteristics of study population for five Swiss general hospital types.

Table A. Description of the study variables

Variables	Short description
Case identifier	Unique anonymous numeric identifier for each case
Age	Age of patient in five-year groups
Sex	Patient's sex
Hospital identifier	Unique anonymous numeric identifier for each hospital
Hospital type	Hospital FSO classification code
Admission date	Patient's date of hospital admission
Discharge date	Patient's date of discharge from hospital
Primary Diagnosis	ICD-10 GM primary patient diagnosis code
Secondary diagnosis	ICD-10 GM secondary patient diagnosis code
Procedure	CHOP code for procedure taken for patient

FSO: Federal Statistics Office; ICD-10 GM: International Classification of Diseases-version 10 German Modification; CHOP: Swiss Operation Classification

Table B. General characteristics of the study population

Hospital Types	Number of hospitals N (%)	Inpatients (%) N = 1,214,875 Males	Females	Discharges (%) N = 1,198,496 Males	Females	Length of stay (days) N = 1,198,496 Mean (95% CI)	Median (IQR)
University Hospitals	5 (4.9)	116,143 (20.5)	120,506 (18.6)	113,910 (20.4)	118,217 (18.5)	8.76 (8.65–8.86)	4.12 (2.04–9.12)
Tertiary care hospitals	39 (38.2)	339,551 (60.0)	385,766 (59.4)	335,185 (60.1)	380,624 (59.4)	6.06 (6.04–6.08)	3.67 (2.04–6.96)
Large basic hospitals	15 (14.7)	49,930 (8.8)	60,708 (9.4)	49,449 (8.9)	60,123 (9.4)	5.33 (5.29–5.37)	3.33 (1.96–5.96)
Medium basic hospitals	27 (26.5)	52,955 (9.4)	71,624 (11.0)	52,501 (9.4)	7,1029 (11.1)	4.93 (4.9–4.97)	3.25 (1.96–5.38)
Small basic hospitals	16 (15.7)	7,122 (1.3)	10,570 (1.6)	7,051 (1.3)	10,407 (1.6)	7.67 (7.52–7.82)	4.12 (2.04–9.12)
Total	102 (100)	565,701 (100)	649,174 (100)	558,096 (100)	640,400 (100)	6.43 (6.40–6.46)	3.7 (2.00–7.00)

CI: Confidence Interval, IQR: Intra-Quartile Range

Multimedia Appendix 2

Percentage of patient turnover by days, days of the week (Monday to Sunday) and weekdays vs weekend over a year for five Swiss general hospital types.

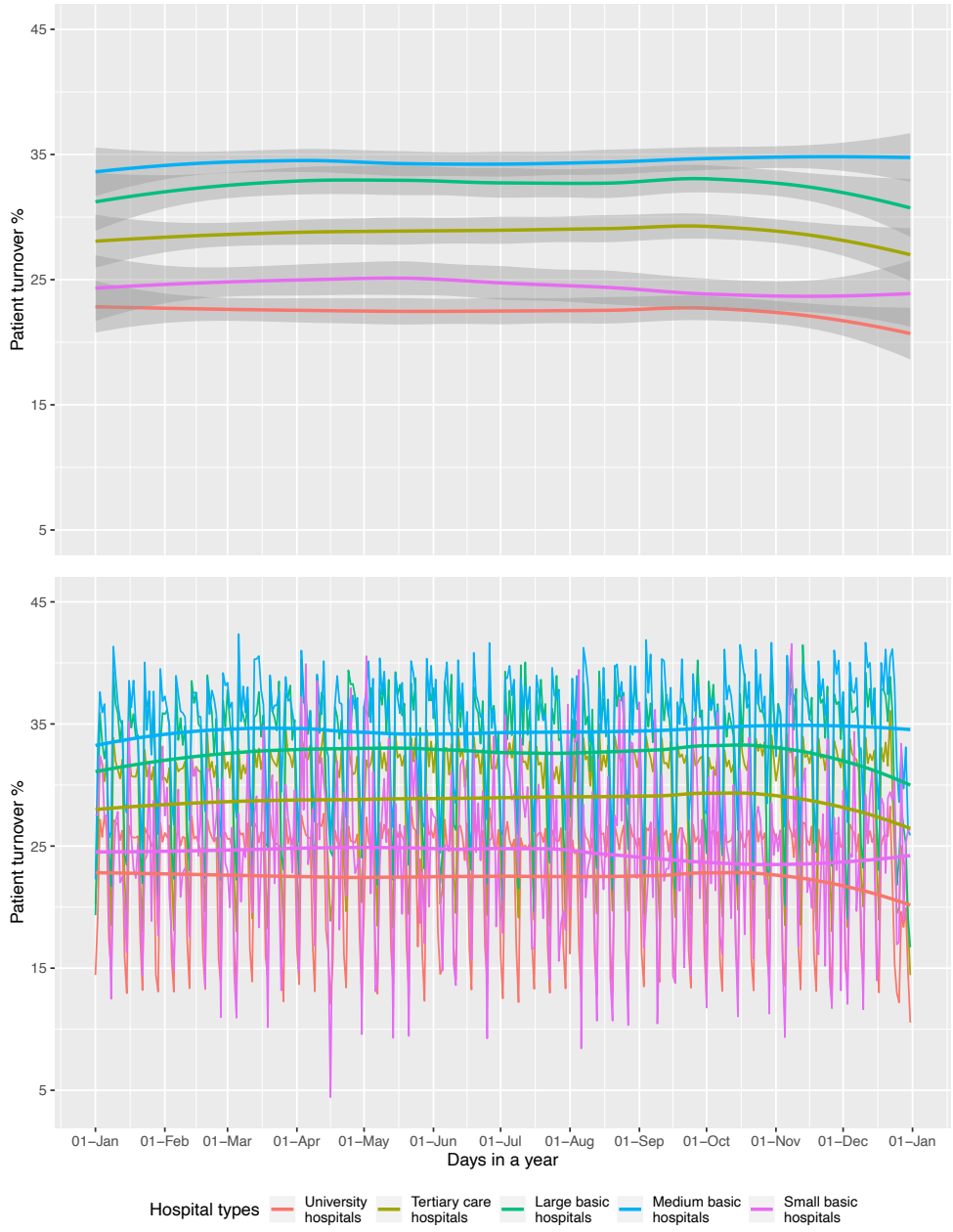


Figure A. Patient turnover of Swiss general hospital types for a year (smooth curve with mean between CI and line chart)

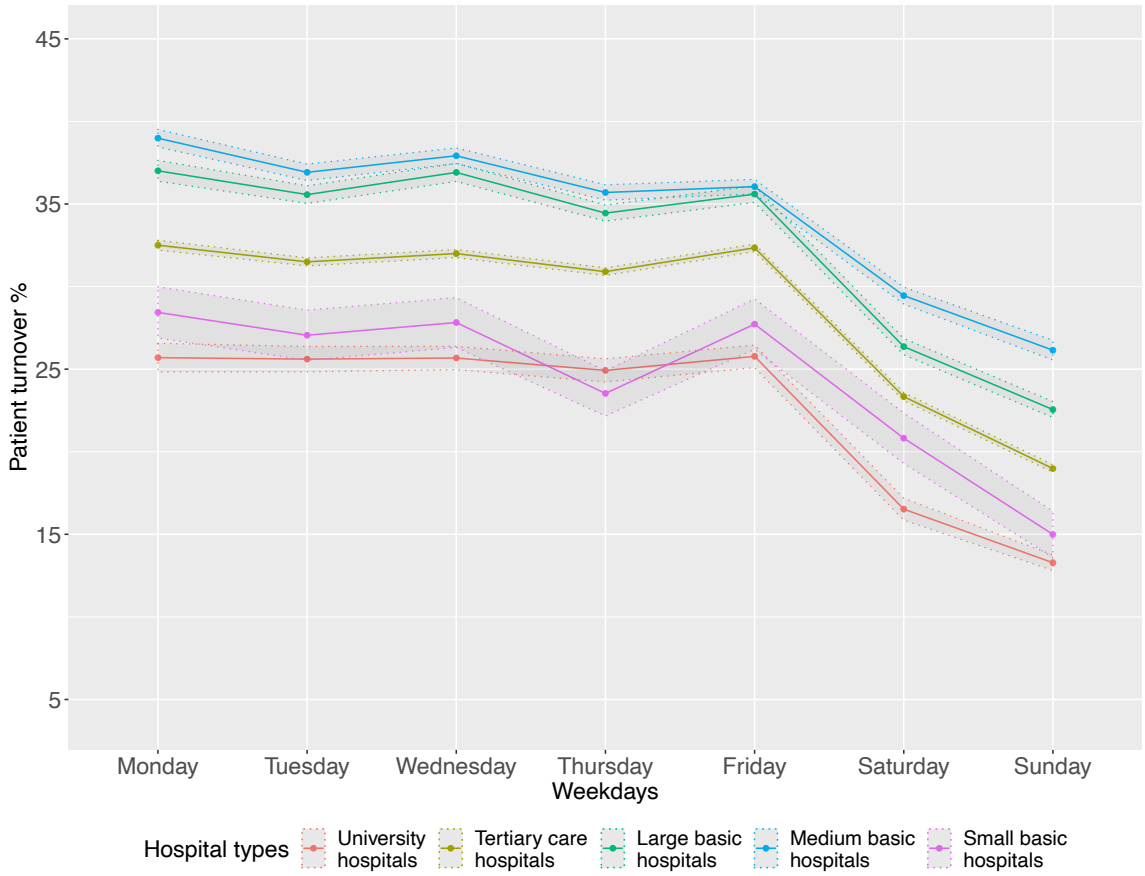


Figure B. Patient turnover of Swiss general hospital types with mean between CI in days of the week

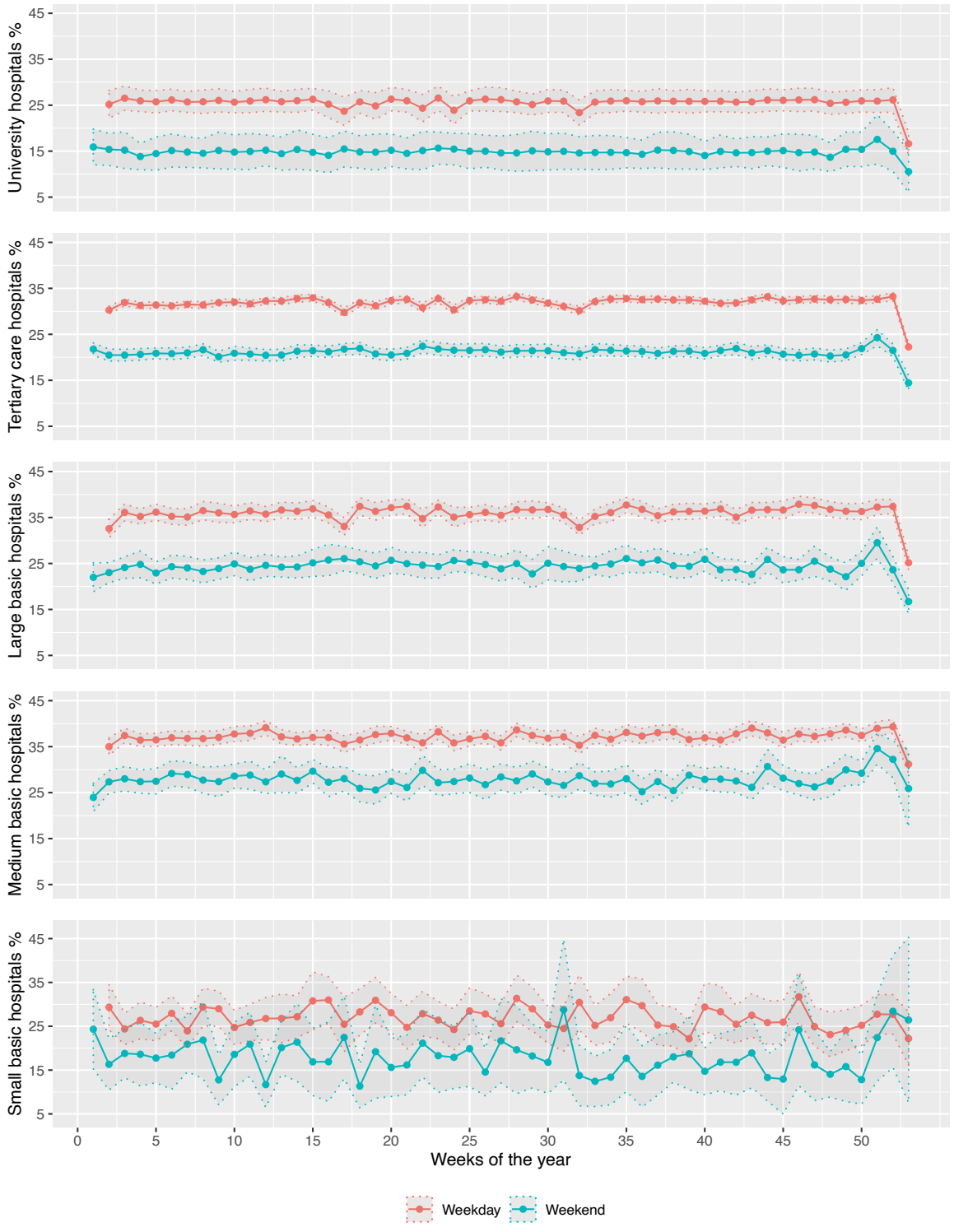


Figure C. Patient turnover of Swiss general hospital types with mean between CI by weekdays and weekend for a year

Multimedia Appendix 3

Average patient clinical complexity levels by days, days of the week (Monday to Sunday) and weekdays vs weekend over a year for five Swiss general hospital types.

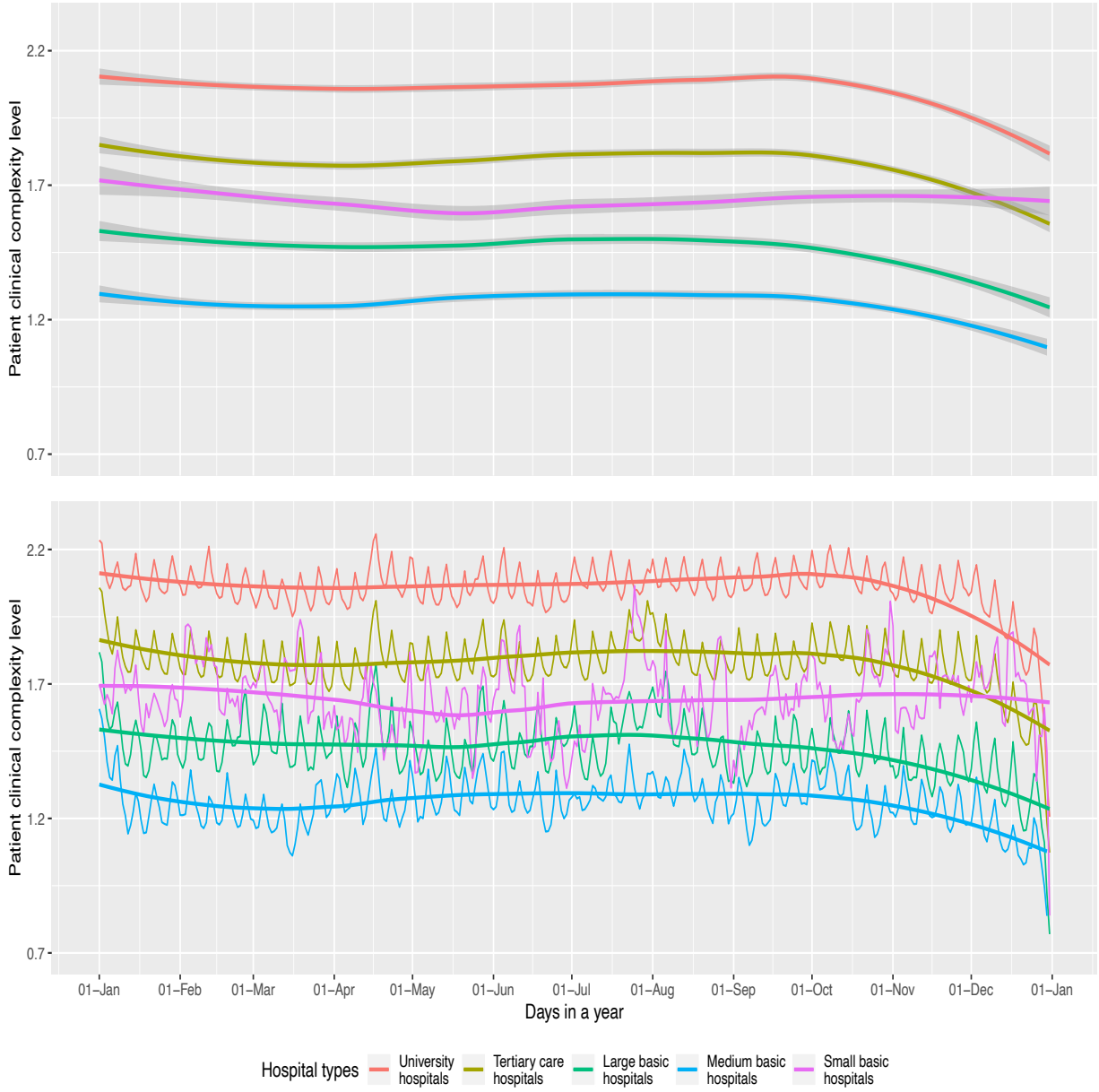


Figure A. Patient clinical complexity level (PCCL) by Swiss general hospital type for one year (smooth curve with mean between CI and line chart)

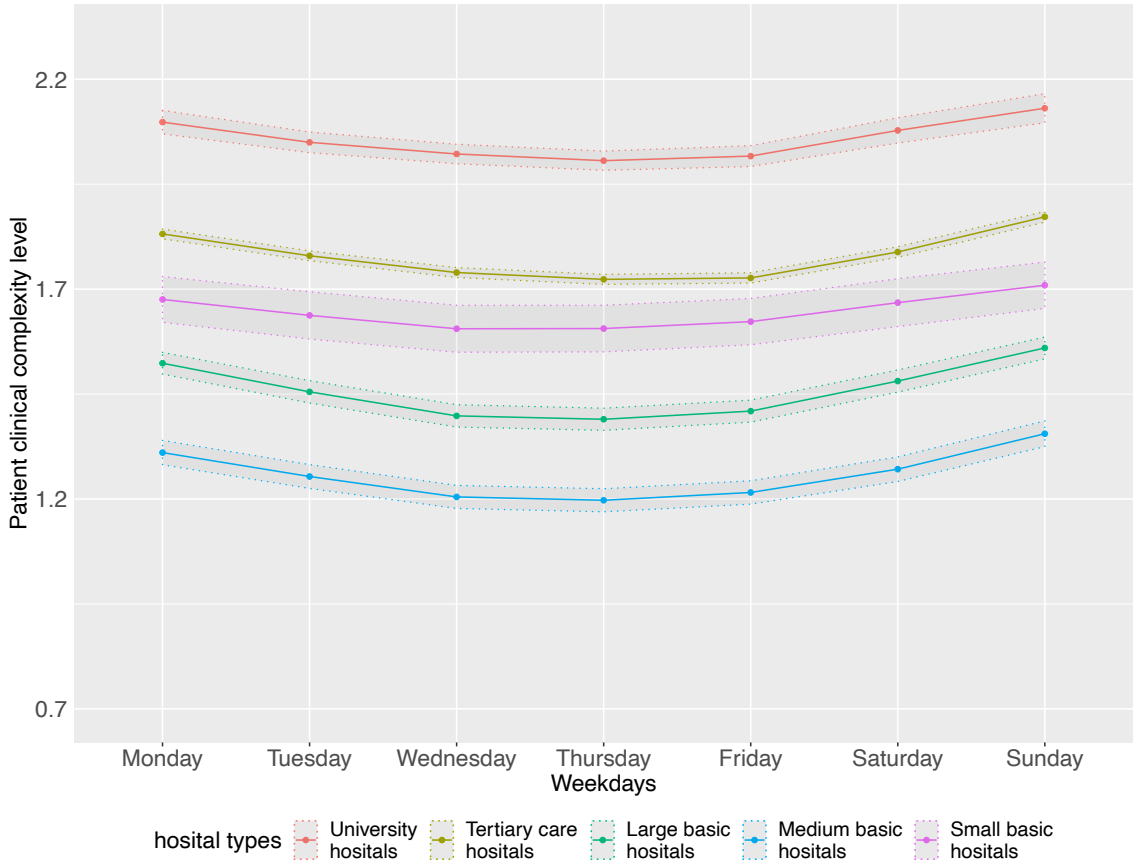


Figure B. PCCL of Swiss general hospital types with mean between CI in days of the week

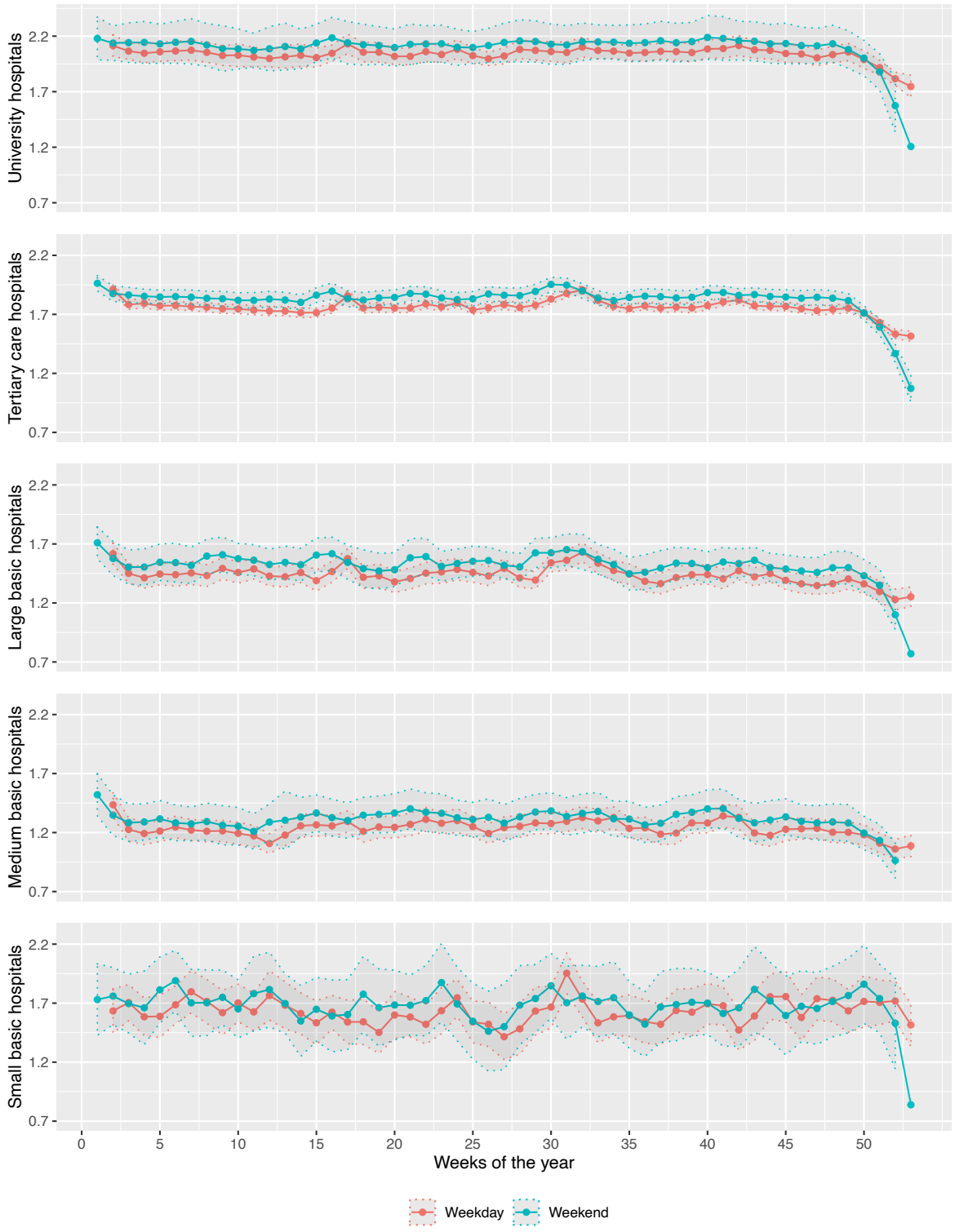


Figure C. PCCL of Swiss general hospital types with mean between CI by weekday and weekend for one year

Multimedia Appendix 4

Distribution of patient clinical complexity levels for last 10 weeks of the year with length of stay and average inpatients on weekly basis (showing complex patient stays over Christmas and the end of the year).

Distribution of PCCL (last 10 weeks/last 70 days of a study year at a university hospital).

PCCL is calculated only with discharged patients, as ICD-10 diagnosis codes are not available for those who are not discharged.

The mean daily PCCL over one study year in a university hospital was 2.38. As the table below shows, patients who were discharged during the last 10 weeks of the year have a lower mean PCCL and shorter length of stay. Thus, the PCCL curve drops at the end of the year. Moreover, capacity utilization also drops at the end of the year, due to fewer admissions and more discharges.

This means patients who were admitted but not discharged during the study year might have higher PCCLs (severity), warranting stays over the Christmas/New Year period.

Table A. Average daily PCCL, LOS and capacity utilization of a university hospital for the last 10 weeks of a study year

Last 10 weeks	Average daily PCCL	Average daily LOS (days)	Average daily inpatients (n)
44	2.27	9.37	918
45	2.22	8.77	929
46	2.21	8.84	915
47	2.19	8.54	944
48	2.23	8.66	942
49	2.27	8.30	893
50	2.17	7.28	883
51	2.03	6.10	838
52	1.94	4.28	644
53	1.78	2.77	277

Observing Length of stay and PCCL through box plot

Patients with higher PCCLs, e.g., PCCL 4, have longer LOSs (IQR 6-20 days); for those with PCCL 3, LOS IQR = 5-10 days, for PCCLs of 2 and 3, LOS IQR = 2-8 days. I.e., patients not discharged during the end of year period likely have higher PCCLs. Daily PCCLs decline sharply at the end of year.

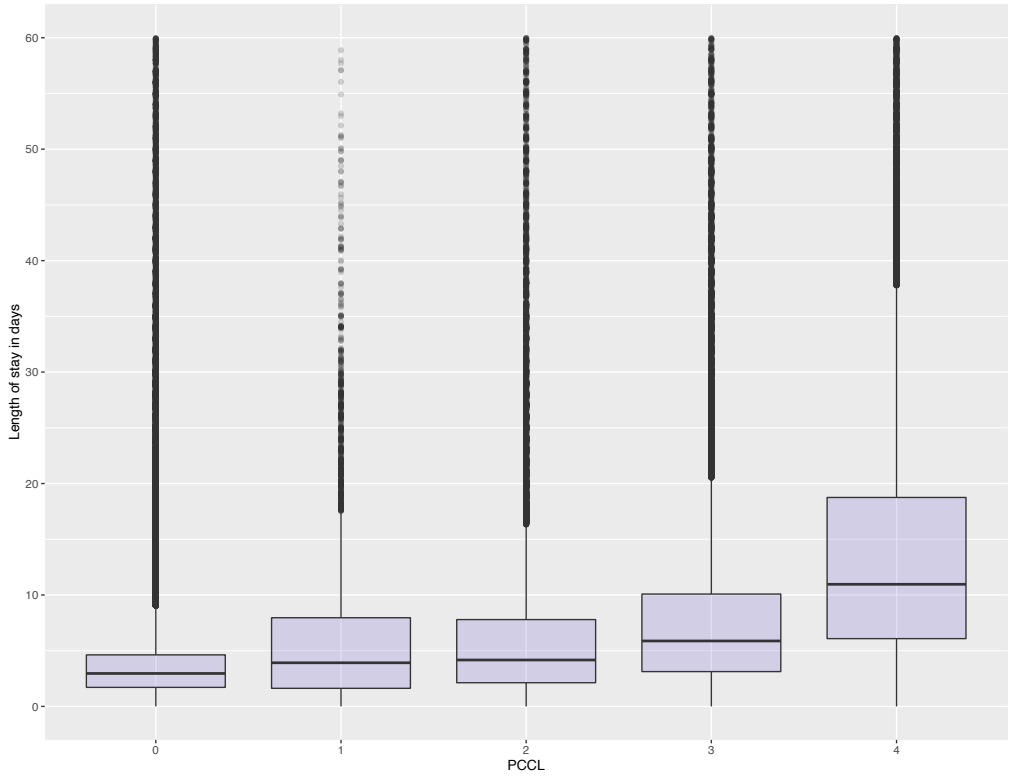


Figure A. Length of stay and PCCL values from all discharge data for one study year

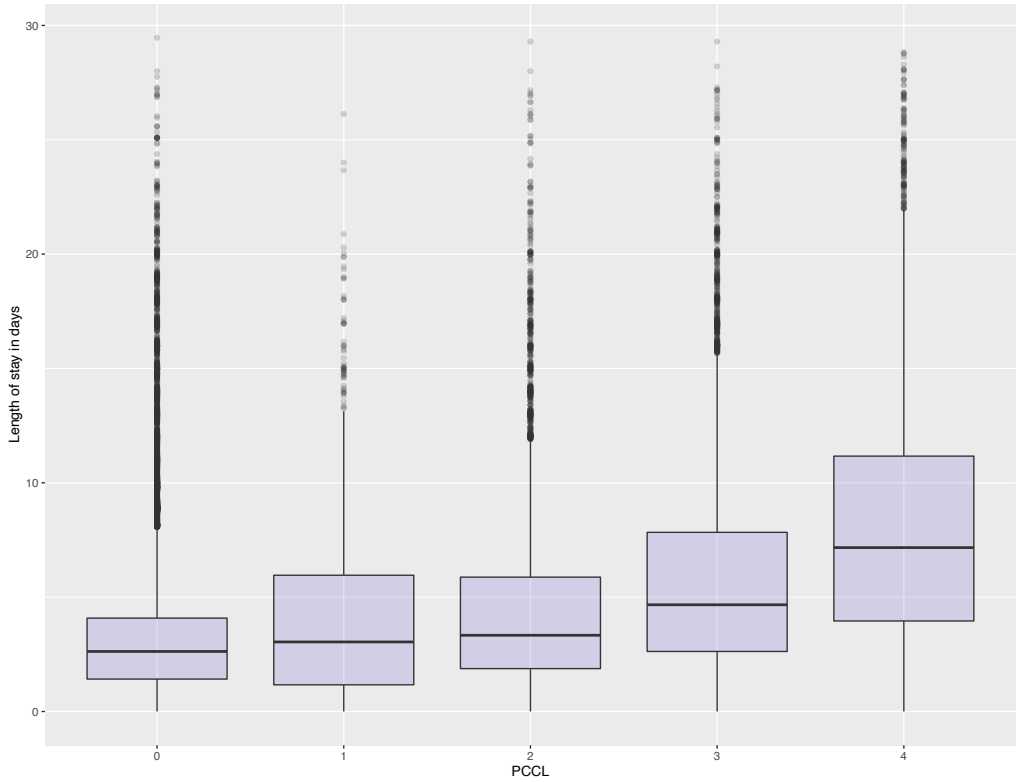


Figure B. Length of stay and PCCL values from all discharge data for November and December

Additional file 1. Supplementary figure and tables

Comparing Charlson, Elixhauser comorbidity indices with different weightings to predict in-hospital mortality: an analysis of national inpatient data

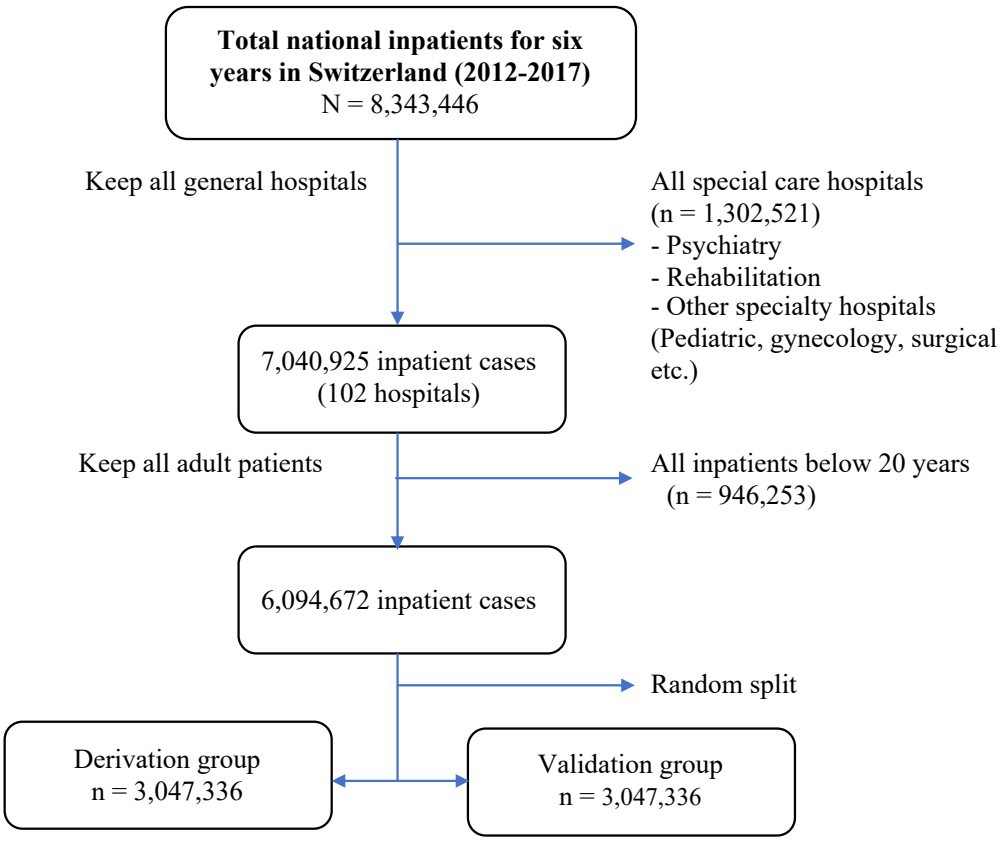


Fig. F1 Flowchart. National inpatient cases of all general hospitals for the analysis

Table S1 The ICD-10 GM codes used in comorbidity package to classify the Charlson and the Elixhauser comorbidity indices

Charlson comorbidity index	**ICD-10 codes**
Myocardial infarction	I21.x, I22.x, I25.2
Congestive heart failure	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5 - I42.9, I43.x, I50.x, P29.0
Peripheral vascular disease	I70.x, I71.x, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9
Cerebrovascular disease	G45.x, G46.x, H34.0, I60.x - I69.x
Dementia	F00.x - F03.x, F05.1, G30.x, G31.1
Chronic pulmonary disease	I27.8, I27.9, J40.x - J47.x, J60.x - J67.x, J68.4, J70.1, J70.3
Rheumatic disease	M05.x, M06.x, M31.5, M32.x - M34.x, M35.1, M35.3, M36.0
Peptic ulcer disease	K25.x - K28.x
Mild liver disease	B18.x, K70.0 - K70.3, K70.9, K71.3 - K71.5, K71.7, K73.x, K74.x, K76.0, K76.2 - K76.4, K76.8, K76.9, Z94.4
Diabetes without chronic complication	E10.0, E10.1, E10.6, E10.8, E10.9, E11.0, E11.1, E11.6, E11.8, E11.9, E12.0, E12.1, E12.6, E12.8, E12.9, E13.0, E13.1, E13.6, E13.8, E13.9, E14.0, E14.1, E14.6, E14.8, E14.9
Diabetes with chronic complication	E10.2 - E10.5, E10.7, E11.2 - E11.5, E11.7, E12.2 - E12.5, E12.7, E13.2 - E13.5, E13.7, E14.2 - E14.5, E14.7
Hemiplegia or paraplegia	G04.1, G11.4, G80.1, G80.2, G81.x, G82.x, G83.0 - G83.4, G83.9
Renal disease	I12.0, I13.1, N03.2 - N03.7, N05.2 - N05.7, N18.x, N19.x, N25.0, Z49.0 - Z49.2, Z94.0, Z99.2
Any malignancy, including lymphoma and leukemia, except malignant neoplasm of skin	C00.x - C26.x, C30.x - C34.x, C37.x - C41.x, C43.x, C45.x - C58.x, C60.x - C76.x, C81.x - C85.x, C88.x, C90.x - C97.x
Moderate or severe liver disease	I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5, K76.6, K76.7
Metastatic solid tumor	C77.x - C80.x
AIDS/HIV	B20.x - B22.x, B24.x

Elixhauser comorbidity index	**ICD-10 codes**
Congestive heart failure	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5 - I42.9, I43.x, I50.x, P29.0
Cardiac arrhythmias	I44.1 - I44.3, I45.6, I45.9, I47.x - I49.x, R00.0, R00.1, R00.8, T82.1, Z45.0, Z95.0
Valvular disease	A52.0, I05.x - I08.x, I09.1, I09.8, I34.x - I39.x, Q23.0 - Q23.3, Z95.2 - Z95.4
Pulmonary circulation disorders	I26.x, I27.x, I28.0, I28.8, I28.9
Peripheral vascular disorders	I70.x, I71.x, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9

Hypertension (uncomplicated)	I10.x
Hypertension (complicated)	I11.x - I13.x, I15.x
Paralysis	G04.1, G11.4, G80.1, G80.2, G81.x, G82.x, G83.0 - G83.4, G83.9
Other neurological disorders	G10.x - G13.x, G20.x - G22.x, G25.4, G25.5, G31.2, G31.8, G31.9, G32.x, G35.x - G37.x, G40.x, G41.x, G93.1, G93.4, R47.0, R56.x
Chronic pulmonary disease	I27.8, I27.9, J40.x - J47.x, J60.x - J67.x, J68.4, J70.1, J70.3
Diabetes, uncomplicated	E10.0, E10.1, E10.9, E11.0, E11.1, E11.9, E12.0, E12.1, E12.9, E13.0, E13.1, E13.9, E14.0, E14.1, E14.9
Diabetes, complicated	E10.2 - E10.8, E11.2 - E11.8, E12.2 - E12.8, E13.2 - E13.8, E14.2 - E14.8
Hypothyroidism	E00.x - E03.x, E89.0
Renal failure	I12.0, I13.1, N18.x, N19.x, N25.0, Z49.0 - Z49.2, Z94.0, Z99.2
Liver disease	B18.x, I85.x, I86.4, I98.2, K70.x, K71.1, K71.3 - K71.5, K71.7, K72.x - K74.x, K76.0, K76.2 - K76.9, Z94.4
Peptic ulcer disease, excluding bleeding	K25.7, K25.9, K26.7, K26.9, K27.7, K27.9, K28.7, K28.9
AIDS/HIV	B20.x - B22.x, B24.x
Lymphoma	C81.x - C85.x, C88.x, C96.x, C90.0, C90.2
Metastatic cancer	C77.x - C80.x
Solid tumor without metastasis	C00.x - C26.x, C30.x - C34.x, C37.x - C41.x, C43.x, C45.x - C58.x, C60.x - C76.x, C97.x
Rheumatoid arthritis/collagen vascular diseases	L94.0, L94.1, L94.3, M05.x, M06.x, M08.x, M12.0, M12.3, M30.x, M31.0 - M31.3, M32.x - M35.x, M45.x, M46.1, M46.8, M46.
Coagulopathy	D65 - D68.x, D69.1, D69.3 - D69.6
Obesity	E66.x
Weight loss	E40.x - E46.x, R63.4, R64
Fluid and electrolyte disorders	E22.2, E86.x, E87.x
Blood loss anaemia	D50.0
Deficiency anaemia	D50.8, D50.9, D51.x - D53.x
Alcohol abuse	F10, E52, G62.1, I42.6, K29.2, K70.0, K70.3, K70.9, T51.x, Z50.2, Z71.4, Z72.1
Drug abuse	F11.x - F16.x, F18.x, F19.x, Z71.5, Z72.2
Psychoses	F20.x, F22.x - F25.x, F28.x, F29.x, F30.2, F31.2, F31.5
Depression	F20.4, F31.3 - F31.5, F32.x, F33.x, F34.1, F41.2, F43.2

.x means all the sub-codes from the main ICD-10 code

Table S2 Comorbidity index of Charlson, van Walraven and Swiss weights in the total study population

Parameters	Alive cohort (%)	Mortality cohort (%)	SMD
Charlson index (Charlson weights)			1.277
0	3,642,650 (61.2)	17,465 (12.2)	
1–4	1,922,213 (32.3)	72,876 (51.1)	
>=5	387,142 (6.5)	52,326 (36.7)	
Elixhauser index (van Walraven weights)			1.303
<0	376,560 (6.3)	1,216 (0.9)	
0	3,236,440 (54.4)	14,311 (10.0)	
1–4	546,349 (9.2)	9,300 (6.5)	
>=5	1,792,656 (30.1)	117,840 (82.6)	
Elixhauser index (Swiss weights)			1.322
<0	1,154,573 (19.4)	4,946 (3.5)	
0	2,561,720 (43.0)	11,607 (8.1)	
1–4	426,880 (7.2)	6,121 (4.3)	
>=5	1,808,832 (30.4)	119,993 (84.1)	

Abbreviations: SMD, standardized mean difference between alive and mortality cohort

Table S3 Prevalence of Charlson comorbidities from the total population and Swiss derivation sample

Charlson comorbidities	Total population	Swiss derivation sample		
	N (%)	Alive cohort (%)	Mortality cohort (%)	SMD
	6,094,672 (100)	2,975,887 (97.7)	71,449 (2.3)	
Myocardial infarction	195,981 (3.2)	91,930 (3.1)	6,138 (8.6)	0.236
Congestive heart failure	360,599 (5.9)	163,685 (5.5)	16,333 (22.9)	0.514
Peripheral vascular disease	294,919 (4.8)	141,051 (4.7)	6,912 (9.7)	0.192
Cerebrovascular disease	254,526 (4.2)	118,841 (4.0)	8,465 (11.8)	0.294
Dementia	189,091 (3.1)	89,536 (3.0)	5,078 (7.1)	0.188
Chronic pulmonary disease	358,407 (5.9)	170,770 (5.7)	8,269 (11.6)	0.209
Rheumatic disease	83,179 (1.4)	40,366 (1.4)	1,061 (1.5)	0.011
Peptic ulcer disease	34,831 (0.6)	16,305 (0.5)	1,144 (1.6)	0.102
Mild liver disease	983,70 (1.6)	46,028 (1.5)	3,203 (4.5)	0.172
Diabetes without chronic complication	523,029 (8.6)	253,022 (8.5)	9,333 (13.1)	0.147
Diabetes with chronic complication	123,336 (2.0)	59,298 (2.0)	2,492 (3.5)	0.092
Hemiplegia or paraplegia	133,663 (2.2)	61,546 (2.1)	5,153 (7.2)	0.246
Renal disease	618,117 (10.1)	289,180 (9.7)	20,528 (28.7)	0.497
Any malignancy, including lymphoma and leukemia, except malignant neoplasm of skin	658,483 (10.8)	302,801 (10.2)	26,900 (37.6)	0.680
Moderate or severe liver disease	29,538 (0.5)	12,442 (0.4)	2,331 (3.3)	0.213
Metastatic solid tumor	276,426 (4.5)	119,667 (4.0)	18,907 (26.5)	0.657
AIDS/HIV	4,894 (0.1)	2,300 (0.1)	85 (0.1)	0.013

Abbreviations: SMD, standardized mean difference between alive and mortality cohort; The total cohort percentages can exceed 100%, as each admission contributes to one or more comorbidities.

Table S4 Prevalence of Elixhauser comorbidities from the total and derivation sample

Elixhauser comorbidities	Total population	Swiss derivation sample		
	N (%)	Alive cohort (%)	Mortality cohort (%)	SMD
	6,094,672 (100)	2,975,887 (97.7)	71,449 (2.3)	
Congestive heart failure	360,599 (5.9)	163,685 (5.5)	16,333 (22.9)	0.514
Cardiac arrhythmias	724,927 (11.9)	341,280 (11.5)	20,754 (29.0)	0.448
Valvular disease	248,362 (4.1)	117,450 (3.9)	6,568 (9.2)	0.213
Pulmonary circulation disorders	116,734 (1.9)	53,292 (1.8)	4,813 (6.7)	0.247
Peripheral vascular disorders	294,919 (4.8)	141,051 (4.7)	6,912 (9.7)	0.192
Hypertension (uncomplicated)	1,383,747 (22.7)	676,609 (22.7)	15,692 (22.0)	0.019
Hypertension (complicated)	459,645 (7.5)	218,656 (7.3)	11,003 (15.4)	0.256
Paralysis	133,663 (2.2)	61,546 (2.1)	5,153 (7.2)	0.246
Other neurological disorders	256,353 (4.2)	120,045 (4.0)	8,011 (11.2)	0.273
Chronic pulmonary disease	358,407 (5.9)	170,770 (5.7)	8,269 (11.6)	0.209
Diabetes, uncomplicated	508,108 (8.3)	245,817 (8.3)	9,059 (12.7)	0.145
Diabetes, complicated	137,602 (2.3)	66,161 (2.2)	2,763 (3.9)	0.096
Hypothyroidism	258,542 (4.2)	126,062 (4.2)	3,454 (4.8)	0.029
Renal failure	617,834 (10.1)	289,047 (9.7)	20,526 (28.7)	0.497
Liver disease	111,396 (1.8)	49,916 (1.7)	5,822 (8.1)	0.303
Peptic ulcer disease, excluding bleeding	12,012 (0.2)	5,808 (0.2)	258 (0.4)	0.032
AIDS/HIV	4,894 (0.1)	2,300 (0.1)	85 (0.1)	0.013
Lymphoma	53,679 (0.9)	25,049 (0.8)	1,759 (2.5)	0.127
Metastatic cancer	276,426 (4.5)	119,667 (4.0)	18,907 (26.5)	0.657
Solid tumour without metastasis	583,683 (9.6)	268,298 (9.0)	24,046 (33.7)	0.631
Rheumatoid arthritis/collagen vascular diseases	97,335 (1.6)	47,305 (1.6)	1,254 (1.8)	0.013
Coagulopathy	200,102 (3.3)	90,551 (3.0)	9,528 (13.3)	0.382
Obesity	137,896 (2.3)	68,155 (2.3)	1,011 (1.4)	0.065
Weight loss	216,541 (3.6)	98,545 (3.3)	9,527 (13.3)	0.369
Fluid and electrolyte disorders	550,094 (9.0)	257,618 (8.7)	17,440 (24.4)	0.434
Blood loss anaemia	40,893 (0.7)	19,759 (0.7)	685 (1.0)	0.033

Deficiency anaemia	148305 (2.4)	72,290 (2.4)	1,886 (2.6)	0.013
Alcohol abuse	199964 (3.3)	96,708 (3.2)	3,086 (4.3)	0.056
Drug abuse	77640 (1.3)	38,044 (1.3)	583 (0.8)	0.045
Psychoses	59909 (1.0)	29,598 (1.0)	404 (0.6)	0.049
Depression	355765 (5.8)	173,898 (5.8)	3,715 (5.2)	0.028

Abbreviations: SMD, standardized mean difference between alive and mortality cohort; The total cohort percentages can exceed 100%, as each admission contributes to one or more comorbidities.

Table S5 Performance measures of the base, Charlson, van Walraven and Swiss weights models for in-hospital mortality in derivation, validation and all cases

	C Statistic (95% CI)								
	Derivation group	Validation groups						Validation group	All cases
	n = 3,047,336	n_1 = 491,962	n_2 = 496,684	n_3 = 504,741	n_4 = 514,267	n_5= 520,277	n_6 = 519,405	n = 3,047,336	N= 6,094,672
	All years (2012–2017)	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	All years (2012–2017)	All years (2012–2017)
Base model	0.757 (0.755–0.759)	0.758 (0.752–0.759)	0.758 (0.754–0.762)	0.756 (0.753–0.760)	0.752 (0.748–0.756)	0.756 (0.752–0.759)	0.750 (0.746–0.754)	0.754 (0.753–0.754)	0.757 (0.755–0.759)
Charlson weights model	0.850 (0.847–0.851)	0.849 (0.846–0.852)	0.852 (0.849–0.855)	0.854 (0.851–0.857)	0.849 (0.846–0.852)	0.854 (0.851–0.857)	0.844 (0.841–0.847)	0.849 (0.848–0.851)	0.850 (0.849–0.851)
VW weights model	0.863 (0.862–0.864)	0.862 (0.859–0.865)	0.866 (0.863–0.869)	0.867 (0.864–0.869)	0.863 (0.860–0.866)	0.869 (0.866–0.872)	0.862 (0.859–0.864)	0.863 (0.862–0.864)	0.863 (0.862–0.864)
Swiss weights model	0.867 (0.865–0.868)	0.865 (0.862–0.868)	0.869 (0.866–0.871)	0.871 (0.868–0.873)	0.866 (0.863–0.869)	0.872 (0.869–0.874)	0.865 (0.862–0.867)	0.866 (0.865-0.867)	0.867 (0.865–0.868)

Abbreviations: VW van Walraven, CI confidence interval
Base model: age groups, sex, hospital types
Charlson weights model: base and Charlson weights
VW weights model: base and Elixhauser/ van Walraven weights
Swiss weights model: base and Elixhauser/ Swiss weights
Note: Predicted probability of null model (c-statistic) is 0.50 of dying, which reflect the mean value of 2.3% from our data.

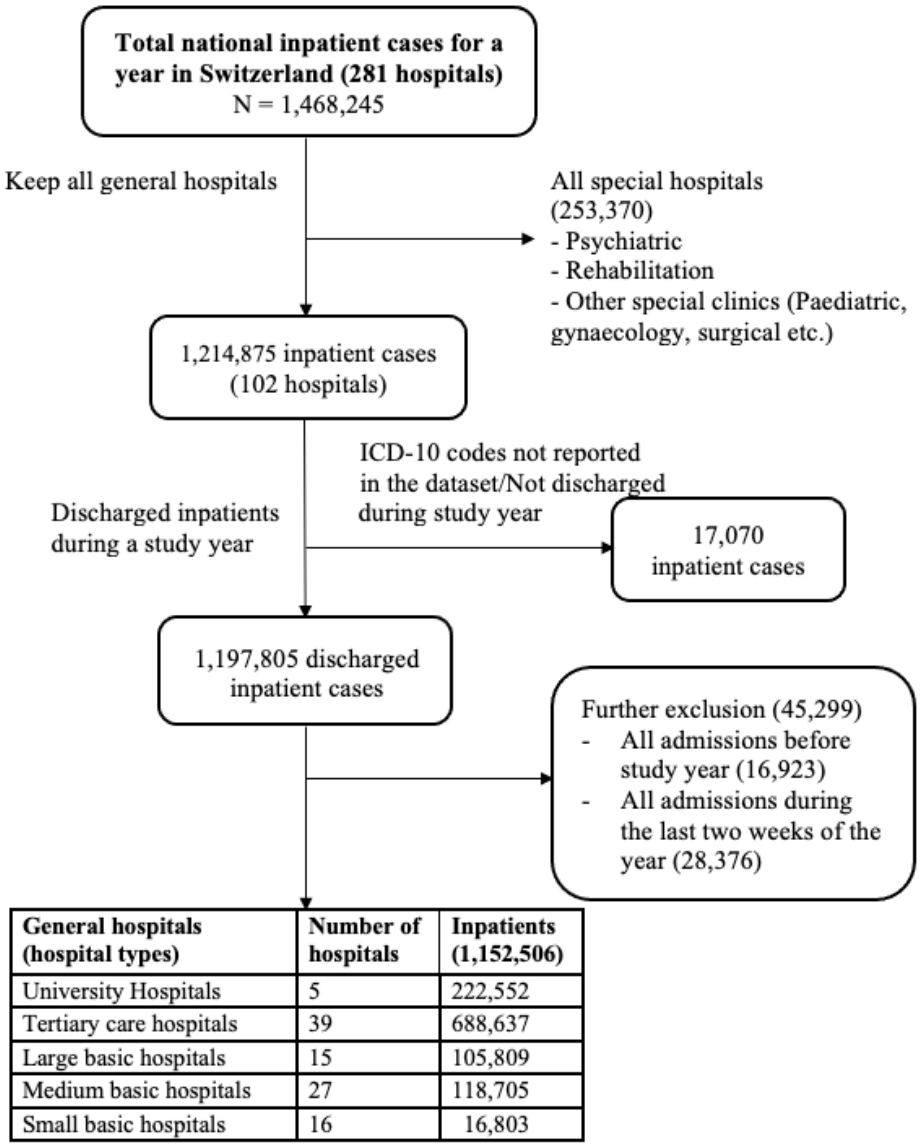
Table S6. The predicted and observed risk of patients from the highest percentage from the predicted values

Highest percentage from the predicted values from the derivation sample				
	1%	2%	5%	10%
Population-based weighting model (pre)	0.332	0.260	0.176	0.124
Population-based weighting model (obs)	0.209	0.182	0.147	0.120
VW weighting model (pre)	0.319	0.251	0.172	0.123
VW weighting model (obs)	0.204	0.181	0.137	0.119
Charlson weighting model (pre)	0.293	0.231	0.160	0.116
Charlson weighting model (obs)	0.185	0.167	0.137	0.112
Base model (pre)	0.110	0.099	0.085	0.072
Base model (obs)	0.107	0.094	0.084	0.072

Pre: mean predicted risk of dying, obs: mean observed risk of dying,
Note: Top 1% of admissions have higher risk of death.

Causal effect of capacity utilization on in-hospital mortality: evidence of safety tipping points

Online supplementary materials



General hospitals (hospital types)	Number of hospitals	Inpatients (1,152,506)
University Hospitals	5	222,552
Tertiary care hospitals	39	688,637
Large basic hospitals	15	105,809
Medium basic hospitals	27	118,705
Small basic hospitals	16	16,803

Figure S1. Flow diagram of the study participants from the study year for the analysis

Table S1. Description of the study variables

Variables	Short description
Individual-level	
Case identifier	Unique anonymous numeric identifier for each case
Age	Age of patient in five-year groups
Sex	Patient's sex (male or female)
Hospital identifier	Unique anonymous numeric identifier for each hospital
Hospital types	Hospital FSO classification code, Five classifications: university hospitals, tertiary care hospitals, large basic hospitals, medium basic hospitals and small basic hospitals
Admission date	Patient's date of hospital admission
Discharge date	Patient's date of discharge from hospital
Primary Diagnosis	ICD–10 GM primary patient diagnosis code (up to 2)
Secondary diagnosis	ICD–10 GM secondary patient diagnosis code (up to 25)
Procedure code	CHOP code for procedure taken for the patient
14–day in-hospital mortality	Death during the first 14 days of hospital stay (death or alive)
Measured variables	
PCCL (0–4; no severity to very high severity)	Patient's individual severity/complexity was measured using SwissDRG version 6 for batch grouping using diagnosis and procedural codes, using online grouping system https://grouper.swissdrg.org/swissdrg/batch
Date related variables	Using the admission date of the study year, time-related variables were created as necessary for the analysis (e.g., Days of the week, weekdays vs weekend)
Length of stay (LOS)	Discharge data minus admission date (in days)
Elixhauser comorbidities (31 comorbidities)	Number of Elixhauser comorbidities count (0, 1, 2,..15) Number of Elixhauser comorbidities with Swiss comorbidity weights,[24] classified as (<0, 0, >0-4, >4)
Hospital level	
Daily capacity utilization	Percentage of inpatients per day from the maximum daily inpatients (via admission and discharge dates)
Daily patient turnover	Percentage of admitted and discharged patients from the actual inpatients per day (via admission and discharge dates)
Daily PCCL	PCCL value for each individual was transformed to average daily PCCL value per day using admission and discharge dates
Unmeasured variables	e.g., variation in hospital leadership, staffing (nurses or doctors) might influence confounders, exposures and outcomes

FSO Federal Statistics Office; *ICD-10 GM* International Classification of Diseases-version 10 German Modification; *CHOP* Swiss Operation Classification; *PCCL* Patient Clinical Complexity Level

Building causal DAGs for time-varying variables to estimate causal effect of capacity utilization on in-hospital mortality

The literature indicates a clear relationship between capacity utilization (exposure) and in-hospital mortality (outcome) [1,2]. Problematically, though, capacity utilization and mortality could be confounded by disease severity (patient clinical complexity level (PCCL)) and patient turnover. I.e., these variables can open up backdoor paths to apparent causality. To block these paths, it is necessary to identify and adjust for confounders.

To explore potential confounding relationships, we used DAGs with time-fixed variables (Figure S2 A). DAGs may clarify relationships between selected variables in a cross-sectional study. They can also be used for time-varying studies such as this one, but multiple measurement points have to be included.

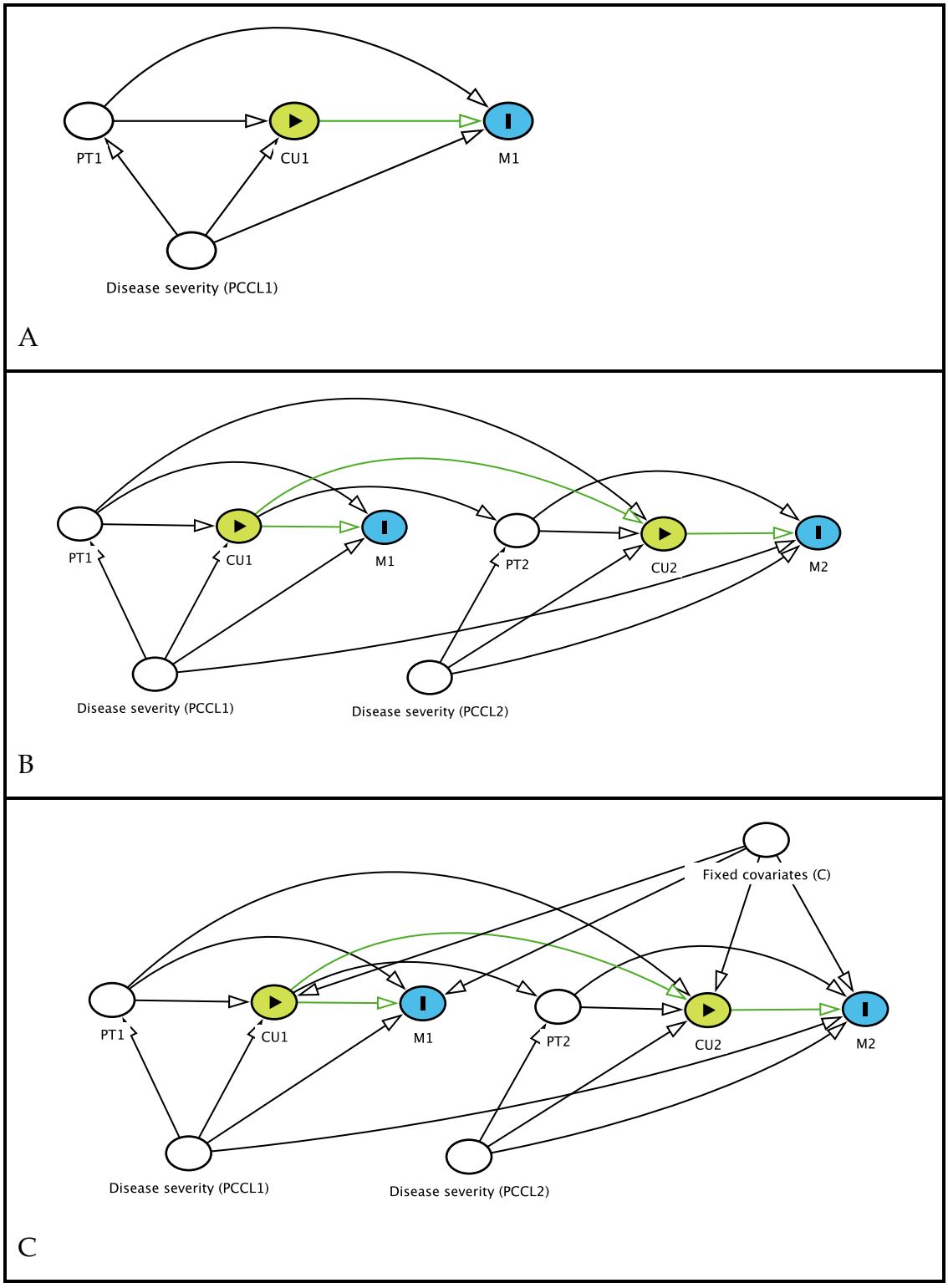
To explore time-varying variables and limit our DAGs' visual complexity, we used two consecutive days as measurement points (day one (baseline): admission day; day two: the following day) (Figure 1B). Assuming that the same relationships would apply at later time points, we drew paths from the confounders to exposure and outcome for each day of the patient's stay, i.e., every day's capacity utilization and mortality figures are influenced by daily patient turnover and daily PCCL values.

Additionally, day-one capacity utilization influences day-two patient turnover. For example, if capacity utilization is high on day one, the hospital may admit fewer and/or discharge more patients on day two; prospective admissions' PCCL scores can also influence both admissions and discharges. Therefore, as Figure 1B shows, we placed an arrow from CU1 to PT2 to indicate feedback between the treatment and the confounder, i.e., TCF. As our DAG considered measurements for all of our dataset's potential confounders, we were able to estimate the hypothesized causal effect simply by implementing an adjusted analysis.

Again, the given DAG is a valid causal model only if it includes every potential confounder. To account for additional individual covariates (e.g., age, sex, Swiss comorbidity weights, weekday/weekend) we built a causal DAG (see Figure 1C) including all additional time-fixed confounders for which adjustment was necessary. This step allowed us to estimate the total causal effect of capacity utilization on in-hospital mortality.

Additionally, we considered the possible relevance of unmeasured variables not included in our dataset (e.g., supportive leadership or staffing). These may also influence the measured variables, e.g., staffing may affect both patient turnover[3] and in-hospital mortality.[4] In this condition, feedback (TCF) between day-one capacity utilization and day-two patient turnover persists. That is, the unmeasured variables (U) leave an open back-door path (CU1-PT2-U-M1) (Figure 1D). Thus, this condition could lead to bias via two pathways: 1. Adjusting for patient turnover (PT2) to block the confounding CU2–PT2–M2 path will open another (through CU1, PT2, U, and M2) by conditioning on a collider; or 2. If we do not adjust for patient turnover (PT2) we leave a back door path open through CU2-PT2-M2.

Traditional methods (e.g., stratification, outcome regression) cannot adjust for both of these confounders at the same time [5,6]. Therefore, innovative methods to handle time-varying variables and TCF are necessary. We chose J. Robins' proposed G-methods. In addition to G-computation [7] and g-estimation for Structural Nested Models, [8] these include the use of inverse probability of treatment weighting (IPTW) for Marginal Structural Models (MSMs) [9].


PT1
CU1
M1
Disease severity (PCCL1)
A
PT1
CU1
M1
PT2
CU2
M2
Disease severity (PCCL1)
Disease severity (PCCL2)
B
Fixed covariates (C)
PT1
CU1
M1
PT2
CU2
M2
Disease severity (PCCL1)
Disease severity (PCCL2)
C

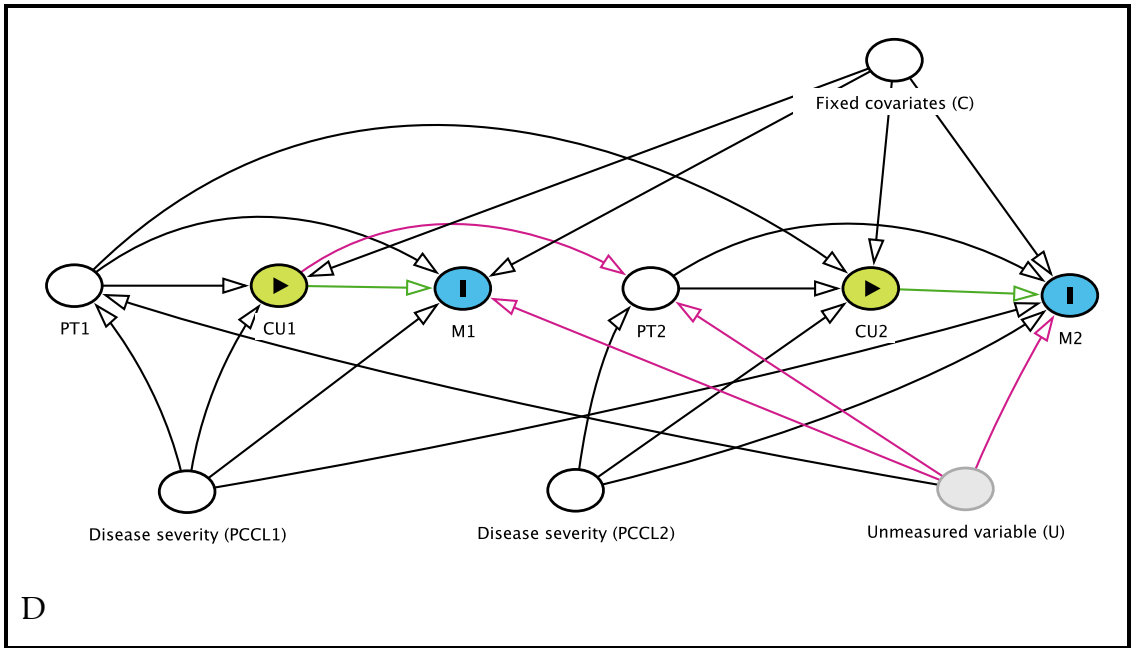


Figure 2: Causal directed acyclic graphs (DAGs) showing causal effects of capacity utilization (CU) on in-hospital mortality (M). Note: the arrows are limited to prevent overcrowding. Green represents causal paths; black represents adjusted (blocked) paths; and pink represents biasing paths.
A: Causal DAG for time-fixed confounders, exposure and outcome (PT1, PCCL1, CU1, M1: respectively, patient turnover, patient clinical complexity level, capacity utilization, and mortality at fixed-day/baseline).
B: Causal DAG for time-varying confounders, exposure and outcome at day one (PT1, PCCL1, CU1, M1) and day two (PT2, PCC2, CU2, M2).
C: Causal DAG for time-varying exposure, confounders and outcome at day one and day two with other fixed covariates (C) (age, sex, comorbidity weights, weeks).
D: Causal DAG for time-varying exposure, confounders and outcome at day one and day two with other fixed covariates (C) and unmeasured variables (U) (e.g., leadership, staffing).

References (Building causal DAGs)

1. Boden D, Agarwal A, Hussain T, Martin S, Radford N, Riyat M, et al. Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital. Emergency Medicine Journal. 2016;33(2):85-90.
2. Kuntz L, Mennicken R, Scholtes S. Stress on the ward: Evidence of safety tipping points in hospitals. Management Science. 2014;61(4):754-771.
3. Park SH, Blegen MA, Spetz J, Chapman SA, De Groot H. Patient turnover and the relationship between nurse staffing and patient outcomes. Research in nursing & health. 2012;35(3):277-288.
4. Griffiths P, Maruotti A, Saucedo AR, Redfern OC, Ball JE, Briggs J, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ quality & safety. 2019;28(8):609-617.
5. Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.
6. McGrath S, Lin V, Zhang Z, Petito LC, Logan RW, Hernán MA, et al. gfoRmula: An R Package for Estimating the Effects of Sustained Treatment Strategies via the Parametric g-formula. Patterns. 2020:100008.
7. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. Mathematical modelling. 1986;7(9-12):1393-1512.
8. Vansteelandt S, Joffe M. Structural nested models and G-estimation: the partially realized promise. Statistical Science. 2014;29(4):707-731.
9. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Statistical models in epidemiology, the environment, and clinical trials. Springer; 2000:95-133.

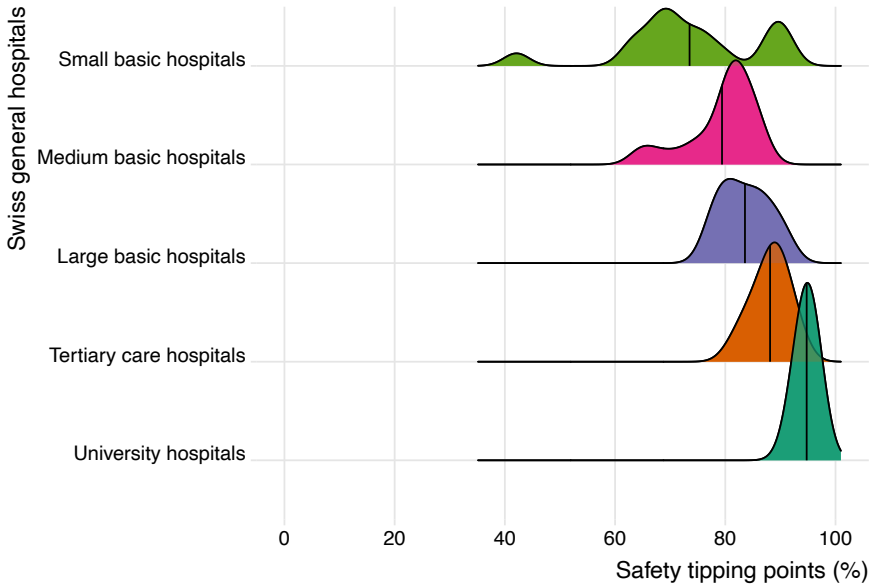


Figure S2. Distribution of safety tipping points and vertical line is an average tipping point across Swiss hospital types

Table S2. Patients' treatment-exposure day distribution by hospital type

Parameters	**Total exposure days**	**Exposure (capacity utilization) days for below and above tipping points**		**SMD**
		< 85th percentile	**≥ 85th percentile**	
Total	6,867,658	5,588,637 (81.4)	1,279,021 (18.6)	
Hospital types				0.075
- University (level 1)	1,474,338	1,216,182 (82.5)	258,156 (17.5)	
- Tertiary care (level 2)	4,073,652	3,326,103 (81.6)	747,549 (18.4)	
- Large basic (level 3)	579,174	463,261 (80.0)	115,913 (20.0)	
- Medium basic (level 4)	628,306	497,660 (79.2)	130,646 (20.8)	
- Small basic (level 5)	112,188	85,431 (76.1)	26,757 (23.9)	

SMD, standardized mean difference between < 85th percentile and ≥ 85th percentile

Table S3. Sensitivity analysis to evaluate robustness to different specifications of the capacity utilization tipping points in Swiss hospitals, splitting capacity utilization distribution at 75th, 80th, 85th, 90th, and 95th percentile. The table displays total/short-term effect of daily exposure on 14-day in-hospital mortality as estimated with generalized estimating equations (GEEs), for different values of the thresholds defining potential tipping points.

GEE glm models	Odds Ratio (95% CI)
75th Percentile split	1.04 (1.01–1.08)*
80th Percentile split	1.06 (1.02–1.10)**
85th Percentile split	1.08 (1.04–1.12)***
90th Percentile split	1.07 (1.03–1.13)**
95th Percentile split	1.07 (1.01–1.14)*

Significance codes: ***p <0.001, **p <0.01, *p <0.05

Note: The effect estimate on each split was positive with similar effects at and above 85th percentile, indicating that our capacity utilization split was robust.
Time-fixed covariates (age, sex, hospital types, Elixhauser index-Swiss comorbidity weights, and weekdays) affecting both exposure and outcomes were adjusted in each model.

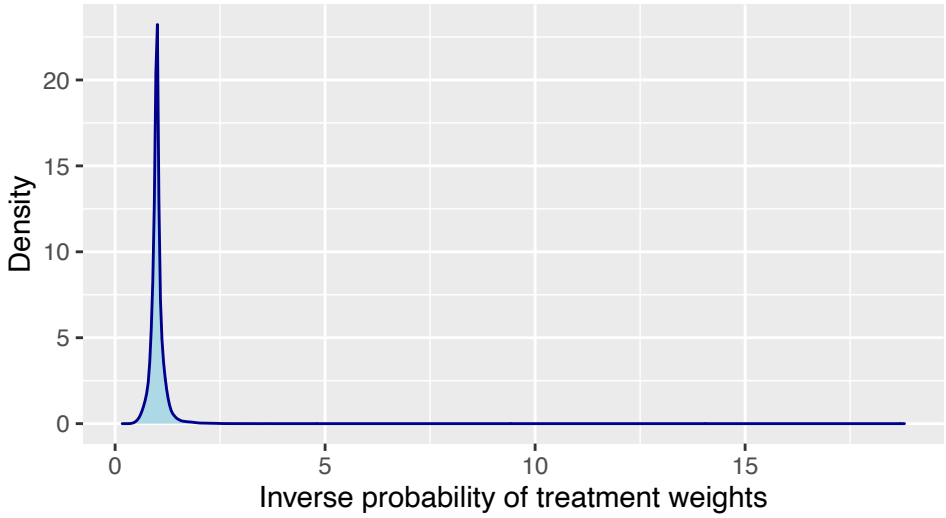


Figure S3. Inverse probability of treatment weight (IPTW) of capacity utilization with two time-varying confounders: patient turnover and patient clinical complexity level (PCCL).

Table S4. Total/short-term effect of time-varying exposure (capacity utilization) on 14-day in-hospital mortality without and with IPTW

	Without IPTW			With IPTW		
	Estimate	p value	Odds Ratio (95% CI)	Estimate	p value	Causal Odds Ratio (95% CI)
Exposure (Capacity Utilization)						
- Below 85th percentile	Reference		1	Reference		1
- 85th percentile and above	0.077	<0.001***	1.08 (1.03 to 1.12)	0.097	<0.001***	1.10 (1.06 to 1.15)
Other adjusted variables						
Weekdays						
- Weekday	Reference			Reference		
- Weekend	0.108	<0.001***	1.11 (1.06 to 1.14)	0.110	<0.001***	1.12 (1.08 to 1.16)
Hospital types						
- University (level 1)	Reference		1	Reference		1
- Tertiary care (level 2)	0.019	0.312	1.02 (0.98 to 1.06)	0.022	0.25	1.02 (0.98 to 1.06)
- Large basic (level 3)	-0.149	<0.001***	0.86 (0.80 to 0.92)	-0.141	<0.001***	0.87 (0.81 to 0.93)
- Medium basic (level 4)	-0.116	<0.001***	0.89 (0.83 to 0.95)	-0.120	<0.001***	0.89 (0.83 to 0.95)
- Small basic (level 5)	0.219	<0.001***	1.24 (1.12 to 1.38)	0.513	<0.001***	1.67 (1.46 to 1.90)
Elixhauser index (Swiss Comorbidity weights)						
- <0	Reference		1	Reference		1
- =0	1.090	<0.001***	2.97 (2.71 to 3.25)	1.090	<0.001***	2.97 (2.71 to 3.26)
- >0 to <5	0.848	<0.001***	2.33 (2.10 to 2.60)	0.866	<0.001***	2.38 (2.13 to 2.64)
- ≥5	2.060	<0.001***	7.82 (7.22 to 8.49)	2.080	<0.001***	8.01 (7.38 to 8.70)
Age (5-year groups)	0.028	<0.001***	1.03 (1.02 to 1.03)	0.028	<0.001***	1.03 (1.02 to 1.03)
Sex						
- Male	Reference		1	Reference		1
- Female	-0.266	<0.001***	0.77 (0.74 to 0.79)	-0.270	<0.001***	0.76 (0.74 to 0.79)

Significance codes: ***p <0.001, **p <0.01, *p <0.05
IPTW: Inverse Probability of Treatment/exposure Weight of Capacity utilization for ≥ 85th percentile for daily varying confounders, patient turnover and PCCL (Patient Clinical Complexity Level)

Table S5. Adjusted causal effect (total/short term and cumulative daily) of time-varying exposure (capacity utilization) on 14-day in-hospital mortality with truncation of the top and bottom 1% of IPTW

	Short term effect with IPTW (MSM with 1% truncation)			Cumulative daily effect with IPTW (MSM with 1% truncation)		
	Estimate	p value	Causal Odds Ratio (95% CI)	Estimate	p value	Causal Odds Ratio (95% CI)
Exposure (Capacity Utilization)						
- <85th percentile	Reference		1	Reference		1
- ≥85th percentile	0.085	<0.001***	1.09 (1.05 to 1.13)	0.012	<0.01**	1.01 (1.01 to 1.03)
Weekdays						
- Weekday	Reference			Reference		
- Weekend	0.106	<0.001***	1.11 (1.08 to 1.16)	0.087	<0.001**	1.09 (1.05 to 1.13)
Hospital types						
- University (level 1)	Reference		1	Reference		1
- Tertiary care (level 2)	0.019	0.31	1.02 (0.98 to 1.06)	0.019	0.31	1.02 (0.98 to 1.06)
- Large basic (level 3)	-0.148	<0.001***	0.86 (0.81 to 0.92)	-0.147	<0.001**	0.86 (0.81 to 0.92)
- Medium basic (level 4)	-0.122	<0.001***	0.88 (0.83 to 0.95)	-0.122	<0.001**	0.89 (0.83 to 0.95)
- Small basic (level 5)	0.275	<0.001***	1.32 (1.18 to 1.47)	0.275	<0.001**	1.32 (1.18 to 1.47)
Elixhauser index (Swiss Comorbidity weights)						
- < 0	Reference		1	Reference		1
- =0	1.090	<0.001***	2.98 (2.72 to 3.27)	1.090	<0.001**	2.98 (2.72 to 3.27)
- >0 to <5	0.863	<0.001***	2.37 (2.13 to 2.64)	0.862	<0.001**	2.37 (2.13 to 2.64)
- ≥5	2.070	<0.001***	7.90 (7.28 to 8.57)	2.070	<0.001**	7.89 (7.38 to 8.70)
Age (5-year groups)	0.028	<0.001***	1.03 (1.02 to 1.03)	0.028	<0.001**	1.03 (1.02 to 1.03)
Sex						
- Male	Reference		1	Reference		1
- Female	-0.270	<0.001***	0.76 (0.74 to 0.79)	-0.270	<0.001**	0.76 (0.74 to 0.79)

Significance codes: ***p <0.001, **p <0.01, *p< 0.05
IPTW: Inverse Probability of Treatment/exposure Weight of Capacity utilization ≥85% for daily varying confounders (patient turnover and PCCL (Patient Clinical Complexity Level).