Introduction
The implementation of rapid response systems (RRS) is based on the knowledge that deteriorating physiological processes are frequently present for hours or days before clear clinical deterioration is recognized [1,2]. It is assumed that this physiological deterioration is often treatable and that treatment will have greater effect when initiated early [3]. The RRS consists of an afferent limb, including “crisis detection” and “response triggering” and an efferent limb, the rapid response team (RRT) [4]. Even though robust evidence to support the effectiveness of the RRS is lacking [5-10] the system has been implemented worldwide. For example, Dutch hospitals are required to implement a patient safety programme including an RRS before 2013 [11]. This article explores the reasons why it is so difficult to prove the effectiveness of an RRS. We discuss the study designs that have been used and the various outcome measures in order to estimate the effects of an RRS. Finally, we make suggestions for future research.

Study design: how to find meaningful control groups?
Study designs used to estimate the effect of a treatment are the randomized controlled trial (RCT) and the non-randomized trial, the so-called quasi experiment [12]. The RCT usually has the most rigorous study design and the advantage of excluding potential bias due to heterogeneity and time trends. To date, the RCT design has been used only twice to estimate the effects of an RRS [13,14]. Both studies used cluster randomization at ward or hospital level, which of course has the disadvantage that e.g. bias due to heterogeneity in standard of care, patient groups, ward staffing ratios and ward staff expertise, cannot be fully eliminated. In addition, with randomization at hospital level, the heterogeneity of RRSs may also influence outcomes. These aspects make it extremely difficult to generalize the outcomes of both studies [13,14]. Due to heterogeneity, cluster randomization also requires the inclusion of a large number of wards or hospitals. The MERIT researchers estimated that over 100 hospitals were probably needed to show a 30% difference in the composite outcome cardiac arrest, unexpected death and unplanned ICU admissions [14]. Furthermore, since patient safety is an important topic in today’s media, increased awareness of the staff to recognize critically ill patients in the control wards

Table 1. Overview of disadvantages of available study designs

<table>
<thead>
<tr>
<th>DESIGN</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT at patient level</td>
<td>Practically impossible</td>
</tr>
<tr>
<td>RCT at ward level</td>
<td>Heterogeneity in standard care, patient groups, ward staffing ratios, ward staff expertise, increased awareness of ward staff on control wards concerning patient safety</td>
</tr>
<tr>
<td>RCT at hospital level</td>
<td>Heterogeneity in standard care, patient groups, ward staffing ratios, ward staff expertise, rapid response system procedures, composition rapid response teams, increased awareness of ward staff on control wards concerning patient safety</td>
</tr>
<tr>
<td>Quasi experiment in general</td>
<td>See RCT at ward level</td>
</tr>
<tr>
<td>Quasi experiment with the use of historical controls</td>
<td>See RCT at ward level, organizational changes such as ward staff ratios, ward staff expertise, improvement of medical treatment</td>
</tr>
<tr>
<td>Meta-analyses and reviews</td>
<td>Heterogeneity</td>
</tr>
</tbody>
</table>

Correspondence
F Simmes
E-mail: Friede.Simmes@HAN.nl
or hospitals might have influenced outcomes. While an RCT with randomization on a patient level would be the ideal design to solve these shortcomings, this is practically impossible to achieve.

Due to the aforementioned problems, the quasi experiment is a potential alternative. However, an important drawback of a quasi experiment is the non-randomized comparison of study groups [12]. The most frequently used quasi-experimental design to estimate the effects of an RRS is the one group before-after design, with the use of historical controls. Almost all those studies have been conducted in single hospitals. Several studies showed a positive effect on mortality [15-20] or incidence of cardiac arrest [15-17,19-23] whereas others found no effect on mortality [24-27] or cardiac arrest [25-28]. Although heterogeneity of patient categories can be partially controlled for, the use of historical control groups offers no way of controlling for other confounding factors, such as improvement in medical treatments and organizational changes [12]. One may therefore question whether the observed changes in outcome are actually due to the RRS.

Finally, systematic reviews and meta-analyses provide an objective method of integrating a number of study results and identifying patterns that otherwise might not have been detected [12]. The drawback of historical control groups was shown in a meta-analysis [9] of quasi-experimental studies as the magnitude of improvement in mortality, cardiac arrests and unplanned ICU admissions in the intervention groups was similar to the control group of the MERIT RCT study. Overall, to date, all reviews and meta-analyses found no or only weak support regarding the effectiveness of an RRS [5-10]. Table 1 shows an overview of the disadvantages of available study designs.

Outcome measures: how to find meaningful outcome measures?

Another reason why it may be difficult or even impossible to show the effectiveness of an RRS is that studies used a variety of outcome measures. The most frequently used outcome measures are the cardiac arrest rate, mortality rate, and number of (unplanned) intensive care unit (ICU) admissions. Unfortunately, the definition of cardiac arrest varies in regard to the type of arrest: cardiac arrest [21] and/or cardiopulmonary arrests [15,17,22,25,26] or cardiac arrest calls [19,28,29]. Also the location of cardiac arrest varies. Most studies used the hospital-wide cardiac arrest rate [7,16,17,19,20,25,26,28,29] thereby including places where the RRS is not active e.g. the operating theatre or the ICU. Others therefore used the out of ICU cardiac arrests [15,23], or cardiac arrests that occurred on the ward [14,22]. Several studies showed a reduction in cardiac arrest rate after the implementation of an RRS. However, this decreased incidence may also be the result of more patients being assigned a do not resuscitate order (DNR) [30-33]. A recent meta-analysis showed that a decline in cardiac arrest rates was not associated with lower hospital mortality [7].

Although the outcome measure mortality appears straightforward, definitions vary among studies. Most studies included all patients who died in the hospital [13,15-17,19,20,23-27,29,30]. Other studies excluded deaths in areas where the RRS was not active e.g. the operating theatre, the ICU or emergency areas [14,18]. However, ward patients may be referred to the ICU in a late stage of deterioration, and die in the ICU. This was the main reason why in our own study we did not exclude patients who died in the ICU following an unplanned IC admission from the ward [34].

Studies that did show a significant reduction in mortality had a high base line mortality incidence of 10 or more per 1000 admissions [13,16-18,23,25]. In the RCT by Priestley et al., baseline mortality was even 57 per 1000 admissions. It is obvious that a reduction in mortality is difficult to prove in settings with a lower baseline incidence. For example, since the baseline mortality rate in patients without a DNR order in our hospital was 3.6 per 1000, the observed decline of 50% of deaths without a DNR order was not statistically significant (Table 2) [34]. The third frequently used outcome measure is the incidence of unplanned ICU admissions. It was hypothesized that implementation of an RRS would decrease the incidence of unplanned ICU admissions due to timely detection and treatment of critically ill patients on the ward [4]. Unfortunately, definitions of ICU admission vary, as some studies included all (planned and unplanned) hospital ICU admissions [17,29] or ICU admissions only from general wards [21] whereas other studies limited inclusion to unplanned ICU admissions [18,35], or unplanned ICU admissions only from the general ward [14,26,36]. Overall, study results are inconclusive; both decreases, [17,21,26,36] no effect [14,28,35], and increases in ICU admissions [29,34] have been found. The hypothesis that the RRS decreases the number of unplanned ICU admissions is questionable, as more ward patients may be detected as

### Table 2. Deaths before and after implementation of an RRS (per 1000 admissions)

<table>
<thead>
<tr>
<th></th>
<th>BEFORE N=1376</th>
<th>AFTER N=2410</th>
<th>OR</th>
<th>95% CI FOR OR</th>
<th>P-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death without DNR</td>
<td>5 (0.36)</td>
<td>4 (0.17)</td>
<td>0.42</td>
<td>0.11–1.59</td>
<td>0.200</td>
</tr>
<tr>
<td>Death with DNR</td>
<td>9 (0.65)</td>
<td>19 (0.79)</td>
<td>1.05</td>
<td>0.46–2.40</td>
<td>0.900</td>
</tr>
</tbody>
</table>

ICU= intensive care unit  IQR= inter-quartile range  LOS=length of stay in days  OR= odds ratio  * Logistic regressions adjusted for age, gender and ASA  CI = confidence interval
critically ill and referred to the ICU. This could explain why we found an increased number of unplanned ICU admissions directly from the ward from 2.5% to 4.2% (OR 1.65, CI 1.07-2.55) after implementation of the RRS [34]. Table 3 shows an overview of what we know and do not know about the measured outcomes.

Remaining issues and future research
To reduce the incidence of cardiac arrests and unexpected mortality in ward patients, we need the timely detection and appropriate treatment of deteriorating patients. First, research is definitely needed on the accuracy and reliability of the ‘track and trigger’ systems, since the sensitivity of most current systems is low [32]. Pryterch et al. showed that using a VIEWS score of ≥5 as a trigger would result in a RRT call in 20% of all the observations, which implicates a substantial workload for the RRT team. However, this would only cover 82% of the deaths that would occur within 24 hours after the observation of the trigger [37]. Also the optimal monitoring frequency of the patient’s vital signs should be explored in more detail [38].

Second, if treatment is started by the ward staff and/or RRT, it would be interesting to analyze if this treatment is appropriate [39]. For example, a study showed inappropriate treatment by the ward staff, despite an accurate diagnosis in 88% (CI 64%-97%) of all preventable adverse events prior to the RRT call [40]. Our own study showed that 20% of the patients, who were referred to the ICU by the RRT, were initially treated by the RRT on the ward for one or two days [34]. This may partly explain why we did not observe a decrease in the median APACHE II score for unplanned ICU admissions after introduction of an RRS. One other study also reported APACHE scores and found no decrease in scores after introduction of an RRS [29]. Apparently, doctors are reluctant to admit a deteriorating patient to the ICU if they feel that he or she does not fulfil obvious admission criteria, like the need for respiratory or inotropic support.

Third, it is important to define the necessary skills of ward personnel [41] and/or responding personnel [39] in different ward or hospital settings. Other solutions for prompt recognition and treatment of deteriorating patients, rather than implementing a rapid response team, may suffice in particular health care settings [14,42,43]. For example, the Denver Health Medical Centre introduced the afferent arm only, including “crisis detection” and “response triggering”. A rapid response team was not introduced since shortage of qualified ward personnel was not a significant issue. Here the patients’ designated house staff delivers the majority of care. Introduction of this system resulted in a significant decrease of cardiopulmonary arrests [44].

Fourth, cost-effectiveness studies, including different aspects of recognition and treatment of critically ill patients, would be helpful in choosing the best interventions. For example, if the main results of RRS would be changes in circumstances of deaths, e.g. more deaths in patients with a DNR order versus deaths in patients without a DNR order, this raises the question whether other measures rather than implementing the complete RRS would suffice.

Finally, non-adherence of the ward staff to set procedures is of serious concern. Even when ‘track and trigger systems’ and

---

Table 3. Overview of what we know and not know about measured outcomes

<table>
<thead>
<tr>
<th>OUTCOME</th>
<th>WHAT DO WE KNOW</th>
<th>WHAT DO WE NOT KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac arrest</td>
<td>Unclear: several before-after studies found a positive effect, other studies, including one RCT at hospital level, found no effect.</td>
<td>Was the outcome influenced by&lt;br&gt;• organizational changes and/or improvement of medical treatment&lt;br&gt;• (some definitions) cardiac arrest calls without resuscitation&lt;br&gt;• (some definitions) cardiac arrest in places where the RRS was not operating?&lt;br&gt;• changes in DNR order policy</td>
</tr>
<tr>
<td>Mortality</td>
<td>Unclear: several before-after studies and one RCT on ward level one RCT on ward level and several before-after studies showed a positive effect, other studies, including an RCT at hospital level, found no effect.</td>
<td>Was the outcome influenced by&lt;br&gt;• heterogeneity between wards?&lt;br&gt;• organizational changes and/or improvement of medical treatment?&lt;br&gt;• (some definitions) mortality in places where the RRS was not operating?&lt;br&gt;• (when defined as deaths without a DNR order) an increase of deaths with a DNR order?&lt;br&gt;Did the outcome&lt;br&gt;• (some definitions) exclude patients who died on the ICU after an unplanned ICU admission?</td>
</tr>
<tr>
<td>ICU admissions</td>
<td>Unclear: several before-after studies showed a decrease, other studies, including one RCT at hospital level, found no effect, and some studies found an increase in ICU admissions.</td>
<td>Was the outcome influenced by&lt;br&gt;• organizational changes and/or improvement of medical treatment?&lt;br&gt;• (some definitions) unplanned ICU admissions from places where the RRS was not operating?&lt;br&gt;• (some definitions) planned ICU admissions on which the RRS has no influence?&lt;br&gt;Is the outcome reliable?&lt;br&gt;• increase of unplanned ICU admissions could be positive as this may be the result of early detection of critically ill patients and prevent patients from dying</td>
</tr>
</tbody>
</table>

RRS=rapid response systems, DNR=do not resuscitate, ICU=intensive care unit
Conclusion

Lack of adequate study designs to show the effectiveness of an RRS makes it almost impossible to show the effectiveness of an RRS. No financial support is declared for this study. Further research should therefore focus on the different aspects of implementing RRS and on the factors that influence the effectiveness of an RRS.

References

Why is it so difficult to prove that rapid response systems improve patient outcome? Directions for further research


