Quality improvement studies – pitfalls of the before and after study design

From a clinical perspective, the problem of postoperative residual paralysis is rather intriguing from several aspects. First and foremost is that it is a difficult phenomenon to define and quantify [1]. Far from being a straightforward response, recovery from pharmacological neuromuscular blockade occurs at different rates for different skeletal muscle, the order of which can vary from one individual to another. Clinical assessments of muscle strength are subjective and tends to be unreliable for assessment of residual paralysis [2]. Objective assessment parameters are indirect and non-linear, being an indicator of receptor occupancy and not of muscle strength. For example, the train-of-four ratio (TOFR) provides discriminating information of the final 30% of receptor occupancy but that does not reliably predict what may be happening in the pharyngeal muscles [3]. We tend to monitor neuromuscular function at sites that are accessible but have minimal clinical consequences, and muscles in which weakness can lead to significant morbidity are not readily accessible.

The pharmacology of neuromuscular blockade reversal is also not straightforward [4]. Prior to the availability of sugammadex, there were no direct agents for antagonising non-depolarising neuromusular blocking agents (NMBAs). Rather, we relied on the antagonism of cholinesterase to increase acetylcholine at the motor endplate in order to affect muscle cell depolorisation. Being reliant on an enzyme system that is not specific for the nicotinic receptor, the cholinesterase inhibitor is prone to a ceiling effect [5] and can induce muscinarinic adverse effects. In addition, neostigmine is known to have a paradoxical effect of increasing residual paralysis when given in excess to what is required. Sugammadex may have gone some way to circumvent these pharmacological complexities. In addition, this agent enables the use of large doses of rocuronium at induction to produce rapid onset and profound paralysis, even for short procedures, without the inconvenience of delayed or inadequate reversal of neuromuscular blockade. This can result in superior operative conditions for different types of surgical procedures [6,7]. On the other hand, the agent is expensive, not widely available, and is only efficacious for aminosteroid neuromuscular blocking agents. Therefore, the combination of neostigmine plus an antimuscarinic agent remain the mainstay pharmacotherapy for the antagonism of NMBAs in the forseeable future.

Standardising the management of reversal of muscle paralysis is no mean feat, as is any other endeavours that involve convincing a group of anaesthestists to unify their practice! In addition to it being a multistep process of assessing the TOFR and then undertaking the calculation of the correct dose, there may be some extra, hidden difficulties with this challenge. If seeing is believing, it may well be the case that anaesthestists are not “seeing” residual muscle paralysis as a previous survey highlights a gross discrepancy between what is estimated by the attending anaesthestist and what is objectively being measured [8]. Perhaps it may be the case that those muscles that when weak would contribute to delayed pulmonary complications are “hidden” from sight (for example, pharyngeal muscles) and the manifestation of weakness are subtle and difficult to observe. One

would have to perform an additional measurement that involves the direct application of a neuromuscular function monitor in order to detect weakness in these muscles. Most pulmonary complications are not obvious in the immediate peri-operative period and thus there is not a strong cognitive link between residual neuromuscular blockade and associated morbidities. Some clinicians may even be inherently sceptical as to whether TOFR monitoring can actually predict and/or prevent postoperative pulmonary complications. It is against this background that Rudolph et al [9] conducted their study to improve the peri-operative management of neuromuscular blockade and to evaluate its effects on postoperative pulmonary complications.

Randomised controlled trials (RTCs) are considered the most appropriate approach for assessing the effectiveness of interventions in the ideal world, but not all interventions can be practically assessed using this method. In addition, the external validity of clinical trials are often poor, and findings from RCTs can often not be generalised to real-world settings [10]. Observational studies can overcome some of the deficiencies of a clinical trial, but confounding variables and difficulties in establishing causation limit the conclusions which can be drawn from such trial designs.

This current study [9] is a prospective observational study comparing the incidence of major postoperative pulmonary complications before and after implementation of a quality improvement initiative. The primary outcome was a dichotomised composite measure of pneumonia, respiratory failure, pulmonary oedema, and/or tracheal intubation. It would not be ethically acceptable to randomise patients to be treated with, or without the quality improvement initiative, and so a quasi-experimental approach was used. Quasi-experiments are studies which aim to evaluate interventions without the use of randomisation and, like an RCT, aim to demonstrate causality between an intervention and an outcome [11]. The lack of randomisation is the major weakness of such a study design, but associations identified in quasi-experiments meet the requirement that the outcome of interest occurred after the intervention, and therefore, a quasi-experiment is not simply an observational study [12].

Threats to internal validity

Internal validity of an experimental outcome is defined as the degree to which observed changes in outcomes can be attributed to the intervention, and when applying a quasi-experimental design, it is important that threats to internal validity are reduced [12]. The first threat to internal validity is selection bias, where systematic differences in conditions in respondent characteristics could also cause the observed effect. The investigators in this current study have reduced this threat by conducting what in essence are two analyses, one on the whole cohort and one on a propensity matched cohort; this ensures that the patient characteristics in the pre-and-post-intervention cohorts are well-matched. The ‘calliper’ used by the investigators for matching is a strict 0.1. This means that the maximum tolerated difference between matched subjects in a "non-perfect" matching intention is set at 0.1 standard deviations of the propensity score. Results from the propensity

matched group confirmed the results seen in the whole cohort, which strengthens the finding that neostigmine dose used significantly decreased after the intervention.

Another threat to internal validity, known as maturation, is when naturally occurring changes over time is confused with an intervention effect [13]. The investigators have handled maturation by analysing the data using an interrupted time-series design. A time-series is a continuous sequence of observations on a population taken repeatedly over time. In an interrupted time series study, a time-series of the outcome is used to establish an underlying trend. The current study took measurements of weekly mean neostigmine dose per case from fifteen months prior to intervention to establish trends in changes. The hypothetical scenario, under which the intervention had not taken place and the baseline trend continues unchanged, is known as 'counterfactual', and the counterfactual scenario provides the comparison for the evaluation of the impact of the intervention.

Figure 1 is a reproduction of Figure 4a in the current study [9], with the added counterfactual line, showing the expected trend of mean neostigmine use had the intervention not taken place. As can be seen, the counterfactual line does not intersect with the post-intervention line, showing that post-intervention use of neostigmine did indeed decrease [12]. When applying the principles of interrupted time-series, it is important to remember that it would only be relevant if the intervention was not introduced gradually. The current study [9] did not strictly adhere to this requirement, because the change in neostigmine vial size was instituted three months before the department-wide educational activities and distribution of cognitive aids. As can be seen in Figure 1, there is a decrease in neostigmine use after week 66, when smaller aliquots sizes were introduced, and a further decrease after week 79, when cognitive aids and other educational activities were instigated.

Instrumentation, where the nature of a measurement may change over time and conditions, is another threat to internal validity. 'Instrumentation' may be an issue in the current study, where the 'instruments' are the clinicians documenting the neostigmine used and TOFR. During the intervention, clinicians were offered a financial bonus for documenting TOFR before neostigmine use, and this financial incentive may have resulted in superior documentation. A summary of the steps recommended for the implementation of a quality improvement study is given in Table 1.

In conclusion, this study has shown that a quality improvement intervention to improve the peri-operative management of neuromuscular blockade does reduce the number of postoperative pulmonary complications, but a financial incentive may be required to encourage the uptake of this new practice.

Table 1

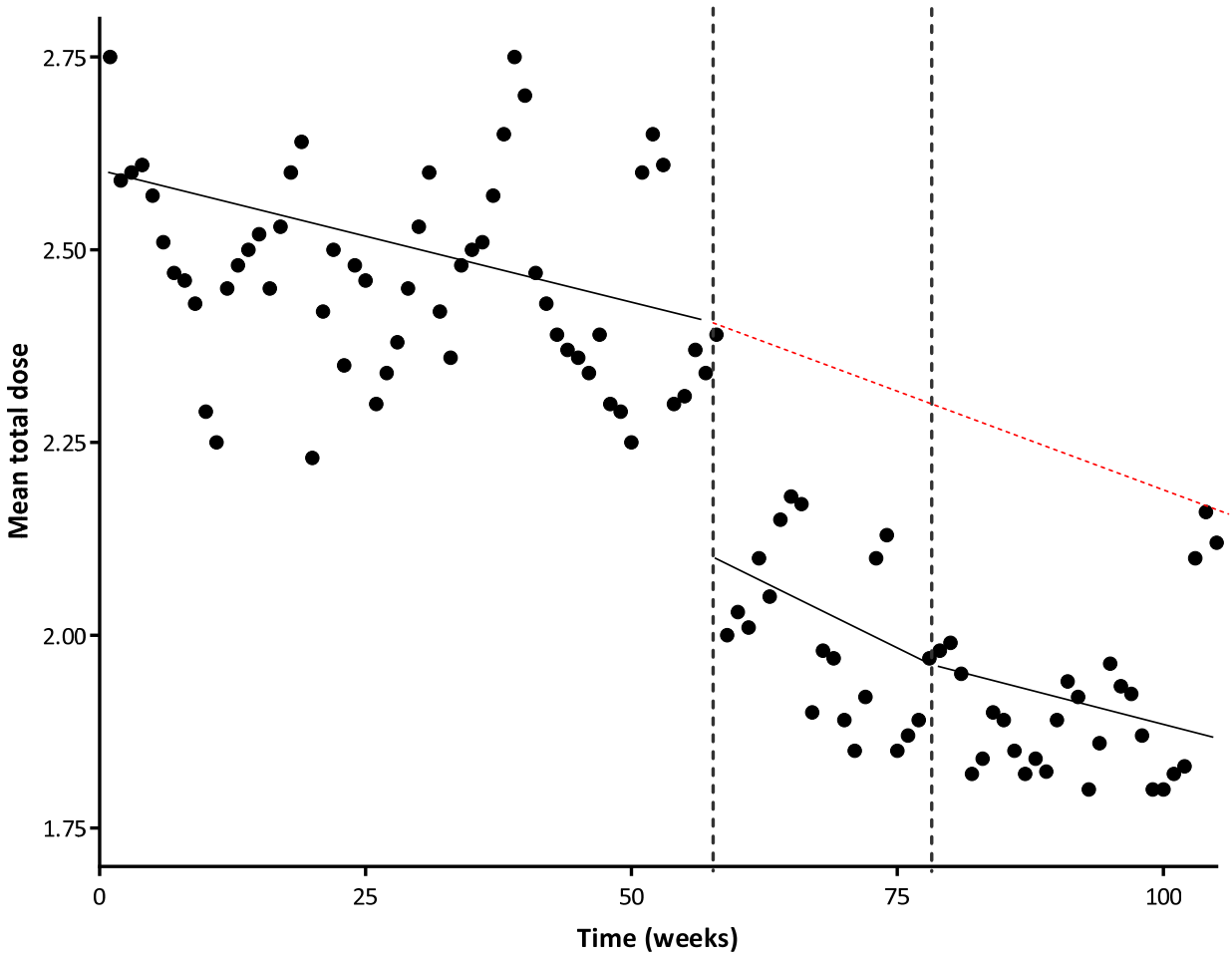
Steps in a Quality Improvement Study

Stage	Action
Planning stage	Outline objectives, timelines and accountability Define details of the proposed change and the implementation Make a prediction about the effect the change is expected to have, this should be done before study commences as it decreases the likelihood that an effect is detected due to random fluctuations or by chance Decide what to measure : a) *Outcome measures* – these occur at the patient level b) *Process measures* – such as measures of provision of appropriate services c) *Balancing measures* – these measures track unintended consequences of the intervention
Implementation stage	Implementation of the change and the provision of channels for those involved to provide qualitative feedback Analysis of the results should occur contemporaneously with the conduct of the project Interrupted time-series, which takes into consideration previous performance trends should be used A scatter plot of the time series, which can help to identify underlying trends, seasonal patterns and outliers should be constructed. Descriptive statistics should also be given Testing and identifying improvements and addressing methodological issues: – Controlling for seasonality and other longer-term trends, for example, stratifying by calendar months. At this time, data should be assessed for autocorrelation, (for example, by using the Durbin-Watson test). If autocorrelation does exist, this should be adjusted for using methods such as Prais regression -Consider the use of control groups where there might be time-varying confounders which are unmeasured or unknown, or adding additional phases so that the intervention is first introduced and then withdrawn to establish if there is a reversal of effect Several cycles of planning and implementation might be required before improvements are identified
Final stage	Sustain any positive changes

Legend

Figure 1: A reproduction of figure 4a [9]. The red dotted line shows the counterfactual scenario, depicting a decreasing baseline trend in neostigmine dose used. Since the counterfactual scenario does not cross the decreased concentrations reached after the intervention, it can be concluded that the intervention is a causative factor in decreased neostigmine dose.

Figure 1



Acknowledgements

SC is statistical advisor to Anaesthesia. No other competing interests declared.

S. W. Choi

Research Associate,

Oral and Maxillofacial Surgery,

The University of Hong Kong,

Hong Kong, HKSAR

Email: htswchoi@hku.hk

G. T. C. Wong

Clinical Associate Professor,

Department of Anaesthesiology,

Queen Mary Hospital,

Hong Kong, HKSAR

References

1. Srivastava A, Hunter JM. Reversal of neuromuscular block. *British Journal of Anaesthesia* 2009; **103:** 115-29.
2. Naguib M, Brull SJ, Johnson KB. Conceptual and technical insights into the basis of neuromuscular monitoring. *Anaesthesia* 2017; **72:** 16-37.
3. Murphy GS. Neuromuscular Monitoring in the Perioperative Period. *Anesthesia & Analgesia* 2018; **126:** 464-8.
4. Hristovska A-M, Duch P, Allingstrup M, Afshari A. The comparative efficacy and safety of sugammadex and neostigmine in reversing neuromuscular blockade in adults. A Cochrane systematic review with meta-analysis and trial sequential analysis. *Anaesthesia* 2018; **73:** 631-41.
5. Morasch KC, Aaron CL, Moon JE, Gordon RK. Physiological and neurobehavioral effects of cholinesterase inhibition in healthy adults. *Physiology & Behaviour* 2015; **138:** 165-72.
6. Choi ES, Oh AY, Koo BW, et al. Comparison of reversal with neostigmine of low-dose rocuronium vs. reversal with sugammadex of high-dose rocuronium for a short procedure. *Anaesthesia* 2017; **72:** 1185-90.
7. Koo BW, Oh AY, Na HS, et al. Effects of depth of neuromuscular block on surgical conditions during laparoscopic colorectal surgery: a randomised controlled trial. *Anaesthesia* 2018; doi: 10.1111/anae.14304.
8. Naguib M, Kopman AF, Lien CA, Hunter JM, Lopez A, Brull SJ. A survey of current management of neuromuscular block in the United States and Europe. *Anesthesia & Analgesia* 2010; **111:** 110-9.
9. Rudolph MI, Chitilian HV, Ng PY, et al. Implementation of a new strategy to improve the peri-operative management of neuromuscular blockade and its effects on postoperative pulmonary complications. *Anaesthesia* 2018; doi: 10.1111/anae.14326.
10. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *British Journal of Cancer* 2014; **110:** 551-5.
11. Harris AD, McGregor JC, Perencevich EN, et al. The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. *Journal of the American Medical Informatics Association* 2006; **13:** 16-23.
12. Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. *British Medical Journal* 2015; **350**:h2750
13. Penfold RB, Zhang F. Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements. *Academic Pediatrics* 2013; **13:** S38-S44.