STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2017–2018

Early detection alerts in sepsis treatment

An empirical study of patient treatment and outcomes

Mara Barschkett (41058)

Abstract

With rising antimicrobial resistance and an increasing number of sepsis incidence, it is essential to examine ways to improve sepsis patient outcomes. The purpose of this study is to evaluate the impact of an early detection sepsis alert on clinicians' behaviour and patient outcomes. During the observation period the alert system has an active and a silent running period. The active alert notifies clinicians when patients are potentially infected with sepsis. During the silent period clinicians are not notified but the system documents when the alert should have been triggered. Using repeated cross-sectional patient-level data, I gauge the effect of the alert system on clinicians' behaviour and patient outcomes employing three different regression forms. This allows me to estimate the differences in the outcome variables between patients that were exposed to the active alert and patients that were not. The results indicate a time trend leading to reduced mortality and length of stay (LOS) in hospital, but further research is needed to disentangle the effect of the alert from the time trend. Hence, this study provides suggestive evidence that the alert system has the potential to reduce mortality and LOS.

Keywords: Alert system, Sepsis, Early detection, Survival analysis, Fixed effects
JEL: D01, D04, D12, I10, I12, I19

Supervisor: Erik Lindqvist
Date submitted: 13/05/2018
Date examined: 28/05/2018
Discussant: Cristina Altomare
Examiner: Mark Sanctuary

Acknowledgements

I am most grateful to my supervisor Erik Lindqvist for his valuable feedback, helpful comments, and guidance throughout this thesis. I would like to thank Imperial College Healthcare NHS Trust Business Intelligence Unit – Clinical Informatics and Analytics for making this thesis possible by granting me access to their data resources. In particular, I would like to thank Erik Mayer and Ben Glampson for continous support and advice throughout this thesis. Furthermore, I would like to thank Mats Kröger, Jonas Nauerz and Hedda Nielsen for being part of a productive feedback group. Lastly, a special thanks to Tobias Pingel and my family, in particular to my parents and sister, without their support this thesis would not have been possible.

Data were accessed via a secure processing environment provided by the Imperial College London Big Data and Analysis Unit (BDAU), data provided by the Trust were de-identified and approved for processing by Imperial Trust Information Governance as part of the Imperial College Clinical returns and data analysis programme. The views expressed in this publication are those of the author and not necessarily those of the Imperial College London, Imperial College Healthcare NHS Trust or the NHS.

Contents

1 Introduction

Sepsis is a major health threat in both the developed and the developing world. Sepsis is defined as "the body's systemic inflammatory response to microbial infection which can cause organ damage, shock, and eventual death" (NHS England, 2015). In the UK, sepsis counts 123,000 incidence per year, whereof approximately 30% are associated with death. Sepsis is thereby the second biggest cause of death (after cardiovascular disease) in the UK (NHS England, 2015). Evidence from the US shows a similar picture: Armen et al. (2016) report 751,000 annual sepsis cases whereof 29% result in death. They expect the number of incidence to grow up to 1,110,000 cases per year by 2020. The mortality rate of sepsis patients is approximately 30% in developed countries and considerably higher in developing countries (45%) (Westphal et al., 2010). In 2017, sepsis was declared a major health care issue by the World Health Organization (WHO). Particularly, increasing antimicrobial resistance leads to clinical unresponsiveness to treatment. The subsequent lack of effective treatments for bacterial infections does not allow to preventing the rapid development of infections to sepsis and septic shock (WHO, 2017). In light of this background, improving existing and developing new treatment methods increases in relevance.

Sepsis also poses a financial burden on health care systems. In 2013, sepsis was the most expensive reason for hospitalisation, accounting for 6.2% of total US hospital costs (Armstrong et al., 2017). For the UK, Padkin et al. (2003) estimate that up to 46% of intensive care unit (ICU) bed days are used by patients with sepsis, where each ICU bed costs around 1,700 GBP per day (Robson and Daniels, 2008).

In the early 2000's the "Surviving Sepsis Campaign" (SSC) was launched in response to a clinical trial suggesting that sepsis outcomes might be improved by care standardisation (Rivers et al., 2001). Care bundles that provide guidelines on how to treat sepsis patients in the first hours after the diagnosis were developed and evaluated in the literature. A number of studies have demonstrated the success of bundle care in terms of both patient outcomes and financial implications (e.g. Barochia et al., 2010; Gao et al., 2005; Levy et al., 2010). However, more recent studies suggest that patient outcomes do not improve for patients that were treated according to the suggested bundle care (e.g. ARISE Investigators and the ANZICS Clinical Trials Group, 2014; Mouncey et al., 2015; Yu et al., 2016). The potential for improving patient outcomes seems exploited once clinicians comply to SSC guidelines. Therefore, other means are necessary to improve the situation. One approach to change the current situation for the better is increasing clinicians' guideline compliance rates which in earlier years were found to be low (e.g. Gao et al., 2005; Nguyen et al., 2007). In addition, there is evidence that earlier detection of sepsis and hence, earlier initiation of treatment,

improves patient outcomes (e.g. Inada-Kim et al., 2017).

The goal of this study is to evaluate an early detection sepsis module, during a pilot testing phase, in a NHS Healthcare Trust in London, UK. The alert system has two functions. First, it notifies clinicians when a patient is potentially infected with sepsis, and second it documents and tracks information on the timing of the alert. During the observation period, the alert has an active and a silent running period. Only during the active running period the alert notifies clinicians about potential infections. In contrast, during the silent running period the alert only documents when an alert should have been triggered. I compare patient outcomes between patients who were exposed to the active alert and patients who were not.

The main research question of this study is thus:

Can sepsis patient outcomes be improved by the early detection alert?

I explain the mechanism through which the alert affects clinicians' behaviour and ultimately patient outcomes with the use of a classical principal-agent contracting problem with asymmetric information: Health outcome of sepsis patients is stochastic (i.e. not only determined by clinicians' behaviour). Due to the asymmetric information distribution and the stochastic outcome, the care provider cannot monitor clinicians' action. The alert system, which saves information on the alert and the clinician's response, is expected to reduce the asymmetric information, and this in turn should lead to higher compliance to the guidelines. Furthermore, the alert system allows to detect infections earlier and therefore initiate the treatment earlier.

Data were accessed via a secure processing environment provided by the Imperial College London Big Data and Analysis Unit (BDAU), data provided by the Trust were de-identified and approved for processing by Imperial Trust Information Governance as part of the Imperial College Clinical returns and data analysis programme. I first assess the *change in behaviour* of clinicians as a response to the alert system. Second, I examine the impact of the alert system on *patient outcomes*. My main model specification is an Ordinary Least Squares (OLS) regression with time and ward fixed effects. This specification allows to identify a causal treatment effect of the alert system on the outcome variables. I expand the analysis by running a Cox regression to evaluate the impact of the alert on total mortality.

In the main specifications, I do not find the alert system to significantly improve *patient outcomes*, namely *mortality* and *length of stay (LOS)* in hospital. However, I identify a time trend leading to reduced *mortality* and *LOS*. The results suggest parts of this time trend may be attributable to the active alert. To confirm and quantify the effect, further investigations of the time trend and its relation to the alert is needed. In addition, I do not

find a significant and robust effect of the alert on *clinicians' behaviour* measured by *time to first test order* and *frequency of test orders*.

With this thesis, I contribute to existing academic knowledge in the following ways: First, to the extent of my knowledge, this evaluation of an early-detection sepsis module is the first of its kind. In the context of other diseases, alerts have been demonstrated to improve patient outcomes (e.g. Lurio et al., 2010). Second, I expand the literature concerning the relation of early detection of sepsis infections and patient outcomes. Third, my paper adds to the literature on clinicians' compliance to sepsis care guidelines. The application of a principal-agent model to explain the impact of an alert system on clinicians' behaviour is novel and may form a promising basis to better understand and monitor clinicians' behaviour.

This study is organised as follows: Section 2 presents background information on sepsis and the introduced early detection sepsis module and specifies my research question. Following, in Section 3, I provide a review of previous literature on sepsis treatment and alert systems. The theoretical framework is presented in Section 4. The data and retained sample used in the study and my hypotheses are outlined in Section 5. In Section 6, I provide evidence on randomized assignment of control and treatment group. Section 7 presents descriptive results for the sample. The empirical strategy used in this study is reviewed in Section 8 and in Section 9 the main results are presented and analysed. This is followed by a discussion including interpretation of the results, policy implications, limitations as well as the validity of the study in Section 10. Finally, Section 11 provides a conclusion of the study.

2 Background

Even though sepsis is a major global health care issue, data on sepsis epidemiology is scarce. The threat of sepsis infections becomes especially clear in light of increasing antimicrobial resistance leading to clinical unresponsiveness to treatment which in turn leads to a lack of effective treatment for bacterial infections. Subsequently, the rapid evolution of infections to sepsis and septic shock cannot be prevented (WHO, 2017). This development poses a burden to infected patients including high mortality rates. Thus, examining new ways of sepsis treatment is essential to counteract this evolution and improve sepsis patient outcomes.

2.1 Sepsis

Sepsis is also known as the "silent killer" as it is hard to detect with its symptoms being similar to less serious diseases (e.g. influenza) (NHS England, 2015). Therefore, it is crucial to

neatly define sepsis, raise awareness, identify patients at risk and train the medical staff to react appropriately when dealing with patients at risk. However, there is no gold standard for diagnosing sepsis, and the definition of (severe) sepsis has been revised multiple times (Mellhammar et al., 2016). The latest definition of sepsis is designated sepsis-3 and defines sepsis as a "life-threatening organ dysfunction caused by a dysregulated host response to infection" (Singer et al., 2016). This definition means when infected with sepsis the body's immune system "goes into overdrive" which leads to multiple reactions of the body (e.g. widespread inflammation, swelling and blood clotting). These reactions can imply a significant decrease in blood pressure leading to a reduction in blood supply to vital organs that consequently lack oxygen. Finally, multiple organ failures and eventual death may follow (NHS England, 2015). Under the latest definition of sepsis, "septic shock" is defined as "a subset of sepsis in which underlying circulatory and cellular metabolism abnormalities are profound enough to substantially increase mortality" (Singer et al., 2016). Sepsis is mostly induced by bacterial infections[1] and therefore treated with antibiotics. The most common bacterial infections that develop into sepsis are "Diseases of the Respiratory System", "Diseases of the Digestive System" and "Diseases of Genitourinary System" (Inada-Kim et al., 2017).

A number of studies point out that the number of registered incidence of sepsis are projected to increase (e.g. Rhee et al., 2015; Tiru et al., 2015). The reasons for this development are multi-fold and, among others, include an aging population with more comorbid and chronic diseases; an increased number of performed medical and surgical interventions; and greater recognition (Levy et al., 2010; NHS England, 2015; Rhee et al., 2014).

2.2 The early detection sepsis module

The sepsis module was introduced in a silent running mode in 2016 where staff were not presented with an alert and as a pilot in several areas in 2017. This enables before and after evaluation. This early detection sepsis module can be categorised as an electronic health records (EHR) system. EHR is defined as an electronic set of patient records that "contain information on a patient's medical history, demographics, laboratory data, medication and other important medical information" (Latha et al., 2012). EHR does not only allow to track patients' health and data electronically but also provides an opportunity to incorporate notifications (or alerts) for clinicians for a range of purposes, including reminders for clinicians to check specific aspects of a patient's medical history to raise awareness that a patient may be at risk of a particular disease.

[1]There are also non-bacterial causes of sepsis, e.g. virus, fungal or parasite infections, that are far less common and responsive to treatment (Inada-Kim et al., 2017).

The introduced alert system has essentially two functions:

1. Notify clinicians when a patient is potentially infected with sepsis.
2. Document and track information on when the alert is triggered, i.e. when a patient is potentially infected.

The sepsis module triggers a clinician alert based on "Systemic Inflammatory Response Syndrome" (SIRS) criteria: Body temperature; respiratory rate; heart rate; glucose level and white cell count. The alert is triggered when either three of the SIRS criteria; or two of these criteria with at least one additional measure indicating organ damage (raised lactate, hypotension, raised bilirubin, raised creatinine) meet a certain threshold. When the alert fires, clinicians are notified with a warning popping up in the electronic health system when accessing the respective patient's record. Since for each action performed on the patient the medical staff needs to access the patient's record, it is likely that they open a patient's record frequently and thus receive the alert shortly after it fired. All patients that stay on the wards where the alert system was installed are exposed to the system.

Besides notifying clinicians when a patient is potentially infected with sepsis, the alert also saves information on the timing of the alert. The timing of the alert gives an indication when a patient developed a potential sepsis infection. After receiving the notification in a patient's record, clinicians need to enter a result description that states a "quick diagnosis", e.g. "suspected sepsis". Matching this information with the record on clinicians' actions performed on a patient provides a detailed overview of the timing of a patient's condition and clinicians' responses.

The installed alert has an *active* and a *silent* running period. During the *silent* running period, no alert is fired, but the system saves information on when it would have fired, i.e. when the measurement values indicate a potential infection. Hence, in the *silent* running period clinicians are not notified. Thus, only the second function (documenting infection time) of the alert is in place but the first function of the alert (notifying the clinician) is not in use. In contrast, during the *active* running period, both functions are activated, i.e. clinicians are notified and information is documented. The time of the (silent) alert can be used as a reference to define infection time as this information is available during both, the silent and the active running period. On all wards where the alert is installed there was a silent running period in the beginning. However, only on some wards the alert has been turned active yet and the "live"-dates vary across wards. The selection of wards for this study and the timing of the live-date are more extensively discussed in Section 5.2.

2.3 Research question

The aim of this study is to evaluate the introduced alert system on basis of *clinicians' behaviour* and *patient outcomes*. In a first step, I assess the change in behaviour of the medical staff triggered by the alert. Subsequently, I evaluate in how far patient outcomes respond to the potential change in behaviour of the medical staff triggered by the introduction of the alert system. Hence, the overall research question can be formulated as follows:

Can sepsis patient outcomes be improved by the early detection alert?

My two-step strategy to answer this question leads to the following two "sub-research questions", which are visualised in Figure 1:

1. *How does the alert affect clinicians' behaviour in treating sepsis patients?*
2. *How does the alert affect patient outcomes through a change in clinicians' behaviour?*

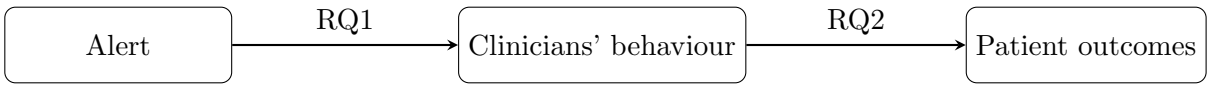


Figure 1: Research question (RQ)

To understand the impact of the alert and the mechanisms through which the alert triggers a change in behaviour of clinicians as well as to be able to infer hypotheses on its effect, I put the situation with and without the alert system into a theoretical framework presented in Section 4.

3 Literature review

The two functions of the alert system, namely notifying clinicians and documenting information on patients' conditions and clinicians' actions, combine two related fields in the literature: First, notifying clinicians leads to earlier detection of potential infections which has been identified as crucial in the literature on sepsis treatment.

Second, tracking clinicians' response to the alert allows to observe whether they follow prescribed guidelines and I thereby expand the literature on clinicians' compliance to care guidelines. There is a large body of research concerning the relatively low observed compliance rates to standardised care protocols of clinicians. This study tries to provide additional evidence on how alert systems can potentially mitigate this problem. In addition, the application of a principal-agent model to explain the impact of an alert system on clinicians' behaviour, presented in Section 4, is novel and therefore may form a promising basis to better understand and monitor clinicians' behaviour.

Lastly, to the extent of my knowledge, this evaluation of an early-detection sepsis module is the first of its kind and provides a better understanding of the system's impact on patient outcomes. However, there is literature on alert systems in other contexts where this study adds to.

Thus, this study builds a bridge between the literature on early detection of sepsis, care guideline compliance in sepsis treatment and alert systems in relation to clinicians' behaviour and patient outcomes. These three fields in the literature are carefully reviewed in the following sections.

3.1 Early detection in sepsis treatment

Recent literature suggests early recognition/early initiation of antibiotic treatment as one way to tackle high mortality rates and generally undesired patient outcomes for sepsis patients (e.g. Inada-Kim et al., 2017; NHS England, 2015). Dellinger et al. (2013) add, the initiation of antimicrobial therapy within an hour of presentation with severe sepsis and septic shock is crucial in surviving sepsis.

The hypothesis that early detection and hence early initiation of treatment improves patient outcomes is also supported by empirical findings. Kumar et al. (2006) point out the onset of hypotension (low blood pressure) is an indicator for sepsis and a critical marker of increased mortality in patients with sepsis. They examine the relationship between the delay in initiation of effective antimicrobial therapy from initial onset of recurrent or persistent hypotension and survival in septic shock. Their results suggest first, the existence of substantial delays in delivery of effective antimicrobials to patients with septic shock. Second, they find that over the first six hours after onset of recurrent or persistent hypotension, each hour of delay in initiation of effective antimicrobial therapy is associated with a mean decrease in survival of 7.6%, whereas delays exceeding six hours are even more harmful. These results are in line with Larché et al. (2003) who demonstrate that a delay of >2 hours is associated with an increased mortality in critically ill cancer patients with septic shock. In addition, Houck et al. (2004) as well as both Iregui et al. (2002) and Mathevon et al. (2002) suggest an association between delayed antimicrobial therapy and the mortality rate for community acquired pneunomia and nosocomial pneumonia, respectively.

3.2 Care guidelines

In response to a clinical trial suggesting sepsis outcomes might be improved by standardising care (Rivers et al., 2001), in 2002, the "Surviving Sepsis Campaign" (SSC) was launched by the European Society of Intensive Care Medicine, the International Sepsis Forum and

the Society of Critical Care Medicine. SSC contains evidence-based guidelines for the management of severe sepsis and septic shock developed through a formal and transparent process. Among other things, the recommendations include the implementation of a bundle strategy for the management of septic patients (Dellinger et al., 2013; Levy et al., 2010; Westphal et al., 2010). A bundle includes certain steps to be taken by the medical staff within a targeted timeframe, i.e. a "care bundle is a group of interventions or elements of care related to a specific disease process or an invasive procedure that, when implemented together, produce a better outcome for the patient than if elements were implemented individually" (Robson and Daniels, 2008). The care bundles include among others, the precept to obtain blood cultures, to administer broad spectrum antibiotics, and measure and monitor hypotension (Surviving Sepsis Campaign, 2016).

Evidence on whether the SSC bundles reduce mortality rates and LOS is, however, controversial. Some studies show the implementation of bundled care significantly improves patient outcomes. However, observed compliance rates are relatively low, reducing the success of the guidelines. In contrast, other studies suggest there are no significant differences in outcomes for patients that were treated according to the recommended bundles compared to the control group. The efficacy of the care bundles, however, forms the basis for the argumentation that higher compliance rates have the potential to improve patient outcomes.

Evidence on success of bundle-care

The first study evaluating the success of bundle care was conducted by Rivers et al. (2001). They assess the efficacy of the application of bundle care in a US hospital and find significant improvements of patients outcomes such as a reduced in-hospital mortality rate for the group assigned to the bundle care compared to the control group (30.5% vs. 46.5%). Similarly, Gao et al. (2005) show that following the guidelines leads to a reduced hospital mortality rate of about 50% in two UK hospitals. The positive impact of bundled care on patient outcomes is further confirmed by Barochia et al. (2010) and Chou et al. (2014).

Evidence on decreasing success of bundle-care

However, more recent studies show a different picture. Mouncey et al. (2015) run a randomised trial in 56 NHS hospitals across the UK to assess whether the bundled care suggested by Rivers et al. (2001) improves patient outcomes. They do not find any significant improvements in patient outcomes of patients treated according to the bundle care. Similarly, ProCESS Investigators (2014) and ARISE Investigators and the ANZICS Clinical Trials Group (2014) (in the US and Australasia respectively) find similar results. Yu et al. (2016) run a meta-analysis of trials evaluating bundle care and find neither a significant decrease in the mortality rate for treated patients compared to the control group nor dif-

ferences in secondary outcomes (e.g. length of hospital stay). One explanation for the decreasing success of bundle care in recent years is that the techniques used in usual resuscitation have evolved over the past 15 years and the mortality rate for septic patients has decreased since the landmark paper of Rivers et al. (2001), probably partly due to the introduction of standardised protocols. Park et al. (2017) confirm this conclusion. They find bundled care is associated with lower mortality rates when the mortality rate of the usual care group is greater than 30%, but no effect when the mortality rate in the usual care group is less than 30%.

Compliance rate

Gao et al. (2005) find in their study initially low compliance rates to bundled care but observe an increase over the course of the study period of two years. Nguyen et al. (2007) observe a similar pattern in the US: They recognise a progressive improvement in bundle compliance over the study period. In this study, the medical staff was not only asked to comply to the introduced bundles but compliance was also measured and the medical staff received feedback. Measuring compliance and giving feedback is a way to overcome the principal-agent problem between provider and clinician since it decreases the asymmetric information distribution. The application of the principal-agent model to this context is more extensively discussed in Section 4. Chou et al. (2014) confirm the findings of Nguyen et al. (2007) and claim that usually the adherence to new strategies is best directly after the introduction (after a learning curve) and the adherence decreases after some time when no educational refreshing measures are done. Reasons for low compliance rates to guidelines are provided in Section 4.

3.3 Alert systems

Early detection as well as compliance to the care bundle guidelines are identified as critical factors improving patient outcomes. One way to foster this is to install EHR alert systems such as the alert system evaluated in this study.

On the one hand, alerts have the potential to bring relevant information to the clinician, and have been demonstrated to positively impact prescribing behaviour as well as improve aspects of performance (Lurio et al., 2010; Terrell et al., 2009). On the other hand, by documenting the patients' condition matched with the clinicians' actions alert systems may reduce the asymmetric information distribution between clinicians and care provider which in turn should lead to higher compliance to prescribed guidelines. This mechanism is reviewed in detail in Section 4.

However, the performance of alerts has not been carefully evaluated and is likely to differ

in different clinical settings. The impact of IT interventions[2] on screening methods to improve antimicrobial prescribing in hospitals has been recently reviewed in a meta-analysis by Baysari et al. (2016). They find three randomised control trials amongst 45 studies evaluating IT based interventions. Study quality varied and no clear evidence of effect on mortality, length of stay or appropriate antimicrobial prescribing was seen amongst high quality studies. They conclude, the evidence of IT initiated sepsis screening benefits is limited but it has the potential to improve patient outcomes. However, similar to low compliance rates to care guidelines, alerts may be overridden by clinicians, which diminishes their success (Baysari et al., 2017). Therefore, a better understanding of the performance of the algorithm and possible modifications as well as an assessment of its impact on outcomes is needed to inform clear policy recommendations in different clinical settings.

4 Theoretical framework

The effect of an active alert on the clinician's behaviour, and in turn also on patient outcomes, can run through two mechanisms:

1. *Earlier detection leads to earlier treatment which leads to better patient outcomes.* The introduced early-detection module supports clinicians in detecting sepsis at an earlier stage which enables them to begin earlier with the appropriate treatment (suggested by the bundled care). This in turn is expected to lead to better patient outcomes as reasoned in Section 3.1. Hence, I expect that the alert triggers a change in *clinicians' behaviour*, i.e. the time between infection and beginning of treatment is reduced.

2. *The alert system reduces asymmetric information and thereby increases compliance to guidelines.* The alert system saves information on the timing of the alert. Furthermore, information on each action that is performed on the patient is documented. Hence, the alert system allows the care provider to observe clinicians' actions and therefore recognise whether clinicians responded to the alert by performing immediate action as suggested by the bundled care. In the next section, I explain this mechanism in more detail.

The two mechanisms are visualised in Figure 2, which specifies Figure 1 in more detail.

To explain the second mechanism, I model the situation with and without the alert system as a principal-agent contracting problem. The principal-agent model was introduced by Jensen and Meckling (1976) and is related to problems such as moral hazard, hidden action

[2]In light of the increased use of electronic medical record systems, IT interventions such as computerised approval systems and surveillance systems offer support for clinicians to treat their patients appropriately (Baysari et al., 2016)

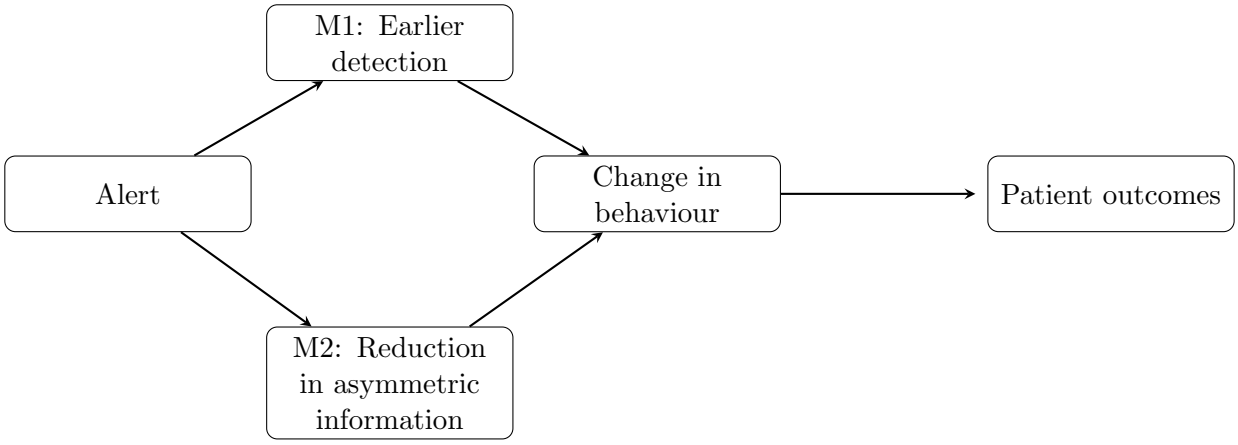


Figure 2: The two mechanisms

and hidden information (Arrow, 1984, pp. 1–24; Arrow, 1963; Pauly, 1968, 1974) which can occur in many different contexts.

4.1 Application of a principal-agent model to the health care context

In the health care context, we observe two principal-agent relationships with asymmetric information: (1) between patient (principal) and clinician (agent); and (2) between insurance provider (principal) and clinician (agent) (Blomqvist, 1991). Hence, the clinician acts in the role of a "double agent", which is visualised in Figure 3.

Figure 3: Clinician in the role of a double agent

The patient (principal) requests actions from the clinician (agent) who has for the patient non-observable information regarding the health status and the required treatment options (since mostly patients lack clinical knowledge). Furthermore, this is a "hidden action" problem, since even with accurate diagnostic information, there is considerable variation in relation between health care and outcomes, i.e. health outcome is stochastic and not only determined by agents' actions (Blomqvist, 1991).

Similarly, the relationship between provider and clinician is usually characterised by asymmetric information since the clinician has information on the actual services needed to perform on the patient whereas the insurance provider can only observe the health outcome which is not deterministic. The insurance provider wants the clinician to economise on the use of health services to reduce costs while providing quality care (Blomqvist, 1991).

4.2 Application to the sepsis context

In the interplay of septic patients/patients at risk, clinicians and the care provider, the clinician acts in the role of a "double agent". However, the introduction of the alert system does not affect the relationship between patient and clinician and is therefore trivial. For simplicity, I focus on the contracting problem between provider and clinician.

According to NHS England (2015) many deaths caused by sepsis are avoidable, clearly indicating the current situation is not optimally solved. One explanation for the unsatisfactory solution could be the contracting problem in the principal-agent relationship. The health care provider expects the clinician to treat the patient in the most cost-efficient way which is suggested by provided guidelines on bundled care. In the British context, the NHS acts as the health care provider. If there were no inefficiencies in this contracting problem, there may be considerably less death cases that could have been avoided.

One example of the contracting problem between clinician and provider is the observed low compliance with standardised guidelines by clinicians. Not complying to guidelines does not necessarily imply selfish preferences of the clinician but can have multiple reasons. Cabana et al. (1999) identified the following barriers to implement clinical guidelines in practice:

- Failure of dissemination strategies (e.g. publication of guidelines does not outreach to majority of clinicians or the format is not practical and therefore not read by clinicians),
- Lack of agreement with one or more recommendations,
- Lack of outcome expectancy (if clinician does not expect interventions to lead to improved patient outcomes, it is unlikely to be implemented),
- Inability to overcome the inertia of previous views and practice or lack of motivation to change,
- External barriers (e.g. guideline-related barriers such as perception that guidelines are inconvenient or difficult to use).

Brown (2002) adds clinicians may also choose to ignore guidelines for non-clinical reasons such as financial incentives or fear of litigation. Some of the listed reasons suggest there is some kind of effort cost associated with following guidelines, e.g. exerting effort to read guidelines or external barriers that are inconvenient to overcome. Hence, the asymmetric information distribution may incentivise clinicians to (partly) not comply to guidelines. Note, however, it is debatable whether guidelines always suggest the best way to treat a patient. For example Brown (2002) points out that the national guidelines for the treatment of adult patients with bacterial meningitis in the Netherlands did not cover the case of

antibiotic resistance suspicion and hence for some patients it proved better to deviate from the guidelines.

However, no matter how non-compliance is motivated, it is the result of the asymmetric information distribution resulting in a hidden action (moral hazard) problem. Assuming the guidelines (care bundles) represent the optimal treatment (a discussion on the success of bundle care is provided in Section 3.2), the provider cannot observe whether clinicians followed the guidelines as health outcome is not deterministic. Overcoming this asymmetric information problem is difficult. One approach to tackle low compliance rates is on the one hand to try to counteract the reasons why clinicians do not comply to guidelines (e.g. improve the guidelines and give them more freedom in applying them to specific cases) and on the other hand to reduce asymmetric information between provider and clinicians. In this paper, I focus on the latter approach: the introduction of an early-detection module reduces the asymmetric information distribution in the contracting problem between provider and clinician (mechanism two). Furthermore, the alert system supports clinicians in detecting sepsis at an earlier stage (mechanism one). In essence, the alert system is expected to improve clinicians' situation (allows them to detect sepsis earlier) as well as the provider's situation who is able to better monitor clinicians' actions. In the next section, I show formally that full information on clinicians' actions allows to introduce a quality-contingent wage scheme that incentivises clinicians to exert more effort. Increased effort includes complying to the care bundles.

4.3 Model

The provider is both interested in patient outcomes and costs of care provided by clinicians to patients. Patient outcomes depend on quality of care provided by clinicians and a stochastic component as reasoned in Section 4.1. The stochastic component cannot be influenced by provider or clinician and is likely not influenced by the introduction of the alert system.

I assume, the principal's utility function depends on quality, which in turn is dependent on the effort level exerted by the agent, and the wage paid to the agent. I further define quality increases linearly in effort, i.e. $q(e) = e$:

$$u_P = q(e) - w = e - w$$

where q is quality, e is effort and w is wage. Furthermore, I presume, the provider is not satisfied with the quality in the initial situation (pre-alert/silent alert).

The agent's utility function depends on the wage they receive and their effort costs:

$$u_A = w - c(e - \underline{e})$$

Suppose the agent's effort cost function takes the following form

$$c(e - \underline{e}) = (e - \underline{e})^2$$

where $\underline{e} > 0$ is the intrinsic effort level, i.e. the effort level agents choose independent of their wage and e is the chosen effort level.

The agent's participation constraint (PC), i.e. the constraint that guarantees the agent does not prefer an outside option (e.g. another employment option), looks as follows:

$$w - c(e - \underline{e}) \geq \underline{u}$$

where $\underline{u}$ is the reservation utility. For simplicity, I set the reservation utility to zero, i.e. $\underline{u} = w - c(e - \underline{e}) = 0$. This assumption implies, in the outside option the agents are paid their reservation wage which is equal to the effort costs and hence gives zero utility.

Based on these assumptions, I present two scenarios to show the effect of the alert on the principal-agent relationship. Scenario 1 represents the baseline case, i.e. before the active alert was introduced and scenario 2 represents the situation after the introduction. In the derivations below I follow Varian (1992, pp. 441–470) and Mas-Colell et al. (1995, pp. 477–506).

Scenario 1: Asymmetric information

I first consider the case before the alert is introduced. In this scenario the monitoring costs for the provider to observe the quality of care which depends on agents' actions are very high. For simplicity, I assume quality of care provided by the agents cannot be observed from patient outcomes. Hence, this is a scenario with complete asymmetric information, i.e. neither can the principal observe the quality level of care (i.e. agents' actions) nor can they infer it from patient outcomes.

The principal wishes to maximise their utility. This, however, depends on the effort level of the agent. The agent maximises their utility and therefore chooses the following effort level obtained by taking the first order condition and setting it equal to zero:

$$\max_e \quad u_A = w - c(e - \underline{e})$$

$$\begin{aligned} u_A' = -c'(e - \underline{e}) &= 0 \\ \Leftrightarrow -((e - \underline{e})^2)' &= 0 \\ \Leftrightarrow -(2e - 2\underline{e}) &= 0 \\ \Leftrightarrow e^* &= \underline{e} \end{aligned}$$

Hence, the agent chooses effort level $e^* = \underline{e}$ as their optimal effort level. Knowing this, the principal faces the following maximisation problem subject to the PC of the agent:

$$\begin{aligned} \max_{w} \quad & u_P = q(e) - w \\ \text{subject to} \quad & w - c(e - \underline{e}) \geq \bar{u} = 0 \end{aligned}$$

The principal needs to make sure the PC holds, i.e. the wage they pay results in at least equal utility for the agent compared to the outside option. This implies the wage paid equalises at least the effort costs which are assumed to be zero when exerting minimum effort:

$$\begin{aligned} w - (\underline{e} - \underline{e})^2 &\geq c(\underline{e}) \\ \Leftrightarrow w^* &\geq c(\underline{e}) = 0 \end{aligned}$$

Since the principal's utility decreases in w, they maximise their utility by setting $w^* = c(\underline{e}) = 0$. In this scenario it is impossible for the principal to monitor the effort level of the agent and therefore it is optimal to pay their effort costs. Regardless of the wage paid, the agent will always choose $\underline{e}$ as their effort level, which results in quality outcome $q = \underline{e}$. Hence, the equilibrium in this scenario is: $equilibrium = (e^*, w^*) = (\underline{e}, 0)$ resulting in quality level $q^* = \underline{e}$.

Scenario 2: Full information

In this scenario, I assume the introduction of the alert allows the principal to perfectly observe the quality of care, i.e. agents' effort. Hence, the principal is able to monitor the quality level provided by the agent by introducing a quality-contingent wage scheme, i.e. $w(q(e))$. Suppose the principal wants to achieve quality level $\hat{q} > q$, requiring the agent to exert effort $\hat{e} > \underline{e}$. In addition to the PC, the agent also faces an incentive compatibility constraint (IC), i.e. in order to exert a higher effort level and receiving the higher wage, the agent's utility needs to be at least equal to the gained utility from exerting lower effort and receiving the lower wage. For simplicity, I assume the following quality-contingent wage scheme that assures the IC to hold in all cases:

$$\begin{cases} \hat{w} > 0 & \text{if } q = \hat{q} \\ w = -\infty & \text{if } q \neq \hat{q} \end{cases}$$

The agents receive $\hat{w}$ if they deliver quality outcome $\hat{q}$ which requires effort $\hat{e}$, but if they deviate they receive a negative wage, i.e. they are punished. This can be for example imagined as lawsuit costs. Assuming the PC holds, the agent will always exert the higher effort level to provide the targeted quality of care level.

Hence, the principal's maximisation problem is only subject to the PC:

$$\begin{aligned} \max_{w} \quad & u_P = q(e) - w(q(e)) \\ \text{subject to} \quad & \hat{w} - c(\hat{e} - \underline{e}) \geq \bar{u} = 0 \end{aligned}$$

The PC holds when $\hat{w} \geq c(\hat{e} - \underline{e})$. Therefore, the principal maximises their utility by introducing the following quality-contingent wage scheme:

$$\begin{cases} \hat{w} = c(\hat{e} - \underline{e}) & \text{if } q = \hat{q} \\ w = -\infty & \text{if } q \neq \hat{q} \end{cases}$$

Hence, the equilibrium in this scenario is: $equilibrium = (e^*, w^*) = (\hat{e}, c(\hat{e} - \underline{e}))$.

Given this equilibrium and $q(e) = e$, I infer the optimal level of quality by solving the principal's maximisation problem:

$$\max_{\hat{e}} \quad u_P = \hat{e} - c(\hat{e} - \underline{e}) = \hat{e} - (\hat{e} - \underline{e})^2$$

$$\begin{aligned} u'_P = 1 - (2\hat{e} - 2\underline{e}) &= 0 \\ \Leftrightarrow \hat{e}^* &= \frac{1}{2} + \underline{e} \end{aligned}$$

Assuming, the provider is willing to pay a higher wage for higher quality, this scenario improves the current situation. This can be shown formally by comparing the quality outcomes from scenario 1 and 2:

$$q(\hat{e}) = \frac{1}{2} + \underline{e} > \underline{e} = q(\underline{e})$$

The quality level is higher in scenario 2 compared to scenario 1. To conclude, I formally showed, the provider can monitor the quality level of care under full information (provided by the alert system) by introducing a quality-contingent wage scheme. Therefore, the provider can induce clinicians to follow guidelines, which is expected to lead to better patient outcomes. With the use of an empirical analysis as presented in Section 8, I test whether the alert system indeed changes clinicians' behaviour and improve patient outcomes.

5 Data and sample retained for analysis

Patient-level data was retrospectively retrieved using the Trust data warehouse between June 2016 and February 2018. Patients stayed either at St. Mary's hospital, at Charing

Cross hospital or at Hammersmith hospital, all located in London, UK. The three hospitals are study sites belonging to the "Imperial College Healthcare NHS Trust". I filtered the extracted data and only included patients who triggered at least one alert of the early detection sepsis module (active or silent) described in Section 2.2.

5.1 Data sets and their variables

Imperial College healthcare Trust provided datasets with pseudonymised identifiers for patients and their encounters for analysis in a secure environment. The first data set includes relevant demographic variables. Among others, it contains age, ethnicity, gender and the date of death (if applicable) of the patients. Data sets two and three contain relevant information on the inpatient episode (i.e. time on a hospital ward and not the emergency department) of patients. This encompasses admission date, discharge date, variables indicating the admission and discharge method (e.g. came to hospital by ambulance or was discharged to another hospital). Patients can be admitted multiple times to hospital during the observation period and hence multiple observations for the same person ID is possible. Furthermore, there are 25 diagnosis variables containing ICD-10 codes (International Classification of Diseases, Tenth Revision[3]). Every patient receives a primary diagnosis (diagnosis 1) and up to 24 secondary diagnoses. The secondary diagnoses also include existing conditions prior to the current hospital stay such as chronic diseases or cancer. To assess whether a diagnosis is sepsis related or not, 43 ICD-10 codes were identified as sepsis related. The selection is based on Inada-Kim et al. (2017) who define a "suspicion of sepsis" target population. In this process they identify the top 25 most common "suspicion of sepsis" diagnoses in the Oxford Academic Health Science Network region in 2013 and 2014. Besides these 25 diagnoses, the selection is further extended by ICD-10 codes that have a sepsis description (e.g. Streptococcal septicaemia). Table A.1 in the Appendix (Section A) includes the full list of selected ICD-10 codes.

Data set four contains information about the emergency department stays of patients. Among others this includes arrival as well as departure date and time, how the patient arrived at the emergency department and on which emergency department the patient is treated.

Data set five includes information about the alert module. It contains the point in time when the alert fired, an alert description, i.e. whether the alert was silent or active, and a result description which is entered by the clinician following the alert. Each alert is constructed

[3]International Classification of Diseases (ICD) is the international standard tool for diagnosis, epidemiology, health management and clinical purposes. Its 10th revised edition was endorsed in May 1990 (WHO, 2018)

as a new observation in the data set such that if a person had ten alerts this leads to ten different observations. However, the first alert is the most critical one since this notifies the medical staff that the patient is potentially infected with sepsis. Therefore, I focus on the first alert in the following analysis. Data set six includes information on microbiology tests performed on the patients. This encompasses the collection date and time, i.e. the point in time when a test sample is taken from a patient, as well as information on the test type, test outcome and whether antibiotics were described. Similar to the alert variable, there are multiple collection dates per person and those are constructed as different observations. Lastly, data set seven includes information on the ward movements of patients. During one hospital stay patients can be moved between different wards. This data set allows to track patients, i.e. it includes information on when a patient was on which ward (ward start date, ward name and ward end date). Also here, each ward movement constitutes a new observation.

5.2 Final sample

In order to achieve a final sample that suits the needs of my analysis approach, I merge all data sets and reshape them into a wide format. Each row represents the pathway of a patient between admission date and discharge date/eventual date of death. Hence, for each patient the data contains multiple points in time where different events take place. In the later analysis, admission date or arrival date on the emergency department if applicable is denoted as time zero for each patient ($t = 0$). The observation end date is either the date of death or 22/02/2018 which is the last date of observation[4]. Apart from date of death, all other events such as the alerts or tests occur between $t = 0$ and discharge. However, they are not evenly distributed between admission and discharge, but the distribution varies across patients. One row, i.e. one hospital stay of a patient, can be viewed as a "time-series" in the sense that a patient moves from $t = 0$ until the observation end date.

To assess the effect of the alert, I divide the sample into control and treatment group. The treatment group includes all patients that had an active first alert, whereas the control group is composed of patients that had a silent first alert. Hence, the difference between control and treatment group lies in the type of the first alert (silent vs. active). To ensure control and treatment group are as similar as possible in terms of pre-treatment characteristics, I exclude all patients that had their first alert not on one of the wards with an active and a silent running period. This selection guarantees that patients in both groups stayed on the same wards for the first alert suggesting their treatment conditions differ only in the alert

[4]On 22/02/2018 date of death was updated for the last time

type. Additionally, I conduct a robustness check with a larger sample including all patients that had an alert (active or silent) independent from the alert location of the first alert.

Figure 4 gives an overview of the active and silent periods of the alert on the included wards in the baseline sample. All included wards are either specialised on acute care or on heamatology[5]. The dashed lines represent the silent period and the solid lines the active period. Hence, patients admitted during the silent running period are part of the control group whereas patients admitted during the active running period are part of the treatment group. For example, a patient admitted to the A&E department at St. Mary's hospital in May 2017 belongs to the control group while a patient admitted to the same A&E department in October 2017 belongs to the treatment group. Whether patients in control and treatment group are randomly allocated is discussed in Section 7.

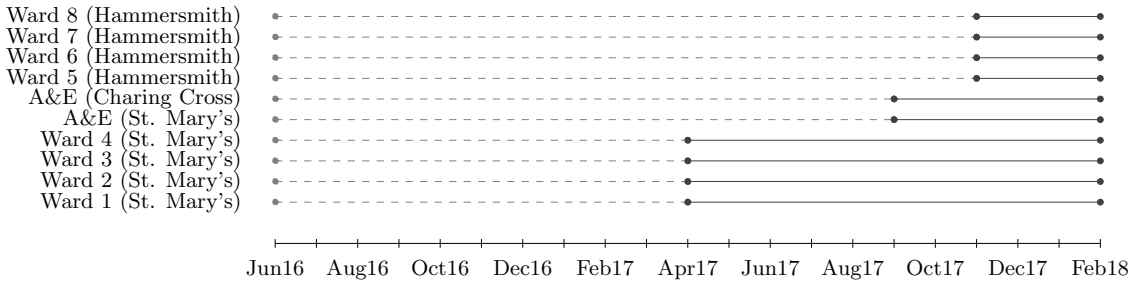


Figure 4: Wards

The final data set includes 2,563 observations and contains mainly repeated cross-sectional data, where each hospital stay is one row. Since only 11 patients stayed a second time in hospital during the observation period, the share of "panel data" is very small. The treatment group includes all patients whose first alert was after the ward specific "live"-date and therefore had an active alert. 938 observations (37%) are in the treatment group, whereas 1,625 observations (63%) are part of the control group.

5.3 Outcome variables

To assess the change in behaviour of clinicians I consider two variables: *time between infection and microbiology test orders*; and *frequency of microbiology test orders.*

In order to treat patients appropriately, clinicians order microbiology tests and prescribe/adjust the (antibiotics) treatment based on the results. Thus, as one component of clinicians' behaviour, I examine whether the time between infection and collection time of microbiology

[5]Heamatology is a discipline within medicine that manages the cause, prognosis, treatment, and prevention of diseases related to blood diagnose disorders of the blood and bone marrow (The Royal College of Pathologists, 2018)

test samples changes. However, various tests are taken on inpatients for various reasons. To extract sepsis related tests, I construct a variable that contains the test date of the first test collected after the first alert. Subsequently, the "Time between infection and microbiology test orders" variable measured in hours is constructed as follows:

$$\begin{aligned} &\textit{Time between infection and microbiology test orders} \\ &= \textit{Collection time of test sample 1} - \textit{Time of Alert 1} \end{aligned}$$

In addition to the timing of the test orders, I also assess the frequency of test orders. One reaction of clinicians to the alert system could be that they order more frequently tests to prescribe the appropriate medication as suggested by the care bundles. This variable is constructed by counting the number of tests collected after the first alert.

To assess patient outcomes I consider two variables: *mortality* and *LOS*. In the literature these variables are frequently used to assess the impact of an intervention on patient outcomes of sepsis patients (e.g. ARISE Investigators and the ANZICS Clinical Trials Group, 2014; Armen et al., 2016; Levy et al., 2010; Nguyen et al., 2007; Rivers et al., 2001).

In the UK, mortality rate of sepsis patients is 30% (NHS England, 2015). However, in cases where patients suffer from multiple conditions or do not die in hospital it is difficult to identify the death cause/relate it to sepsis. Furthermore, Prescott et al. (2016) highlight sepsis also affects the long-term mortality rate. They estimate a 22.1% absolute increase in late mortality relative to adults not in hospital. In addition, they find a 10.4% absolute increase in late mortality for sepsis patients relative to patients admitted with non-sepsis infections. They estimate mortality for sepsis patients remains higher for approximately two years relative to adults not in hospital. Most studies use in-hospital mortality rates and/or 28/30/31-day mortality (e.g. Gao et al., 2005; Nguyen et al., 2007; Yu et al., 2016). To also cover potential long-term effects, I assess overall mortality during the study period[6] next to in-hospital and 31-day mortality.

LOS impacts patients' well-being as well as contributes to the financial burden associated with sepsis patients. Hence, it is a commonly used indicator to assess the effect of an intervention. I construct LOS measured in days by subtracting the admission date (or arrival date on the emergency department if applicable) from the discharge date.

$$\textit{LOS} = \textit{discharge date} - \textit{admission/arrival date}$$

[6]The data contains records on date of death from June 2016 until 22/02/2018

5.4 Hypotheses

The two outlined mechanisms explaining the effect of the alert on clinicians' behaviour, earlier detection and a solution for the contracting problem, are both expected to change clinicians' behaviour in a way that improves patient outcomes. However, the data set at hand does not allow to differentiate between the effect of the two mechanisms since the variables used to measure behaviour are the same for both mechanisms. For example, when assessing whether test samples are collected earlier, it cannot be inferred whether this can be attributed to earlier detection or compliance to the bundle care (higher exerted effort by the clinician). For this reason, the share of each mechanism's influence cannot be measured.

Following this, I formulate two hypotheses regarding the effect of the alert on clinicians' behaviour, measured by *time between infection and microbiology test orders*; and *frequency of microbiology test orders*:

H1 *The active alert leads to reduced time between infection and microbiology test orders*

H2 *The active alert leads to increased frequency of microbiology test orders*

Patient outcomes, measured by *mortality* and *length of stay in hospital*, are indirectly affected by the alert through the expected change in clinicians' behaviour. Reduced time to test order and increased frequency of test orders should facilitate earlier and appropriate antibiotic treatment initiation which in turn improves patient outcomes. There are reasons to believe that *LOS* decreases due to the alert system since patients are expected to be cured faster. However, decreased mortality could also counteract this effect since a patient that would have died without the alert may survive and therefore stay longer in hospital (Westphal et al., 2010). However, based on the literature (e.g. Armen et al., 2016) I expect the former effect to outweigh the latter. Hence, I formulate the following two hypotheses concerning patient outcomes:

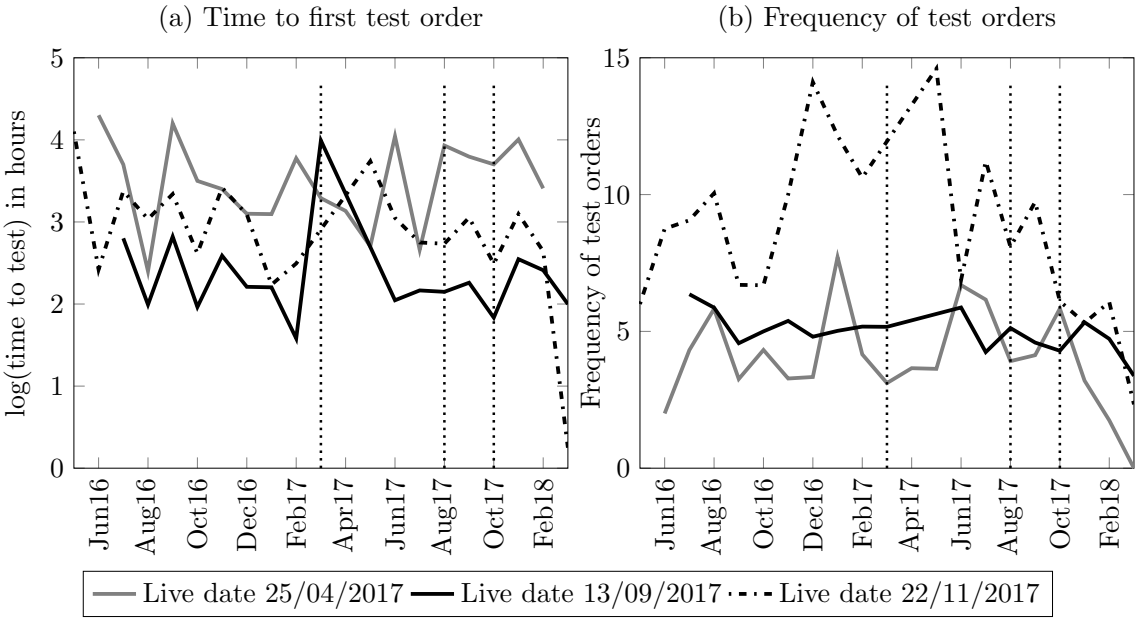
H3 *The active alert leads to reduced mortality*

H4 *The active alert leads to a reduced length of stay in hospital*

6 Descriptive results

I first graphically analyse the development of *clinicians' behaviour* and *patient outcomes* over time. This analysis allows me to understand whether the alert caused significant changes in the trend of the outcome variables. Second, I employ different regression forms

to further examine the relationship between alert and outcome variables as presented in Section 8.



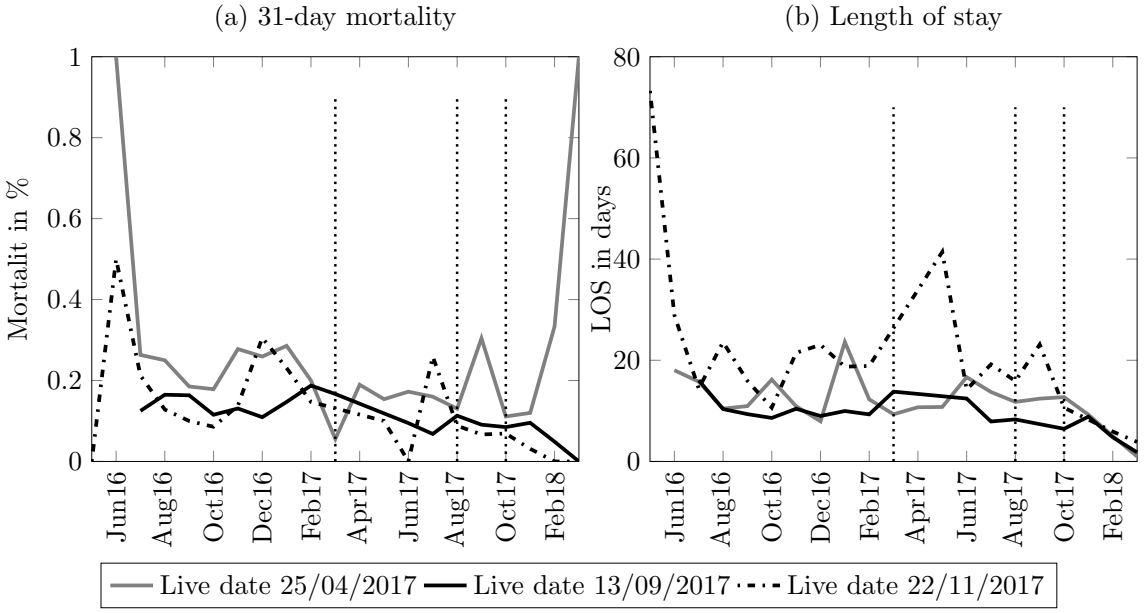
Source: Author's rendering of NHS data (2018).

Figure 5: Time trend of clinicians' behaviour

Figure 5 illustrates how clinicians' behaviour, namely the logarithm of *time to first test order* and *frequency of test orders*, developed during the observation time on ward level. The vertical dotted lines mark the live-date of the alert on the respective wards. The spikes in both variables during June–July 2016 and between April 2017 and June 2017 are likely caused by the low number of admissions during these periods (see Section 7). Apart from these spikes, the evolution of *time to first test order* is relatively volatile but the trend seems to be stable over time. Examining the development on ward level suggests no major changes after the respective live-dates indicating the treatment does not affect *time to first test order* significantly. Similarly, investigating the graph on the right hand side reveals *frequency of test orders* is relatively volatile. Furthermore, rather than dropping/increasing sharply after the respective live-dates, the trend of *frequency of test orders* seems to be relatively constant suggesting no significant impact of the alert on this variable.

Patient outcomes, measured in *31-day mortality* and *LOS*, seem to improve during the observation period. Figure 6 suggests that after a spike in June – July 2016 (caused by the initial low number of observations as outlined in Section 7) *31-day mortality* and *LOS* decrease over time. For patients encountered first on the emergency departments (live date 13/09/2017), *LOS* declines from on average about 18 days in August 2016 to about 10 days in December 2017, whereas *31-day mortality* drops from about 18% in August 2016

to about 10% in December 2017. Both graphs suggest the level and evolution of *LOS* and *31-day mortality* is similar across wards, but slightly higher and more volatile on the wards that are turned on in November 2017.



Source: Author's rendering of NHS data (2018).

Figure 6: Time trend of patient outcomes

There are no sharp drops after the respective live-dates of the alert but instead both LOS and mortality have a negative time trend. Hence, graphical examination does not resolve whether the negative time trend (i.e. improvements) in patient outcomes can be (partly) attributed to the introduction of the active alert. This is further investigated in Section 9.

Patients who are still in hospital at the end of the observation period are excluded in the average calculations of *LOS* per month. Similarly, patients admitted in January - February 2018 were not under observation for 31 days and hence if they did not die before 22/02/2018, they are counted as "not dead" in the average 31-day mortality computation. These "censoring" issues, i.e. excluding non-discharged patients and counting patients as not dead during 31 days even though they were not followed for 31 days, explain the drops of *LOS* and *31-day mortality* at the end of the observation period. Including these observations leads potentially to an upward bias of the effect of the alert. To mitigate this bias, I adjust the sample depending on the dependent variable as described in Section 8.

Table 1 shows the means for a selected number of pre-treatment characteristics. Reviewing the sample in terms of age, I find the average age in the sample to be 65 years, indicating that the sample consists mostly of pensioners. This is expected since sepsis is more prevalent among older patients (e.g. Martin et al. (2006)). Furthermore, the full sample includes

equal numbers of male and female patients in line with literature that indicates there are no clear gender-specific differences in sepsis prevalence (e.g. Failla and Connelly (2017)). Moreover, the vast majority of patients (88% in the whole sample) suffer from at least four comorbid conditions. On average, patients suffer from 11 comorbid diseases. Comorbidities are defined as coexistent diseases to the disease of interest, which potentially have a direct effect on the prognosis of the disease of interest, or an indirect influence on the choice of treatment (Huang et al., 2014). The high prevalence of comorbidities is not surprising since the average age in the sample is relatively high and older patients tend to suffer from more chronic conditions than younger patients (e.g. Wolff et al. (2002)). Moreover, about 36% of all patients have a sepsis diagnosis as their first diagnosis and about 82% of all patients were first encountered on the emergency department.

Table 1: Pre-treatment characteristics

	Sample Mean
Age	64.787
Male	0.511
Sepsis first diagnosis	0.361
No. of comorbidities	11.287
≥ 4 comorbidities	0.883
Emergency department	0.817

Note, coefficients are rounded to three decimal places.

7 Testing for randomised assignment

In order to infer causal relationships between the alert system and the outcome variables control and treatment group need to be randomly assigned. Random assignment implies that patients in control and treatment group only differ in the type of their alert (active vs. silent) but are comparable in terms of pre-treatment characteristics. This implies the introduction of the alert system should not be correlated with pre-treatment characteristics. To get an overview of the type of patients in control and treatment group, I first test for differences in means of selected pre-treatment characteristics. Second, I test whether pre-treatment characteristics are correlated with the introduction of the alert system. Lastly, I show how the number of admissions developed over the course of the observation period.

In Table 2 I report mean differences in pre-treatment characteristics between control and treatment group using t-tests. There are statistically significant differences (on the 5% or 1% significance level) in all characteristics except for having sepsis as the primary diagnosis. Patients in the control group are slightly younger compared to the treatment group and more often female. Furthermore, patients in the control group suffer from significantly

Table 2: Pre-treatment characteristics between control and treatment group

	Mean Control	Mean Treatment	Diff.
Age	64.136	65.916	-1.78**
Male	0.495	0.542	-0.047**
Sepsis first diagnosis	0.361	0.361	-0.001
No. of comorbidities	11.550	10.832	0.719***
≥ 4 comorbidities	0.922	0.816	0.106***
Emergency department	0.758	0.920	-0.163***

Note, coefficients are rounded to three decimal places.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

more comorbid conditions as well as more often from four or more comorbid conditions. In addition, they are less likely to be initially admitted to the emergency department. The differences are statistically significant, but relatively small in magnitude. The underlying reasons for the arising differences are difficult to detect. To account for these pre-treatment characteristics, I include them as control variables in the regression analysis (described in Section 8.4).

In order to generate unbiased estimates of the treatment effect, it is especially important that pre-treatment characteristics are uncorrelated with the introduction of the alert system. To test this assumption, I run Ordinary Least Squares (OLS) regressions on a selection of pre-treatment characteristics where the alert systems serves as explanatory variable. In line with my main model specification (see Section 8) I control for time and ward fixed effects to account for time and ward-specific trends in the sample. The specification to test for the correlation between pre-treatment characteristics and the alert system looks as follows

$$Y_{itw} = \beta_0 + \beta_1 T_i + \alpha_t + \delta_w + \epsilon_{itw}$$

where Y_{itw} is a vector of pre-treatment characteristics including age, gender, ethnicity, number of comorbidities and the presence of four or more comorbidities; T_i is a dummy variable indicating whether the alert was silent or active; and α_t and δ_w are time and ward fixed effects, respectively.

The regression results are reported in Table 3. Age, gender, ethnicity and the number of comorbid conditions seem not to be significantly correlated (on 5% or 1% significance level) with the introduction of the alert system. As opposed to the results of the t-tests, the absence of correlation between alert and age, gender and the number of comorbidities suggests the type of the alert, i.e. whether patients belong to control or treatment group, does not predict these characteristics. Furthermore, the magnitude of the effects for age, gender and the number of comorbidities indicated by the t-tests is relatively small (1.78

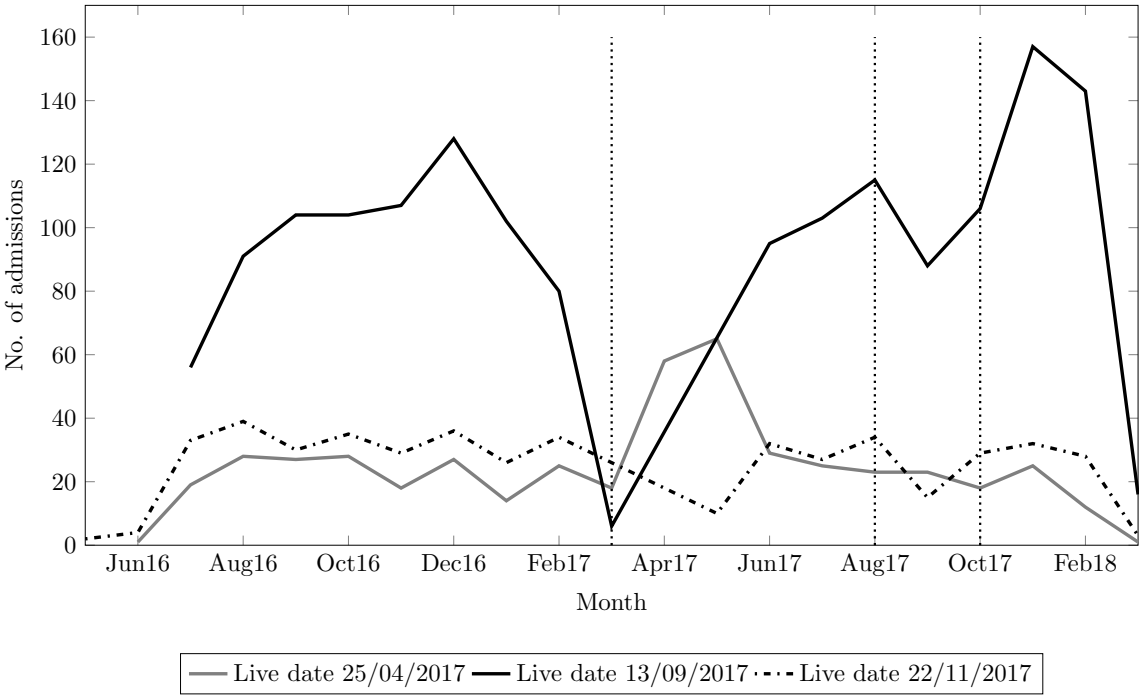
Table 3: OLS results of pre-treatment characteristics

	(1)	(2)	(3)	(4)	(5)
	Age	Male	Ethnicity	No. of comorbidities	≥ 4 comorbidities
Active alert	-0.101	0.057	0.682*	-0.623	-0.048**
	(1.844)	(0.049)	(0.356)	(0.634)	(0.022)
Constant	54.559***	0.440***	6.889***	10.433***	0.929***
	(2.397)	(0.072)	(0.455)	(0.783)	(0.035)
Observations	2563	2563	2222	2563	2563
Time fixed effects	yes	yes	yes	yes	yes
Ward fixed effects	yes	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

years in age, 4.7 percentage points in gender and 0.7 comorbidities) suggesting no economic meaningful differences. In contrast, the type of the alert appears to be a predictor of the presence of four or more comorbidities. However, the effect is relatively small in magnitude (an active alert corresponds to a 5 percentage point increase in the probability of having four or more comorbidities) suggesting no economic meaningful correlation between the presence of four or more comorbidities and the alert system.

Figure 7 provides further evidence for a reasonably random assignment of control and treatment group. The figure shows the development of the number of admissions over time on ward-level. Note, apart from low number of admissions in the first two months (June - July 2016) and a drop after the alert turned active in the first group of wards (April 2017 - June 2017) the number of admissions is relatively constant over time suggesting no difference in the number of admissions after the alert turned active. The low numbers of admissions in the beginning as well as between April and June 2017 can be explained by technical problems of the alert system, i.e. during these periods the alert system did not work properly and hence fewer alerts were documented. Since this sample only includes patients that had an alert, the number of admissions in this sample is reduced during times of technical problems. Patients admitted to the emergency department (live date 13/09/2017) account for the largest share of admissions. This is in line with the literature suggesting that a large share of sepsis patients is first encountered on emergency departments (e.g. Nguyen et al., 2007). The figure also suggests the volatility of total admissions is predominantly driven by the volatility in the admissions to the emergency department.



Source: Author's rendering of NHS data (2018).

Figure 7: Time trend number of admissions

Given only minor differences in pre-treatment characteristics and a relatively stable number of admissions over time I conclude it is reasonable to assume randomised assignment of control and treatment group.

8 Empirical strategy

I estimate the difference in *clinicians' behaviour* and *patient outcomes* between patients that had a silent alert and patients that had an active alert. I conduct the estimations using the following three types of models:

1. Ordinary Least Squares (OLS) regression with time and ward fixed effects
2. OLS regression with ward fixed effects
3. Cox proportional hazard model with mortality as the event variable (survival analysis)

In the following three sections I describe and motivate each of these three models in detail.

8.1 OLS regression with time and ward fixed effects

My first regression using OLS with time and ward fixed effects has the following specification:

$$Y_{itw} = \beta_0 + \beta_1 T_{itw} + X\beta + \alpha_t + \delta_w + \epsilon_{itw} \tag{1}$$

The outcome vector Y_{itw} represents the following four dependent variables (see Section 5.3):

- *Time between infection and microbiology test orders* measured in hours,
- *Frequency of microbiology test orders*,
- *Mortality* as a binary variable which is equal to 1 if patient i died and 0 otherwise,
- *Length of stay in hospital* measured in days,

for patient i who was admitted to hospital in month t and had their first alert (no matter if silent or active) on ward w. I use three different measures for *mortality*, namely in-hospital, 31-day and total mortality. To mitigate the aforementioned potential upward bias caused by "censoring" issues, I restrict the sample to patients that were already discharged at the end of the observation period for in-hospital mortality and LOS. To estimate 31-day mortality I exclude patients that were admitted less than 31-days before the end of the observation period[7]. Since it is difficult to find a suitable sample for total mortality, I estimate this variable with the use of a Cox proportional hazards model as explained in Section 8.3.

X is a vector of patient-specific characteristics that allows to control for factors likely influencing the outcome variables. The content of this vector of control variables is described in Section 8.4. Furthermore, time and ward fixed effects, denoted by α_t and δ_w respectively, are included. The time (month) dummies estimate the common difference/change in the outcome variables relative to the first month controlling for ward fixed effects, treatment and patient-specific characteristics. In other words, α_t "takes out" monthly time trends that are independent from the included variables. Similarly, the ward dummies estimate the common difference/change in the outcome variables on ward w relative to ward 1. Hence, they capture unobservable time-invariant variables. Therefore, the only identifying variation (variation not controlled for by the covariates) is due to the wards turning on the alert at different points in time. This variation is captured by T_{itw} which is the main variable of interest. It is a binary variable equal to 1 if patient i had an active alert and equal to 0 if they had a silent alert. My main focus lies on the interpretation of the β_1-coefficient which captures the treatment effect, i.e. the effect of the active alert on the outcomes of

[7] This corresponds to a cut-off on 21/01/2018.

interest. The estimation of the treatment effect is based on the comparison between the silent and active alert periods as well as on the comparison between the different wards.

I control for monthly trends as opposed to weekly or daily trends for two reasons. First, controlling for weekly or daily trends takes out large parts of the variation in the alert variable. However, the effect of the alert becomes apparent through time variation. Therefore I only want to control for time trends not caused by the alert rather than for all variation. Second, controlling for daily or weekly trends requires significantly more computing power.

To assure that the comparison of treated versus non-treated patients is valid and yields unbiased results the main identifying assumption is the *common trends assumption*. The assumption states that control and treatment group have common trends, implying the trends in the outcome of interest would have been identical for treated (active alert) and non-treated (silent alert) patients in the absence of an active alert. This goes back to the problem that there might be unobserved heterogeneity between control and treatment group which is included in the error term and correlated with the treatment indicator T_i (Angrist and Pischke, 2009, pp. 230–231; Lechner et al., 2011). In Section 7, I provide evidence that patients in control and treatment group exhibit only minor differences in pre-treatment characteristics suggesting that control and treatment groups are randomly drawn. Furthermore, from a-priori reasoning follows there is no selection bias into control and treatment group. Patients do not choose when to be infected with sepsis and hence cannot select whether they are treated during an active or silent running period. Based on the evidence suggesting randomised assignment of control and treatment group (presented in Section 7) and a-priori reasoning the *common trends assumption* is reasonable.

Another key assumption in order to generate unbiased estimates when conducting an OLS regression is that the error term ϵ_{itw} is assumed to fulfil the *zero conditional mean assumption*, namely that

$$E(\epsilon_{itw}|T_{itw}, X_i) = 0.$$

This assumption implies that the error term is uncorrelated with the independent variables (Wooldridge, 2013, p. 86), i.e. that patient characteristics are uncorrelated with the introduction of the active alert. As I include time and ward fixed effects to control for any differences not caused by the included variables and the *common trends assumption* holds, the *zero conditional mean assumption* is reasonable. To account for potential arising heteroskedasticity, I use robust standard errors (Wooldridge, 2013, pp. 93–102). Given these assumptions hold the model allows a causal interpretation of the effect of the alert, i.e. of the β_1-coefficient.

Four out of the five outcome variables are continuous, which allows to use "classic" OLS. However, *mortality* is a binary outcome variable. The OLS approach for binary variables is called "Linear Probability Model" (LPM). I argue for the use of this model as opposed to more conventional non-linear models such as the binary logistic regression model, because LPM allows a convenient interpretation of the coefficients since it immediately reports the magnitude of the estimated effects and generates more precise estimators when including fixed effects. Furthermore, Hellevik (2009) and Angrist and Pischke (2009, pp. 104–107) argue that in many applications LPM generates similar estimates to logit models.

The downside of including time fixed effects is that the time dummies "take out" the treatment variations which leads to less precise estimates. In the next section, I discuss this issue and present the model without time fixed effects.

8.2 OLS with ward fixed effects

Including time and ward fixed effects allows me to isolate the effect of the treatment on the outcome variables independent from time and ward-specific trends in the sample. This model is less vulnerable to omitted variable bias compared to a model without fixed effects and therefore allows to generate an unbiased estimate of the effect of the alert. However, including time dummies takes out all time-variation common to all wards in the alert variable. Hence, including time fixed effects leads to an imprecise estimate which is reflected in an increased standard error of the alert variable. Hence, whether to include time fixed effects is a trade-off between bias and precision/efficiency of the estimator.

Clearly, to infer causal relationships, an unbiased estimator is essential. Hence, the model without time fixed effects serves solely as an addition to the base model with a more precise estimate of the treatment effect.

The second regression model excluding time fixed effects has the following specification:

$$Y_{itw} = \beta_0 + \beta_1 T_{itw} + X\beta + \delta_w + \epsilon_{itw} \tag{2}$$

All components of the specification are identical to those in Regression 1. The only difference is that α_t is excluded for the aforementioned reasons. The variation allowed in the treatment variable in this regression is higher than in Regression 1 as it allows for time variation. This approach is expected to generate more precise estimates at the cost of potentially biased estimates.

8.3 Survival analysis

Survival analysis encompasses a wide variety of methods that can be utilised to analyse the timing of events (time-to-event data). It is used to analyse the expected time until one or even multiple events occur. In this context I aim to predict the time between infection (i.e. the first alert) and death of a patient.

As opposed to OLS or logistic regression models survival models allow to take censored data into account. Following Clark et al. (2003), for many observations the survival time is unknown since the studied event did not occur during the study period. For example in this sample, only some patients die during the observation period, but of the remaining patients some will die of sepsis at a later point in time. This kind of censoring, where the event time of the observation is beyond the end of the study, is called *right censoring* (Kartsonaki, 2016). In addition, in survival models it is possible to deviate from the normality distribution assumption and make assumptions regarding the distribution of the event variable in the data (Cleves et al., 2010).

To conduct survival analysis, it is essential to define times (time origin and endpoint or event of interest) appropriately to provide a solid framework for the analysis. I define time origin as the admission time (or arrival time at the emergency department if applicable) and death as the event variable. Hence, all patients that did not die during the observation period are censored.

Survival and hazard function

According to Clark et al. (2003) survival data are usually described in terms of two related probabilities: *survival* and *hazard*. The survival probability is also called survivor function and measures the likelihood that an observed object survives from the time origin until a specified time in the future (e.g. end of the observation period). In this study this is the probability that a patient survives from the point of being infected with sepsis (i.e time of the alert) until the end of the study period. In contrast, the hazard function gives the conditional failure rate, i.e. the instantaneous rate at which individuals experience events given they have survived until time t (Rodríguez, 2017). Survivor function and hazard function are related and can be obtained from each other (Kartsonaki, 2016; Rodríguez, 2017). These functions provide the basis for survival models which consist of two parts. The first part is characterised by the baseline hazard function which describes how the risk of the occurrence of an event changes over time based on baseline levels of the covariates. The second part consists of the effect parameters, which describe how the hazard changes in response to explanatory covariates (Cox and Oakes, 1984, pp. 62–77).

One very common survival model is the Cox proportional hazard model. I use this semi-parametric model in this context for two reasons: First it allows to assess the influence of covariates on the survival probability as opposed to non-parametric models and, second, the use and interpretation of the Cox proportional hazard model is much easier compared to parametric models (Cleves et al., 2010; Kartsonaki, 2016).

Cox proportional hazards model

The goal of the Cox proportional hazards model is to examine how specific factors influence the occurrence rate of a certain event at a precise point in time. The main assumption that Cox proportional hazard models rely on is the proportional hazard assumption, i.e. that the hazard ratio (ratio of the hazard function to the baseline hazard) is constant over time (Kartsonaki, 2016). Hence, one only makes assumptions about the baseline hazard $h_0(t)$ but not on the distribution of the event variable (Rodríguez, 2017).

The Cox proportional hazards model takes an exponential form which leads to the following hazard function (Cox, 1972):

$$h(t) = h_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n) \tag{3}$$

where t represents the survival time, x is a vector of explanatory variables, β is a vector of coefficients (it measures the impact of each explanatory variable on the hazard probability) and h_0 is the baseline hazard. The baseline function is assumed to be some function $h_0(t)$ and is therefore free of any assumptions, whereas the covariates are assumed to follow a linear function in the form of $exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n)$. In my model specification, x includes the alert variable, the time and ward dummies as well as the patient-specific characteristics described in Section 8.4.

When running a Cox-regression, Stata reports the hazard ratios (HR), i.e. the values $exp(b_i)$. A HR of 1.05 for the variable Male, for example, indicates that males face a 5% higher hazard than females.

8.4 Control variables

The choice of control variables is based on previous studies where the selected variables have proven to be essential in the context of sepsis patient outcomes (e.g. ARISE Investigators and the ANZICS Clinical Trials Group, 2014; Armen et al., 2016; Wang et al., 2012; Westphal et al., 2010).

In all regressions, (1), (2) and (3), X is a vector of patient-specific characteristics that are time-invariant including:

- A binary variable indicating if a patient is male.
- A continuous variable indicating age of the patients in years.
- An interaction term of age and gender to control for interactions between age and gender.
- A continuous variable indicating age squared as well as the interaction of this variable with gender.
- A categorical variable indicating ethnicity of the patient.
- A continous variable indicating the number of comorbidities a patient suffers from.
- A binary variable indicating whether the patient has four or more comorbidities.

Control variables such as gender, age and ethnicity are standard controls in the literature. Age and gender are known to influence patient outcomes in different medical contexts but also in the sepsis context (e.g. Armen et al., 2016; Westphal et al., 2010; Whittaker et al., 2015). Also ethnicity seems to have a major influence on sepsis patient outcomes: Esper et al. (2006) find significant differences in LOS and hospital discharge status between races.

The literature on sepsis patient outcomes suggests that comorbidities including chronic medical conditions play a major role in predicting patient outcomes, in particular mortality. For example, Yang et al. (2010) show that comorbidities (captured in the Charlson Comorbidity Index (CCI)[8]) are significant predictors of hospital mortality as well as LOS. Similarly, Heung and Koyner (2015) suggest that the recognition of the complex interplay between comorbid conditions such as chronic kidney disease or diabetes mellitus, sepsis and acute kidney injury, a severe complication that often occurs in the sepsis context, is the key to successfully manage these syndromes. Whittaker et al. (2015) point out that age and oncology diagnosis (common to sepsis patients and captured in the comorbidities variable) are risk factors associated with higher 28-day mortality rate from sepsis. This is also confirmed by Esper et al. (2006) who find that the presence of comorbidities is (a) differently distributed across gender and race and (b) associated with the development of acute organ dysfunction.

Since it is not feasible to control for each different comorbid condition individually, I control for the number of comorbidities and construct an indicator variable denoting whether patients suffer from four or more comorbid conditions. This approach is suggested by Whiles et al. (2016) who find that between sepsis patients with 0 and 1–3 comorbidities there is no

[8]CCI is the most widely used comorbidity index. Orginially, it was developed to predict one-year mortality among 604 patients based on comorbidity data (Charlson et al., 1987; de Groot et al., 2003).

difference in mortality while the odds of death elevated with the increase of comorbidities to 4 (or even more comorbiditities).

9 Results and analysis

Using the aforementioned regressions, I present and analyse the effect of the alert on *clinicians' behaviour* and *patient outcomes*. Additionally, I conduct robustness checks with other samples to back the results. Interpretations of the findings are provided in Section 10.

9.1 Clinicians' behaviour

The results of the OLS regressions on *clinicians' behaviour* are reported in Table 4. The first two columns include the results on *time to first test order* and the two last columns show the results on *frequency of test orders*. Columns (1) and (3) represent the main specification, i.e. with time and ward fixed effects, whereas columns (2) and (4) contain the results of the specification only with ward fixed effects.

In both specifications *time to first test order*, i.e. time between the alert (infection time) and the first microbiology test sample collection after the alert, seems not to be impacted by the treatment (silent vs. active alert). In the main specification, the alert coefficient is negative (-0.9) as expected, but the effect is not significant. The standard error is relatively large leading to a large confidence interval that ranges between -13 and 11 and hence an insignificant coefficient. Thus, in line with the graphical analysis presented in Section 6, I do not find support for hypothesis **H1**: The active alert leads to reduced time to first test order.

The effect of the active alert on *frequency of test orders* is positive and significant in the main specification indicating that an active alert leads to an increase in the frequency of test orders of 1.2 tests. The confidence interval of the alert coefficient ranges between 0.1 and 2.2. The positive effect is in line with expectations and supports hypothesis **H2**, i.e. the active alert leads to an increased frequency of microbiology test orders. The positive effect contradicts the graphical evidence in Figure 5. However, since the graphs are very volatile, it is difficult to detect a clear trend.

Comparing the results of *time to first test order* and *frequency of test orders* with and without time fixed effects, it becomes apparent the alert coefficients differ substantially in magnitude, direction and significance, whereas the coefficients of the controls are reasonably similar. Including time fixed effects, standard errors of the alert variable become three to four times larger. In contrast, standard errors for the controls do not differ significantly

Table 4: OLS results of clinicians' behaviour

	(1) Time to first test order	(2) Time to first test order	(3) Frequency of test orders	(4) Frequency of test orders
Active alert	-0.916	0.180	1.163**	-0.027
	(6.177)	(1.555)	(0.521)	(0.202)
Age	0.152	0.164	0.050	0.052
	(0.233)	(0.231)	(0.039)	(0.040)
Male	15.150	17.240	0.671	0.398
	(12.158)	(12.159)	(1.766)	(1.779)
Male × Age	-0.540	-0.591	-0.012	-0.004
	(0.469)	(0.472)	(0.061)	(0.061)
Age^2	-0.001	-0.001	-0.000	-0.000
	(0.002)	(0.002)	(0.000)	(0.000)
Male × Age^2	0.005	0.005	-0.000	-0.000
	(0.004)	(0.004)	(0.000)	(0.000)
No. of comorbidities	0.493**	0.470**	0.232***	0.223***
	(0.200)	(0.205)	(0.026)	(0.026)
≥ 4 comorbidities	-5.106**	-5.442***	-1.441***	-1.551***
	(2.192)	(2.078)	(0.346)	(0.335)
Constant	20.498**	15.998*	-0.931	-2.351**
	(8.253)	(8.460)	(1.136)	(1.110)
Observations	2079	2079	2222	2222
Time fixed effects	yes	no	yes	no
Ward fixed effects	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

across specifications. In line with the reasoning in Section 8, adding time dummies takes out variation in the alert variable. Thus, whether to include time dummies is a trade-off between bias and precision of the estimator.

9.2 Patient outcomes

The results of the OLS regressions on patient outcomes are reported in Table 5. The first two columns show the results on *LOS*, columns (3) and (4) on *in-hospital mortality* and the last two columns on *31-day mortality*. In all main specifications, the alert coefficients are not significant. The alert coefficient on *LOS* is 0.3 with a confidence interval between -2.7 and 3.3. For *in-hospital mortality* and *31-day mortality* the alert coefficients are -2.3 and -3 percentage points with confidence intervals ranging between -8 and 3 and between -7 and 6, respectively.

Table 5: OLS results of patient outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	LOS	LOS	In-hospital mortality	In-hospital mortality	31-day mortality	31-day mortality
Active alert	0.338	-2.494***	-0.023	-0.033***	-0.003	-0.052***
	(1.527)	(0.519)	(0.030)	(0.012)	(0.034)	(0.014)
Age	0.136	0.129	-0.003	-0.003	-0.004*	-0.004*
	(0.103)	(0.104)	(0.002)	(0.002)	(0.002)	(0.002)
Male	2.679	2.117	0.068	0.049	0.002	-0.021
	(4.016)	(4.023)	(0.077)	(0.077)	(0.084)	(0.084)
Male × Age	-0.070	-0.048	-0.002	-0.002	-0.000	0.001
	(0.148)	(0.148)	(0.003)	(0.003)	(0.004)	(0.004)
Age^2	-0.001	-0.001	0.000*	0.000**	0.000**	0.000**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Male × Age^2	0.000	0.000	0.000	0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
No. of comorbidities	0.852***	0.844***	0.013***	0.013***	0.017***	0.016***
	(0.070)	(0.069)	(0.001)	(0.001)	(0.001)	(0.001)
≥ 4 comorbidities	-3.370***	-3.389***	-0.087***	-0.087***	-0.072***	-0.076***
	(0.807)	(0.790)	(0.015)	(0.015)	(0.014)	(0.014)
Constant	19.962***	15.951***	0.168**	0.157**	0.181**	0.163**
	(3.773)	(3.743)	(0.082)	(0.072)	(0.090)	(0.080)
Observations	2170	2170	2170	2170	2109	2109
Time fixed effects	yes	no	yes	no	yes	no
Ward fixed effects	yes	yes	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The highly significant effect of the alert in the specifications without time fixed effects suggest, there is a time trend leading to reduced *mortality* and *LOS*. The depicted time trend is in line with the graphical evidence presented in Figure 6. LOS is reduced by 2.5 days, in-hospital mortality by 3.3 percentage points and 31-day mortality by 5.2 percentage points in the active-alert period compared to the silent-alert period. Due to the tripled standard errors in the main specifications, the effects (coefficients) of the alert need to be quite large in order to be significant. Since controlling for time trends takes out variation in the alert variable, it is reasonable the effect of the alert is not big enough to be significant and hence parts of the time trend (reduced mortality and LOS) can be attributed to the active alert. Thus, I find suggestive support for hypothesis **H3** and **H4**, i.e. the active alert reduces mortality and LOS. In order to confirm and quantify the effect of the alert, further investigations of the time trend and its relation to the alert system is needed. However, such analysis exceeds the scope of this thesis.

The results of the Cox proportional hazards regression are presented in Table 6. The reported hazard ratio of the alert variable measures the effect of the alert on the probability that a treated patient dies compared to the probability that a non-treated patient dies. In the main specification, i.e. with time fixed effects, the hazard ratio is insignificant but suggests an increased hazard probability. However, the standard error is quite large, leading to a wide confidence interval (0.6 to 1.6) and hence an unreliable estimate.

Table 6: Results of Cox-ph regressions

	(1)	(2)
	Event: death	Event: death
Active alert	1.025	0.69***
	(0.245)	(0.087)
Age	1.073**	1.074**
	(0.038)	(0.037)
Male	1.487	1.135
	(2.346)	(1.776)
Male × Age	0.994	1.001
	(0.047)	(0.047)
Age^2	1.000	1.001
	(0.000)	(0.000)
Male × Age^2	1.000	1.000
	(0.000)	(0.000)
No. of comorbidities	1.121***	1.111***
	(0.01)	(0.01)
≥ 4 comorbidities	0.861	0.756
	(0.306)	(0.260)
Observations	2218	2218
Time fixed effects	yes	no
Ward fixed effects	yes	yes
Ethnicity	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results of the Cox regression confirm that there is a time trend in the sample leading to reduced *total mortality* shown by the specification without time fixed effects. The results in column two suggest, patients who are exposed to the active alert face only 69% of the hazard compared to patients who had a silent alert. Similar to the OLS results, it is likely that a small effect of the alert is not indicated by the main specification due to the large standard error. In line with the interpretation of the OLS results, I interpret the Cox regression results as suggestive evidence towards a negative impact of the alert on mortality. As such, further research is needed to confirm and quantify the effect.

9.3 Robustness checks

I conduct two robustness checks to explore whether my results are ward-specific or biased by technical issues during the observation period.

9.3.1 Inclusion of more wards

The baseline sample only includes patients that had their first alert on a ward that has both an active and a silent running period. To investigate whether the results hold across other wards I conduct a robustness check with a sample including also patients that had their first alert on a ward where the alert system has not been turned active during the observation period. This results in an enlarged control group (2,264 observations) leading to a total sample size of 3,202 observations. The results for *clinicians' behaviour* (Table B.1) and *patient outcomes* (Table B.2 and B.3) are presented in the Appendix in Section B.

Comparing the results on *clinicians' behaviour* based on the larger sample with the results based on the baseline sample suggests the effect of the alert is robust to the increased control group for *time to first test order* but not for *frequency of test orders*. In the main specification the effect of the alert on *frequency of test orders* is no longer significant and the alert coefficient is smaller. This finding indicates the effect found in the baseline sample is ward-specific and does not apply when other wards are added. To understand in which contexts the alert significantly increases *frequency of test orders* further inspection of the characteristics of the different wards is needed. However, the results on *patient outcomes* are robust to the increased sample as the results are similar to the baseline results. Therefore, *frequency of test orders* does not seem to have a major impact on *patient outcomes*.

9.3.2 Exclusion of periods with technical issues

As mentioned in Section 7 in June–July 2016 as well as between April 2017 and June 2017 there were technical problems with the alert system leading to a significantly reduced number of observations in these months. This also results in spikes/drops in the outcome variables during these periods. To ensure the estimates obtained from the baseline sample are not biased due to the technical problems, I conduct a robustness check where I exclude all observations admitted in the problematic months.

The reduced sample includes 2,399 observations whereof 1,597 are part of the control group and 802 belong to the treatment group. The results of the analysis are presented in the Appendix in Section C in Table C.1 (*clinicians' behaviour*) and Table C.2 and C.3 (*patient outcomes*).

The results both for *clinicians' behaviour* and *patient outcomes* are robust to excluding the months with technical problems. Hence, the results in the baseline sample seem not to be biased by the observations during the periods with low number of admissions.

10 Discussion

In the main specifications, my results indicate no significant effect of the alert on patient outcomes, namely *mortality* and *LOS*. However, the included time dummies take out variation in the alert variable reflected by increased standard errors. Hence, potential smaller effects appear to be insignificant. Since the results without time fixed effects suggest that there is a time trend in the sample leading to improved patient outcomes, it is reasonable that parts of this time trend can be attributed to the active alert. As suggested by the specifications without time fixed effects, *LOS* decreased by 2.5 days, *hospital-mortality* by 3.3 percentage points and *31-day mortality* by 5.2 percentage points in the period with the active alert. Since in OLS regressions censored data is not accounted for, I estimate the effect of the alert on *total mortality* with a Cox regression. The reduction in the risk of dying depicted by the Cox regression corresponds to a decrease of about 31%. The results are robust to changes in the sample composition, suggesting the effect of the alert is not biased by technical issues in the beginning of the observation period and also apply to other wards.

Due to the apparent time trend, I am not able to obtain a definitive effect of the alert or the size of the effect. In order to do so, further investigations of the time trend and its correlation to the alert is required. For example, it would be interesting to assess whether the installation of the alert system went along with another intervention/change that also affects mortality and LOS and depends on the introduction of the alert system. Finally, the results serve as suggestive evidence that alert systems have the potential to reduce *mortality* and *LOS*, but further research is needed to confirm the effect.

Furthermore, I find only limited support for significant changes in the variables measuring *clinicians' behaviour*, which is likely due to the fact that I only assess two variables although clinicians' behaviour is multifaceted. Hence, there is likely a large share of unobservable behavioural changes that further contribute to the trend of reduced *mortality* and *LOS* that may be partly driven by the active alert. The shares of the change in *mortality* and *LOS* that can be attributed to the treatment effect are caused by the changed behaviour of clinicians. As reasoned in Section 4, the underlying reasons for the change in behaviour of clinicians are likely a mixture of earlier recognition and higher compliance to the guidelines induced by the reduced asymmetric information distribution between clinicians and care

provider. Other components of clinicians' behaviour that may influence patient outcomes are, i.a. the timing of the initiation of antibiotics treatment and movement of patients to wards with higher level of care (more intense surveillance).

One reason why I do not observe a significant effect of the alert on *time to test order* could be that the microbiology test variable includes a variety of tests and from most patients test samples are frequently collected. Hence, if test samples are already collected in a timely manner during the control period, the active alert does not change the timing but might induce to collect different types of tests that are more relevant to sepsis patients. However, since I do not have information on the type of test conducted, this is a hypothesis that needs further investigation and justification. In contrast, the results on *frequency of test orders* suggest that *frequency of test orders* is increased in the active alert period. However, this effect is not robust to including other wards to the sample whereas the results on patient outcomes are robust to the expansion of the sample. Hence, the effect of the alert on *frequency of test orders* is ward-specific and does not have a large impact on patient outcomes.

One explanation why most control variables in the regressions on *clinicians' behaviour* are insignificant could be that clinicians follow routines and guidelines (i.e. order tests as soon as they suspect a certain condition) which are reasonably independent of patient characteristics such as age and gender. In the OLS regressions on *patient outcomes* both age and gender are insignificant. This might be explained by the short time horizon of this study: Age and gender are both factors that influence mortality in the long-run but not necessarily in the short run. Since censored data are not included in OLS regressions, the long-term effect is not captured as opposed to the Cox proportional hazards model. This hypothesis is supported by the results of the Cox regression which show that an increase in age leads to an increased risk of dying. In line with expectations and the literature (e.g. Whiles et al., 2016), with increasing number of comorbidities the risk of dying increases as well.

10.1 Comparison with other findings and policy implications

Since this is the first study of its kind, i.e. evaluating an early-detection sepsis alert, there is no direct answer to whether my findings align with previous literature. However, I compare my results with findings of studies that evaluate alert systems in other disease contexts or concerning other outcome measures as well as with findings from studies that implemented guidelines to improve patient outcomes. Furthermore, I provide ideas for future research. Lastly, I derive policy recommendations from my findings.

10.1.1 Alert systems in other contexts

Literature on the effect of alert systems on clinicians' behaviour and patient outcomes is rather scarce and does most often not provide precise estimates of the effect of the alert system. Baysari et al. (2016) review different IT interventions and their effect on antimicrobial prescribing in hospitals. They find that high quality studies indicate an increase in the appropriateness of antimicrobial prescribing, which corresponds to a change in clinicians' behaviour. This suggests that alert systems do have the potential to alter clinicians' behaviour. Hence, it may be worthwhile to further assess the effect of the early detection sepsis module evaluated in my study on other components of clinicians' behaviour such as clinicians' antimicrobial prescribing behaviour and the timing of the initiation of the antibiotic treatment following an alert. In terms of patient outcomes, Baysari et al. (2016) suggest a positive effect on mortality and LOS but note that further research is needed to confirm their findings. My study provides suggestive evidence of the positive effect found in their study, but to confirm and quantify the effect, further investigation of the time trend is required.

Other studies on alert systems focus on the "alert overriding problem", i.e. clinicians ignoring alerts. Similarly to the problem that clinicians do not comply to guidelines, there is evidence, alerts may be overridden by the medical staff (e.g. Baysari et al., 2017). One reason why clinicians tend to override alerts is a phenomenon called "alert fatigue", i.e. that clinicians are overwhelmed by too many alerts/guidelines and therefore do not distinguish between important and non-important alerts but rather overrule them (Van Der Sijs et al., 2006). Alert overriding may also be a problem arising in the context of my study and therefore limit the effectiveness of the sepsis alert system. For further research, it would be interesting to have a closer look on the reaction of clinicians on the alert system and how the potential problem of alert overriding could be mitigated.

10.1.2 Implementation of guidelines

Many studies assessing the compliance and subsequently the success of bundle care guidelines for sepsis patients (e.g. Gao et al., 2005) find initially low compliance rates which is in line with evidence from other diseases (e.g. acute myocardial infarction (Wolff et al., 2004) or management of stroke (Vikman et al., 2004)). However, over the course of the study period Gao et al. (2005) observe that compliance rates with the care bundles rose which also led to a reduced hospital mortality by about 50% for the compliant group. Nguyen et al. (2007) and Chou et al. (2014) observe a similar pattern, i.e. compliance is initially low but increases over time and improves patient outcomes. In my study, the post-alert

period is relatively short, for example for the wards where the alert was turned active in November 2017 the post-alert period lasted only for 3 months. Hence, it is likely that similar to compliance to guidelines, working effectively with the alert system requires some time, i.e. a "learning-curve". Future research could conduct a follow-up study to this study to evaluate whether the treatment effect is more clear when the active alert has been in place for a longer time and clinicians had the chance to adapt to the new system.

10.1.3 Policy recommendations

Since my findings are rather suggestive evidence than definitive effects, further research is needed to formulate clear policy recommendations. However, in line with the aforementioned studies, this paper suggests a positive effect of the early-detection sepsis module on mortality and LOS. From a policy perspective, a new intervention needs not only be evaluated based on its effectiveness but also based on its costs. Hence, to come to a final conclusion regarding whether to implement this alert system a cost-effectiveness analysis is recommended. In order to do so, in a first step the treatment effect needs to be disentangled from the general time trend to obtain a valid estimate of the alert's effectiveness. Secondly, the effectiveness needs to be quantified in monetary values, i.e. how much is a life worth, and indirect and direct costs need to be considered. Decreased LOS hints towards decreased hospitalisation costs, but other implications such as whether medication costs increase, whether patients are more frequently treated on higher level of care units and the purchase and maintenance costs of the alert system need to be assessed. Furthermore, before implementing the system region- or nationwide external validity of this study should be reviewed as discussed in Section 10.3.

10.2 Limitations and internal validity

A major concern when studying the effect of an intervention on patient outcomes is whether there are other factors influencing patient outcomes during the observation period, distorting the causality of the observed effects. Furthermore, the prevalence of a selection bias into control and treatment group could lead to biased estimates. However, I believe my estimations to infer causality for two reasons. First, I control for time and ward specific trends in the sample mitigating the risk of an omitted variable bias and discuss the differences between the estimates with and without time fixed effects. Second, the implementation of the active alert can be thought of as an exogenous shock which suggests that patients could not self-select into control and treatment group. Furthermore, I provide evidence that patients in control and treatment group are reasonably similar in their pre-treatment characteristics.

Therefore, the common trends assumption required to infer causality is arguably fulfilled. There are, however, other concerns with the internal validity put forward in the following three paragraphs.

10.2.1 Death cause

One issue concerning the data is that the death cause of a patient is not reported implying, my measures of mortality may include patients who died for other reasons. The percentage of patients included in the mortality measures that have non-sepsis death causes likely increases with the scope of the mortality measure, i.e. hospital mortality contains a smaller share of patients whose death cause is not sepsis related than the total mortality. Despite this drawback, I regard assessing total mortality as a valuable addition as sepsis has long-term impacts on mortality (e.g. Prescott et al., 2016).

The bias arising from this issue likely leads to an underestimation of the true treatment effect. The alert has potentially little effect on patients who suffer from severe other conditions in the late stage but the death of these patients is regarded as not prevented by the alert. To mitigate this downward bias, the data set needs to be extended by death cause.

10.2.2 Control wards

Another issue arising from data availability is the choice of the control wards. The analysis only includes wards that have a silent and an active running period of the alert system, i.e. the control wards are turned active at different points in time. Thus, the different live-dates suggest the implementation of the active alert system might not be entirely exogenous. I do not have information about why the system was turned active at different points in time, and hence cannot be entirely sure the live-dates are exogenous. For example, if ward-specific characteristics (e.g. severeness of the disease of patients) determine the live-date, the implementation of the active alert is endogenous. I show that control and treatment group are randomly allocated and therefore patient characteristics do not determine whether a patient was part of the control or treatment group. However, I do not account for the fact that within the control group patient- or ward-specific characteristics could determine the live-dates of the system on the different wards. If this was the case, the results might be subject to a bias. However, since the alert system is a computer-based system which should not be dependent on patient characteristics I believe the risk of endogeneity to be relatively small.

10.2.3 Cox proportional hazards assumption

The estimates of the Cox regression presented in Section 9 are subject to a bias in case the proportional hazards assumption does not hold. Hence, I test the plausibility of the proportional hazards assumption by testing whether the hypothesis of nonzero slope in a generalised linear regression of the scaled Schoenfeld residuals on time for individual covariates and globally can be rejected. This is equivalent to testing that the log hazard-ratio function is constant over time (Grambsch and Therneau, 1994). In addition, I employ graphical evidence in form of a "log-log" plot (i.e. $-ln\{-ln(survival)\}$ curves for each category of a nominal or ordinal covariate versus $ln(analysistime)$.

Testing the null hypothesis of no violation of the assumption, indicates the hypothesis cannot be rejected ($p > 0.05$). However, graphical evidence (presented in Figure D.1 in the Appendix in Section D) suggests, the proportional hazards assumption for being exposed to a silent alert versus an active alert (treatment) is violated. If the two lines were reasonably parallel, this would suggest the assumption is not violated, but this is clearly not the case. The contradictory evidence may arise because the sample size is not large enough to have sufficient power to reject the null hypothesis.

Having this in mind, I carefully interpret the estimates of the Cox regression. Uno et al. (2014) suggest that when the proportional hazards assumption is violated the estimated parameters may not be a meaningful measure of the between group difference. In contrast, Allison (2010, pp. 172ff.) holds the view that violation of the assumption does not cause serious problems as in such cases estimates can be interpreted as "average effects". Stratifying the model or proceeding with a parametric model are potential solutions to overcome the problem. However, in the scope of this thesis I proceed with the Cox proportional hazards model as it is, bearing in mind that the estimates might be biased in either direction. For further research, it would be interesting to compare the estimates when using a stratified or a parametric model.

10.3 External validity

The baseline mortality in this study is relatively low compared to other recent studies, e.g. Chou et al. (2014) report a baseline hospital mortality rate of 28% and Armen et al. (2016) of 23.5% which is respectively 18 and 13.5 percentage points higher compared to this study. The relatively low baseline mortality rate in my sample suggests sepsis patients are already relatively well managed in the three study site hospitals. Therefore, the treatment effect might be underestimated in this study and installing the alert system in other hospitals

and countries might lead to larger treatment effects.

Furthermore, the alert system is currently only installed on wards that are specialised on either acute care or heamatology. Whether it is also effective on other wards and care units, where potentially patients with other needs stay, remains to be assessed by future research.

11 Conclusion

The purpose of this study is to assess the impact of an early detection sepsis alert on patient outcomes. Specifically, in a first step I evaluate the effect of the alert system on clinicians' behaviour and in a second step its effect on patient outcomes. Using repeated cross-sectional patient-level data, I compare the outcomes of a control and treatment group that were exposed to a silent and active alert, respectively. The aim is to answer the question: *Can sepsis patient outcomes be improved by the early detection alerts?*

To theoretically explain the expected change in clinicians' behaviour and ultimately improved patient outcomes, I provide two mechanisms through which the alert affects clinicians' behaviour. First, the alert system helps clinicians to detect infections earlier which allows them initiate the treatment earlier. Earlier initiation of treatment is expected to improve patient outcomes. Second, next to its notification function, the alert system also documents clinicians' actions and patients' conditions which reduces asymmetric information in the principal-agent relationship between care provider and clinician. Thus, the care provider can monitor clinicians' behaviour which is expected to lead to higher compliance to the sepsis care guidelines.

Empirically I employ three different models to answer the research question: an OLS regression including time and ward fixed effects, an OLS regression including only ward fixed effects and a Cox proportional hazard model to account for censored data in the mortality variable. Using these regressions, I estimate the difference in clinicians' behaviour and patient outcomes for patients that were exposed to an active alert (treatment group) and patients that were exposed to a silent alert (control group). I find including time fixed effects generates an unbiased estimator but reduces the variation in the alert variable. This results in imprecise estimates reflected by increased standard errors and hence potential small effects appear insignificant.

The results indicate, there is no significant and robust effect of the alert on the two variables employed to measure clinicians' behaviour. However, since clinicians' behaviour is multifaceted and the literature on alert systems suggests alerts can potentially alter clinicians' behaviour, future work could shed light on the treatment effect on other components of

clinicians' behaviour.

Furthermore, the main specifications do not indicate a significant effect of the alert on patient outcomes. However, I find a time trend in the data leading to reduced mortality and LOS for patients that were exposed to the active alert. The results suggest that parts of this time trend can be potentially attributed to the alert, but further research is needed to disentangle the effect of the alert from other trends leading to improved patient outcomes.

In conclusion, my results can serve as suggestive evidence that sepsis patient outcomes can be improved by early detection alerts. However, I suggest for future research to investigate the time trend to confirm and quantify the effect of the alert on patient outcomes.

References

Allison, P. D. (2010), *Survival analysis using SAS: a practical guide*, Sas Institute, Cary, North Carolina, USA, p. 172 ff.

Angrist, J. D. and Pischke, J.-S. (2009), *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, Princeton, New Jersey, USA, pp. 104–107, 230–231.

ARISE Investigators and the ANZICS Clinical Trials Group (2014), 'Goal-directed resuscitation for patients with early septic shock', *New England Journal of Medicine* **371**(16), 1496–1506.

Armen, S. B., Freer, C. V., Showalter, J. W., Crook, T., Whitener, C. J., West, C., Terndrup, T. E., Grifasi, M., DeFlitch, C. J. and Hollenbeak, C. S. (2016), 'Improving outcomes in patients with sepsis', *American Journal of Medical Quality* **31**(1), 56–63.

Armstrong, B. A., Betzold, R. D. and May, A. K. (2017), 'Sepsis and septic shock strategies', *Surgical Clinics of North America* **97**(6), 1339–1379.

Arrow, K. J. (1963), 'Uncertainty and the welfare economics of medical care', *The American Economic Review* **53**(5), 941–973.

Arrow, K. J. (1984), *The economics of agency.*, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California, USA, pp. 1–24.

Barochia, A. V., Cui, X., Vitberg, D., Suffredini, A. F., O'Grady, N. P., Banks, S. M., Minneci, P., Kern, S. J., Danner, R. L., Natanson, C. et al. (2010), 'Bundled care for septic shock: an analysis of clinical trials', *Critical Care Medicine* **38**(2), 668.

Baysari, M. T., Lehnbom, E. C., Li, L., Hargreaves, A., Day, R. O. and Westbrook, J. I. (2016), 'The effectiveness of information technology to improve antimicrobial prescribing in hospitals: a systematic review and meta-analysis', *International Journal of Medical Informatics* **92**, 15–34.

Baysari, M. T., Tariq, A., Day, R. O. and Westbrook, J. I. (2017), 'Alert override as a habitual behavior – a new perspective on a persistent problem', *Journal of the American Medical Informatics Association* **24**(2), 409–412.

Blomqvist, Å. (1991), 'The doctor as double agent: Information asymmetry, health insurance, and medical care', *Journal of Health Economics* **10**(4), 411–432.

Brown, E. M. (2002), 'Guidelines for antibiotic usage in hospitals', *Journal of Antimicrobial Chemotherapy* **49**(4), 587–592.

Cabana, M. D., Rand, C. S., Powe, N. R., Wu, A. W., Wilson, M. H., Abboud, P.-A. C. and Rubin, H. R. (1999), 'Why don't physicians follow clinical practice guidelines?: A framework for improvement', *Jama* **282**(15), 1458–1465.

Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, C. R. (1987), 'A new method of classifying prognostic comorbidity in longitudinal studies: development and validation', *Journal of Chronic Diseases* **40**(5), 373–383.

Chou, S.-L., Chan, K.-S., Cheng, K.-C., Chou, W., Hung, H.-M. and Chen, C.-M. (2014), 'The implementation of sepsis bundles on the outcome of patients with severe sepsis or septic shock in intensive care units', *International Journal of Gerontology* **8**(2), 60–65.

Clark, T., Bradburn, M., Love, S. and Altman, D. (2003), 'Survival analysis part i: Basic concepts and first analyses', *Br J Cancer* **89(02)**, 232–238.

Cleves, M., Gould, W., Gutierrez, R. G. and Marchenko, Y. (2010), *An Introduction to Survival Analysis Using Stata*, 3rd edn, StataCorp LP, College Station, Texas, USA.

Cox, D. R. (1972), 'Regression models and life-tables', *Journal of the Royal Statistical Society. Series B (Methodological)* **34**(2), 187–220.

Cox, D. R. and Oakes, D. (1984), *Analysis of survival data*, Vol. 1, Chapman & Hall Ltd, London, UK, pp. pp. 62–77.

de Groot, V., Beckerman, H., Lankhorst, G. J. and Bouter, L. M. (2003), 'How to measure comorbidity: A critical review of available methods', *Journal of Clinical Epidemiology* **56**(3), 221–229.

Dellinger, R., Levy, M., Rhodes, A., Annane, D., Gerlach, H., Opal, S., Sevransky, J., Sprung, C., Douglas, I., Jaeschke, R., Osborn, T., Nunnally, M., Townsend, S., Reinhart, K., Kleinpell, R., Angus, D., Deutschman, C., Machado, F., Rubenfeld, G. and Moreno, R. (2013), 'Surviving sepsis campaign-guidelines committee including the pediatric subgroup. surviving sepsis campaign: International guidelines for management of severe sepsis and septic shock: 2012', *Critical Care Medicine* **41**, 580–637.

Esper, A. M., Moss, M., Lewis, C. A., Nisbet, R., Mannino, D. M. and Martin, G. S. (2006), 'The role of infection and comorbidity: Factors that influence disparities in sepsis', *Critical Care Medicine* **34**(10), 2576.

Failla, K. R. and Connelly, C. D. (2017), 'Systematic review of gender differences in sepsis management and outcomes', *Journal of Nursing Scholarship* **49**(3), 312–324.

Gao, F., Melody, T., Daniels, D. F., Giles, S. and Fox, S. (2005), 'The impact of compliance

with 6-hour and 24-hour sepsis bundles on hospital mortality in patients with severe sepsis: a prospective observational study', *Critical Care Medicine* **9**(6), R764–R770.

Grambsch, P. M. and Therneau, T. M. (1994), 'Proportional hazards tests and diagnostics based on weighted residuals', *Biometrika* **81**(3), 515–526.

Hellevik, O. (2009), 'Linear versus logistic regression when the dependent variable is a dichotomy', *Quality & Quantity* **43**(1), 59–74.

Heung, M. and Koyner, J. L. (2015), 'Entanglement of sepsis, chronic kidney disease, and other comorbidities in patients who develop acute kidney injury', *Seminars in Nephrology* **35**(1), 23–37.

Houck, P. M., Bratzler, D. W., Nsa, W., Ma, A. and Bartlett, J. G. (2004), 'Timing of antibiotic administration and outcomes for medicare patients hospitalized with community-acquired pneumonia', *Archives of Internal Medicine* **164**(6), 637–644.

Huang, Y.-q., Gou, R., Diao, Y.-s., Yin, Q.-h., Fan, W.-x., Liang, Y.-p., Chen, Y., Wu, M., Zang, L., Li, L. et al. (2014), 'Charlson comorbidity index helps predict the risk of mortality for patients with type 2 diabetic nephropathy', *Journal of Zhejiang University Science B* **15**(1), 58–66.

Inada-Kim, M., Page, B., Maqsood, I. and Vincent, C. (2017), 'Defining and measuring suspicion of sepsis: an analysis of routine data', *BMJ* **7**(6).

Iregui, M., Ward, S., Sherman, G., Fraser, V. J. and Kollef, M. H. (2002), 'Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia', *Chest* **122**(1), 262–268.

Jensen, M. C. and Meckling, W. H. (1976), 'Theory of the firm: Managerial behavior, agency costs and ownership structure', *Journal of Financial Economics* **3**(4), 305–360.

Kartsonaki, C. (2016), 'Survival analysis', *Diagnostic Histopathology* **22**(7), 263–270.

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., Suppes, R., Feinstein, D., Zanotti, S., Taiberg, L. et al. (2006), 'Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock', *Critical Care Medicine* **34**(6), 1589–1596.

Larché, J., Azoulay, É., Fieux, F., Mesnard, L., Moreau, D., Thiery, G., Darmon, M., Le Gall, J.-R. and Schlemmer, B. (2003), 'Improved survival of critically ill cancer patients with septic shock', *Intensive care medicine* **29**(10), 1688–1695.

Latha, N. A., Murthy, B. R. and Sunitha, U. (2012), 'Electronic health record', *International Journal of Engineering* **1**(10).

Lechner, M. et al. (2011), 'The estimation of causal effects by difference-in-difference methods', *Foundations and Trends in Econometrics* **4**(3), 165–224.

Levy, M. M., Dellinger, R. P., Townsend, S. R., Linde-Zwirble, W. T., Marshall, J. C., Bion, J., Schorr, C., Artigas, A., Ramsay, G., Beale, R., Parker, M. M., Gerlach, H., Reinhart, K., Silva, E., Harvey, M., Regan, S. and Angus, D. C. (2010), 'The surviving sepsis campaign: Results of an international guideline-based performance improvement program targeting severe sepsis.', *Critical Care Medicine* **38**(5), 367–74.

Lurio, J., Morrison, F. P., Pichardo, M., Berg, R., Buck, M. D., Wu, W., Kitson, K., Mostashari, F. and Calman, N. (2010), 'Using electronic health record alerts to provide public health situational awareness to clinicians', *Journal of the American Medical Informatics Association* **17**(2), 217–219.

Martin, G. S., Mannino, D. M. and Moss, M. (2006), 'The effect of age on the development and outcome of adult sepsis', *Critical Care Medicine* **34**(1), 15–21.

Mas-Colell, A., Whinston, M. D., Green, J. R. et al. (1995), *Microeconomic theory*, Vol. 1, Oxford University Press, New York, New York, USA, pp. 477–506.

Mathevon, T., Souweine, B., Traoré, O., Aublet, B. and Caillaud, D. (2002), 'ICU-acquired nosocomial infection: Impact of delay of adequate antibiotic treatment', *Scandinavian Journal of Infectious Diseases* **34**(11), 831–835.

Mellhammar, L., Wullt, S., Lindberg, Å., Lanbeck, P., Christensson, B. and Linder, A. (2016), 'Sepsis incidence: A population-based study', *Open Forum Infectious Diseases* **3**(4), ofw207.

Mouncey, P. R., Osborn, T. M., Power, G. S., Harrison, D. A., Sadique, M. Z., Grieve, R. D., Jahan, R., Harvey, S. E., Bell, D., Bion, J. F., Coats, T. J., Singer, M., Young, J. D. and Rowan, K. M. (2015), 'Trial of early, goal-directed resuscitation for septic shock', *New England Journal of Medicine* **372**(14), 1301–1311.

Nguyen, H. B., Corbett, S. W., Steele, R., Banta, J., Clark, R. T., Hayes, S. R., Edwards, J., Cho, T. W. and Wittlake, W. A. (2007), 'Implementation of a bundle of quality indicators for the early management of severe sepsis and septic shock is associated with decreased mortality', *Critical Care Medicine* **35**(4), 1105–1112.

NHS England (2015), 'Improving outcomes for patients with sepsis: A cross-system action plan', *London: NHS England* .

Padkin, A., Goldfrad, C., Brady, A. R., Young, D., Black, N. and Rowan, K. (2003), 'Epidemiology of severe sepsis occurring in the first 24 hrs in intensive care units in England, Wales, and Northern Ireland', *Critical Care Medicine* **31**(9), 2332–2338.

Park, S.-K., Shin, S. R., Hur, M., Kim, W. H., Oh, E.-A. and Lee, S. H. (2017), 'The effect of early goal-directed therapy for treatment of severe sepsis or septic shock: A systemic review and meta-analysis', *Critical Care Medicine* **38**, 115–122.

Pauly, M. V. (1968), 'The economics of moral hazard: Comment', *The American Economic Review* **58**(3), 531–537.

Pauly, M. V. (1974), 'Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection', *The Quarterly Journal of Economics* **88**(1), 44–62.

Prescott, H. C., Osterholzer, J. J., Langa, K. M., Angus, D. C. and Iwashyna, T. J. (2016), 'Late mortality after sepsis: propensity matched cohort study', *BMJ* **353**.

ProCESS Investigators (2014), 'A randomized trial of protocol-based care for early septic shock', *New England Journal of Medicine* **370**(18), 1683–1693.

Rhee, C., Gohil, S. and Klompas, M. (2014), 'Regulatory mandates for sepsis care — reasons for caution', *New England Journal of Medicine* **370**(18), 1673–1676.

Rhee, C., Murphy, M. V., Li, L., Platt, R. and Klompas, M. (2015), 'Comparison of trends in sepsis incidence and coding using administrative claims versus objective clinical data', *Clinical Infectious Diseases* **60**(1), 88–95.

Rivers, E., Nguyen, B., Havstad, S., Ressler, J., Muzzin, A., Knoblich, B., Peterson, E. and Tomlanovich, M. (2001), 'Early goal-directed therapy in the treatment of severe sepsis and septic shock', *New England Journal of Medicine* **345**(19), 1368–1377.

Robson, W. P. and Daniels, R. (2008), 'The sepsis six: Helping patients to survive sepsis', *British Journal of Nursing* **17**(1), 16–21.

Rodríguez, G. (2017), 'Generalized linear statistical models', `http://data.princeton.edu/wws509/notes/c7s1.html`. Accessed: 2018-04-06.

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J.-D., Coopersmith, C. M. et al. (2016), 'The third international consensus definitions for sepsis and septic shock (sepsis-3)', *Jama* **315**(8), 801–810.

Surviving Sepsis Campaign (2016), 'Surviving sepsis campaign bundles', `http://www.survivingsepsis.org/Bundles/Pages/default.aspx`. Accessed: 2018-04-26.

Terrell, K. M., Perkins, A. J., Dexter, P. R., Hui, S. L., Callahan, C. M. and Miller, D. K. (2009), 'Computerized decision support to reduce potentially inappropriate prescribing to older emergency department patients: A randomized, controlled trial', *Journal of the American Geriatrics Society* **57**(8), 1388–1394.

The Royal College of Pathologists (2018), 'Haematology', `https://www.rcpath.org/discover-pathology/i-want-a-career-studying/blood/haematology-careers.html`. Accessed: 2018-05-02.

Tiru, B., DiNino, E. K., Orenstein, A., Mailloux, P. T., Pesaturo, A., Gupta, A. and McGee, W. T. (2015), 'The economic and humanistic burden of severe sepsis', *Pharmacoeconomics* **33**(9), 925–937.

Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., Schrag, D., Takeuchi, M., Uyama, Y., Zhao, L. et al. (2014), 'Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis', *Journal of Clinical Oncology* **32**(22), 2380.

Van Der Sijs, H., Aarts, J., Vulto, A. and Berg, M. (2006), 'Overriding of drug safety alerts in computerized physician order entry', *Journal of the American Medical Informatics Association* **13**(2), 138–147.

Varian, H. R. (1992), *Microeconomic analysis*, W. W. Norton & Company, Inc., New York, New York, USA, pp. 441–470.

Vikman, S., Airaksinen, K. E. J., Tierala, I., Peuhkurinen, K., Majamaa-Voltti, K., Niemelä, M., Tuunanen, H., Nieminen, M. S. and Niemelä, K. (2004), 'Improved adherence to practice guidelines yields better outcome in high-risk patients with acute coronary syndrome without ST elevation: Findings from nationwide finacs studies', *Journal of Internal Medicine* **256**(4), 316–323.

Wang, H. E., Shapiro, N. I., Griffin, R., Safford, M. M., Judd, S. and Howard, G. (2012), 'Chronic medical conditions and risk of sepsis', *PLOS ONE* **7**(10), 1–7.

Westphal, G., Koenig, , Caldeira-Filho, M., Feijó, J., Trindade de Oliveira, L., Nunes, F., Fujiwara, K., Fonseca Martins, S. and Gonçalves, A. (2010), 'Reduced mortality after the implementation of a protocol for the early detection of severe sepsis', *Critical Care Medicine* **26**, 76–81.

Whiles, B., Deis, A., Miller, P. and Simpson, S. (2016), 'Comorbid conditions predict outcomes in patients with severe sepsis', *Chest* **149**(4), A170.

Whittaker, S.-A., Fuchs, B. D., Gaieski, D. F., Christie, J. D., Goyal, M., Meyer, N. J., Kean, C., Small, D. S., Bellamy, S. L. and Mikkelsen, M. E. (2015), 'Epidemiology

and outcomes in patients with severe sepsis admitted to the hospital wards', *Journal of Critical Care Medicine* **30**(1), 78–84.

WHO (2017), 'Improving the prevention, diagnosis and clinical management of sepsis', `https://www.global-sepsis-alliance.org/news/2017/5/26/wha-adopts-resolution-on-sepsis`. Accessed: 2018-04-07.

WHO (2018), 'Classification of diseases', `http://www.who.int/classifications/icd/en/`. Accessed: 2018-03-19.

Wolff, A. M., Taylor, S. A. and McCabe, J. F. (2004), 'Using checklists and reminders in clinical pathways to improve hospital inpatient care', *Medical Journal of Australia* **181**(8), 428.

Wolff, J. L., Starfield, B. and Anderson, G. (2002), 'Prevalence, expenditures, and complications of multiple chronic conditions in the elderly', *Archives of Internal Medicine* **162**(20), 2269–2276.

Wooldridge, J. M. (2013), *Introductory econometrics: A modern approach, 5th edition*, South-Western: Michigan State University Press, Mason, Ohio, USA, pp. 86–87, 93–102, 250–253.

Yang, Y., Yang, K. S., Hsann, Y. M., Lim, V. and Ong, B. C. (2010), 'The effect of comorbidity and age on hospital mortality and length of stay in patients with sepsis', *Critical Care Medicine* **25**(3), 398–405.

Yu, H., Chi, D., Wang, S. and Liu, B. (2016), 'Effect of early goal-directed therapy on mortality in patients with severe sepsis or septic shock: A meta-analysis of randomised controlled trials', *BMJ* **6**(3).

Appendices

A ICD-10 Codes

Table A.1 provides a list of selected ICD-10 codes indicating a sepsis related diagnoses.

Table A.1: List of selected ICD-10 codes

ICD-10	Description
A400	Streptococcal septicaemia
A401	Streptococcal septicaemia
A402	Streptococcal septicaemia
A403	Streptococcal septicaemia
A408	Streptococcal septicaemia
A409	Streptococcal septicaemia
A410	Other septicaemia
A411	Other septicaemia
A412	Other septicaemia
A413	Other septicaemia
A414	Other septicaemia
A415	Sepsis due to other Gram-negative organisms
A418	Other septicaemia
A419	Sepsis, unspecified
R650	Systemic Inflammatory Response Syndrome of infectious origin without organ failure
R651	Systemic Inflammatory Response Syndrome of infectious origin with organ failure
R652	Systemic Inflammatory Response Syndrome of non-infectious origin without organ failure
R653	Systemic Inflammatory Response Syndrome of non-infectious origin with organ failure
R659	Systemic Inflammatory Response Syndrome, unspecified
J039	Acute tonsillitis, unspecified
J180	Bronchopneumonia, unspecified
J181	Lobar pneumonia, unspecified
J189	Pneumonia, unspecified
J22X	Unspecified acute lower respiratory infection
J36X	Peritonsillar abscess
J440	Chronic obstructive pulmonary disease with acute lower respiratory infection
J690	Pneumonitis due to food and vomit
K358	Acute appendicitis, other and unspecified
K37X	Unspecified appendicitis
K610	Anal abscess
K631	Perforation of intestine (nontraumatic)
K800	Calculus of gallbladder with acute cholecystitis
K819	Cholecystitis, unspecified
L022	Cutaneous abscess, furuncle and carbuncle of trunk
L024	Cutaneous abscess, furuncle and carbuncle of limb
L031	Cellulitis of other parts of limb
L050	Pilonidal cyst with abscess
N12X	Tubulo-interstitial nephritis, not specified as acute or chronic
N390	Urinary tract infection, site not specified
N459	Orchitis, epididymitis and epididymo-orchitis without abscess
O234	Unspecified infection of urinary tract in pregnancy
T814	Infection following a procedure, not elsewhere classified
T845	Infection and inflammatory reaction due to internal joint prosthesis

B Robustness checks – large sample

Table B.1 reports the results of the OLS regressions on *clinicians' behaviour* based on the larger sample.

Table B.1: OLS results of clinicians' behaviour (large sample)

	(1)	(2)	(3)	(4)
	Time to first test order	Time to first test order	Frequency of test orders	Frequency of test orders
Active alert	-5.097	-0.093	0.530	-0.556**
	(4.222)	(1.564)	(0.545)	(0.248)
Age	0.240	0.247	0.133***	0.121**
	(0.233)	(0.231)	(0.049)	(0.048)
Male	15.263	17.472	2.412	1.664
	(11.552)	(11.479)	(1.953)	(1.947)
Male × Age	-0.493	-0.541	-0.051	-0.026
	(0.447)	(0.445)	(0.068)	(0.068)
age2	-0.002	-0.002	-0.001***	-0.001***
	(0.002)	(0.002)	(0.000)	(0.000)
Male × Age^2	0.004	0.004	0.000	0.000
	(0.004)	(0.004)	(0.001)	(0.001)
No. of comorbidities	0.493***	0.476**	0.335***	0.325***
	(0.184)	(0.187)	(0.030)	(0.030)
$\geq$ 4 comorbidities	-7.536***	-8.339***	-2.839***	-2.640***
	(2.384)	(2.338)	(0.502)	(0.456)
Constant	66.751***	72.939***	17.867***	16.322***
	(15.041)	(17.930)	(2.805)	(2.851)
Observations	2520	2520	2785	2785
Time fixed effects	yes	no	yes	no
Ward fixed effects	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.2 presents the results of the OLS regressions on *patient outcomes* based on the larger sample.

Table B.2: OLS results of patient outcomes (large sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	LOS	LOS	In-hospital mortality	In-hospital mortality	31-day mortality	31-day mortality
Active alert	1.451	-2.453***	0.005	-0.033***	0.004	-0.053***
	(1.159)	(0.526)	(0.023)	(0.013)	(0.027)	(0.015)
Age	0.198**	0.181*	-0.002	-0.003	-0.003	-0.003
	(0.099)	(0.100)	(0.002)	(0.002)	(0.002)	(0.002)
Male	6.277	5.071	0.108	0.086	0.052	0.036
	(3.861)	(3.855)	(0.073)	(0.073)	(0.079)	(0.078)
Male × Age	-0.169	-0.125	-0.004	-0.004	-0.002	-0.002
	(0.142)	(0.142)	(0.003)	(0.003)	(0.003)	(0.003)
Age^2	-0.002*	-0.001	0.000*	0.000**	0.000*	0.000*
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Male × Age^2	0.001	0.001	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
No. of comorbidities	0.882***	0.871***	0.015***	0.014***	0.017***	0.017***
	(0.064)	(0.064)	(0.001)	(0.001)	(0.001)	(0.001)
≥ 4 comorbidities	-4.023***	-3.772***	-0.098***	-0.096***	-0.079***	-0.081***
	(0.818)	(0.790)	(0.014)	(0.014)	(0.013)	(0.013)
Constant	89.598***	82.876***	-0.083	-0.111*	-0.107	-0.159**
	(26.481)	(27.037)	(0.071)	(0.065)	(0.078)	(0.068)
Observations	2698	2698	2698	2698	2617	2617
Time fixed effects	yes	yes	no	no	yes	no
Ward fixed effects	yes	yes	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3 shows the results of the Cox regression on *mortality* based on the larger sample.

Table B.3: Results of Cox-ph regressions (large sample)

	(1)	(2)
	Event: death	Event: death
Active alert	1.024	0.691***
	(0.191)	(0.088)
Age	1.069**	1.074**
	(0.03)	(0.030)
Male	0.749	0.812
	(1.011)	(1.090)
Male × Age	1.005	1.002
	(0.041)	(0.040)
Age^2	1.000	1.000*
	(0.000)	(0.000)
Male × Age^2	1.000	1.000
	(0.000)	(0.000)
No. of comorbidities	1.119***	1.111***
	(0.009)	(0.009)
≥ 4 comorbidities	0.827	0.737
	(0.264)	(0.222)
Observations	2775	2775
Time fixed effects	yes	no
Ward fixed effects	yes	yes
Ethnicity	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

C Robustness checks – small sample

Table C.1 reports the results of the OLS regressions on *clinicians' behaviour* based on the smaller sample.

Table C.1: OLS results of clinicians' behaviour (small sample)

	(1) Time to first test order	(2) Time to first test order	(3) Frequency of test orders	(4) Frequency of test orders
Active alert	12.617	1.551	1.289**	-0.013
	(13.168)	(1.718)	(0.546)	(0.214)
Age	0.417*	0.387*	0.055	0.057
	(0.229)	(0.216)	(0.041)	(0.042)
Male	8.510	6.379	0.971	0.663
	(10.569)	(10.511)	(1.841)	(1.856)
Male × Age	-0.299	-0.195	-0.026	-0.016
	(0.449)	(0.454)	(0.063)	(0.064)
Age^2	-0.004*	-0.003*	-0.001	-0.001
	(0.002)	(0.002)	(0.000)	(0.000)
Male × Age^2	0.002	0.001	0.000	0.000
	(0.004)	(0.004)	(0.001)	(0.001)
No. of comorbidities	0.415*	0.417*	0.246***	0.237***
	(0.212)	(0.228)	(0.028)	(0.028)
≥ 4 comorbidities	-4.249**	-4.279**	-1.517***	-1.644***
	(2.131)	(1.813)	(0.358)	(0.348)
Constant	0.314	-6.405	-0.985	-2.606**
	(6.728)	(5.545)	(1.200)	(1.165)
Observations	1524	1524	2074	2074
Time fixed effects	yes	yes	no	no
Ward fixed effects	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.2 presents the results of the OLS regressions on *patient outcomes* based on the smaller sample.

Table C.2: OLS results of patient outcomes (small sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	LOS	LOS	In-hospital mortality	In-hospital mortality	31-day mortality	31-day mortality
Active alert	1.141	-2.881***	-0.015	-0.030**	-0.001	-0.049***
	(1.598)	(0.531)	(0.031)	(0.013)	(0.035)	(0.015)
Age	0.079	0.071	-0.002	-0.003	-0.004	-0.004*
	(0.104)	(0.105)	(0.002)	(0.002)	(0.002)	(0.002)
Male	3.053	1.917	0.062	0.043	0.011	-0.018
	(4.073)	(4.106)	(0.077)	(0.076)	(0.085)	(0.085)
Male × Age	-0.087	-0.044	-0.003	-0.002	-0.000	0.001
	(0.150)	(0.151)	(0.003)	(0.003)	(0.004)	(0.004)
Age^2	-0.001	-0.001	0.000	0.000*	0.000**	0.000**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Male × Age^2	0.001	0.000	0.000	0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
No. of comorbidities	0.868***	0.863***	0.013***	0.013***	0.017***	0.016***
	(0.075)	(0.075)	(0.001)	(0.001)	(0.002)	(0.002)
≥ 4 comorbidities	-1.813**	-1.959***	-0.069***	-0.070***	-0.074***	-0.078***
	(0.744)	(0.726)	(0.014)	(0.013)	(0.015)	(0.014)
Constant	18.565***	15.061***	0.132	0.138*	0.165*	0.166**
	(3.793)	(3.729)	(0.083)	(0.075)	(0.092)	(0.082)
Observations	1961	1961	1961	1961	1961	1961
Time fixed effects	yes	no	yes	no	yes	no
Ward fixed effects	yes	yes	yes	yes	yes	yes
Ethnicity	yes	yes	yes	yes	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.3 shows the results of the Cox regression on *mortality* based on the smaller sample.

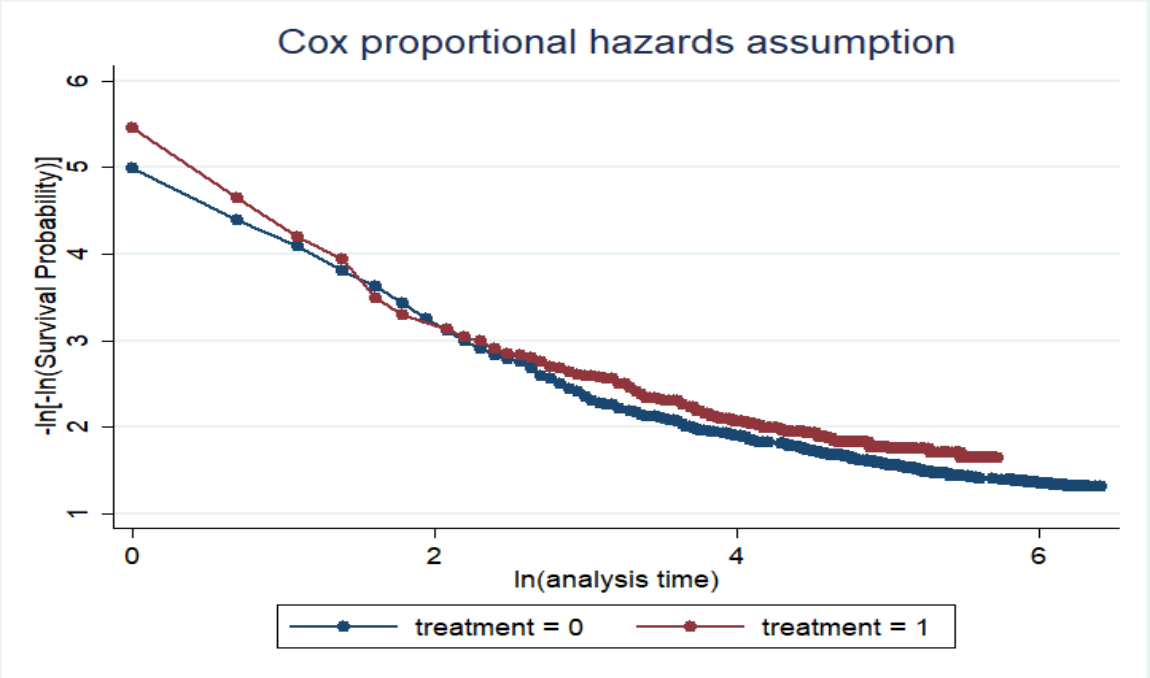
Table C.3: Results of Cox regressions (small sample)

	(1)	(2)
	Event: death	Event: death
ACtive alert	1.023	0.721**
	(0.251)	(0.095)
Age	1.09**	1.091**
	(0.042)	(0.041)
Male	2.806	1.96
	(4.72)	(3.276)
Male × Age	0.975	0.984
	(0.049)	(0.049)
Age^2	1.000*	1.000*
	(0.000)	(0.000)
Male × Age^2	1.000	1.000
	(0.000)	(0.000)
No. of comorbidities	1.125***	1.117***
	(0.011)	(0.010)
≥ 4 comorbidities	0.793	0.695
	(0.282)	(0.239)
Observations	2070	2070
Time fixed effects	yes	no
Ward fixed effects	yes	yes
Ethnicity	yes	yes

Note, coefficients and standard errors are rounded to three decimal places.
Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D Cox proportional hazards assumption

Figure D.1 exhibits graphical evidence on the plausibility of the Cox proportional hazards assumption in form of a "log-log" plot.



Source: Author's rendering of NHS data (2018).

Figure D.1: Cox proportional hazards assumption