This is a repository copy of *Can emergency medicine research benefit from adaptive design clinical trials?*.

Can emergency medicine research benefit from adaptive design clinical trials?

Laura Flight[1], Steven A Julious*[2] and Steve Goodacre[3]

1. Medical Statistics Group, School of Health and Related Research, University of Sheffield, Sheffield, England

2. Medical Statistics Group, School of Health and Related Research, University of Sheffield, Sheffield, England. Email: s.a.julious@sheffield.ac.uk Phone: Tel: +44 (0) 114 222 0709

3. Health Services Research, School of Health and Related Research, University of Sheffield, Sheffield, England.

Abstract

Background

Adaptive design clinical trials use pre-planned interim analyses to determine whether studies should be stopped or modified before recruitment is complete. Many emergency medicine trials are well-suited to these designs as many have a short time to primary outcome relative to the length of recruitment. We hypothesised that the majority of published emergency medicine trials have the potential to use a simple adaptive trial design.

Methods

We reviewed clinical trials published in three emergency medicine journals between January 2003 and December 2013. We determined the proportion that used an adaptive design, as well as the proportion that could have used a simple adaptive design based on the time to primary outcome and length of recruitment.

Results

Only 19 of 188 trials included in the review were considered to have used an adaptive trial design. A total of 154/165 trials that were fixed in design had the potential to use an adaptive design.

Conclusion

Currently there seems to be limited uptake in the use of adaptive trial designs in emergency medicine despite their potential benefits to save time and resources. Failing to take advantage of adaptive designs could be costly to patients and research. It is recommended that where practical and logistical considerations allow adaptive designs should be used for all emergency medicine clinical trials.

Introduction

Background

Traditionally a clinical trial is `fixed' in its design. All aspects of the design, conduct and analysis are outlined before a trial begins. (1) Once the trial has started the data collected are not examined until the trial has ended. An alternative is to use an adaptive design. A review of registered clinical trials by Hatfield *et al* (2016) (2) found over the last decade the number of trials using adaptive methods has increased from 11 per 10,000 registered trials between 2001 and 2005 to 38 per 10,000 registered trials between 2012 and 2013. Adaptive designs have great potential in emergency medicine to reduce the time, number of patients and resources needed to conduct clinical trials.

The Pharmaceutical Research and Manufacturers of America (PhRMA) working group define an adaptive design as using accumulating data to inform changes to the study as it continues, without undermining the 'validity and integrity of the trial'. (3) In an adaptive design data, collected as the trial progresses, are examined at interim analyses and the trial is modified or stopped if there is sufficient evidence to show that the treatment is efficacious, is not working (futility), or is harmful. (4) There are many types of adaptive designs. (5) One appropriate to emergency medicine is the group sequential design, (2) where the data are analysed after a pre-specified proportion of patients reach the primary outcome or amount of time has passed. (6) The study protocol should outline the motivations, the timing and the methods used at the interim analysis. This pre-specification maintains the validity and integrity of the adaptive clinical trial. A*d hoc* examinations of the data can introduce bias, (7) while many clinical trials with have some element of monitoring to ensure patient safety, adaptive designs are a more formalised process with pre-specified methods and rules guiding modifications.

As an example, the RATPAC (Randomised Assessment of Treatment using Panel Assay of Cardiac markers) study evaluated the use of point of care cardiac markers in the emergency department. (8) In a re-analysis by Sutton *et al* (2012) a group sequential design was used to examine the data at three monthly interim analyses. (9) A stopping rule was applied and a decision about whether to stop the trial early based on evidence the new treatment was working (efficacy) or not working (futility) was made. (6) The re-analysis showed that potentially by the second interim analysis data from more patients would be unlikely to change the outcome. The trial could have stopped for efficacy which would have avoided wasting resources (estimated to be $390,000), reduced the

ethical risks of continued recruitment of another 1521 patients and allowed the earlier reporting of trial findings.

Adaptive trials also allow us to update the uncertain assumptions made at the design stage of a trial when we have little information, (10) often the main motivation for the trial in the first place (11). For example, in the trial PRIMO trial that compared the effects of vitamin D in the treatment of patients with chronic kidney disease, (12) before the trial began there was little information available about the variability of their primary outcome (left ventricular mass index). However, by using an adaptive design they were able to analyse the data collected after 50% of patients had reached the outcome and re-calculate their sample size. (13) They found that their initial estimate of 220 patients was sufficient to achieve the desired power and the trial continued to the final analysis.

Adaptive trials work best when the time to the primary outcome is shorter than the total recruitment period of the study, because the researchers can intervene while patients are still being enrolled. (14). Emergency medicine clinical trials often have this short lag time. An interim analysis can be performed early in the recruitment window, leaving plenty of time to make adaptations to the trial before recruitment ends, as illustrated in Figure 1. The RATPAC trial had a primary outcome of successful discharge within four hours and with no return within the following three months. If successful discharge at four hours was used as an intermediate endpoint for the primary outcome, data are available on patients in the trial almost immediately. (9)

Goals of this investigation

This paper aimed to estimate the proportion of recently published emergency medicine trials that used an adaptive study design and the proportion that could have potentially used a simple adaptive design such as a group sequential design (GSD). (5)

Methods

We collected data from three emergency medicine journals:
1. Emergency Medicine Journal (EMJ),
2. Annals of Emergency Medicine (AnEM),
3. Academic Emergency Medicine (AcEM).

EMJ articles were identified by Gibb *et al* (2014) (15) and updated to include articles up to December 2013 by hand searching the journal. AnEM and AcEM articles were identified using Scopus and the

search terms "Clinical Trial" and "Randomised Controlled Trial". The following inclusion/exclusion criteria were used to select the articles adapted from Gibb *et al* (2014). (15)

1. Articles must report results from a randomised controlled trial, with a parallel group design, superiority primary endpoint and involve human trial subjects.

2. Subjects in the trial must have been individually randomised.

3. Articles must be the first published results from the trial in question.

4. Articles must be classified in the journal as original articles or pre-hospital care articles (EMJ only).

5. Articles must be published between January 2003 - December 2013.

6. Articles discussing non-clinical outcomes such as medical education studies and cost analyses are excluded.

8. After initial data extraction articles not reporting sufficient information about the length of recruitment or time to the primary outcome were excluded.

Data extracted included the length of recruitment, time to the primary outcome, sample size, whether the trial stopped early and the main results of the trial. The following definitions were used throughout the data extraction.

1. A trial reporting no examination of the data before it ended was considered a fixed design trial.

2. Any trials with an early, interim, examination of the data for any reason (including practical and safety reasons with no formal analysis of the data) before the trial terminated were identified and categorised as:

   a. Pre-planned – considered to be an adaptive design as it was clear the early examination was pre-specified and there was formal examination of the data.

   b. Unplanned – not considered to be an adaptive design as not clear that the interim analysis was pre-specified.

3. A trial was categorised as having stopped early if it was explicitly stated in the article. For fixed trials that stopped early, the motivation for early termination was investigated. We assumed, when these reasons were given, no analysis of the data was conducted prior to the decision to stop.

4. The length of recruitment was defined to be the length of time patients were recruited into the trial or the study duration. We assumed for trials in emergency medicine these two values are approximately equal. It is likely the study period ends shortly after the last patient has entered the trial.
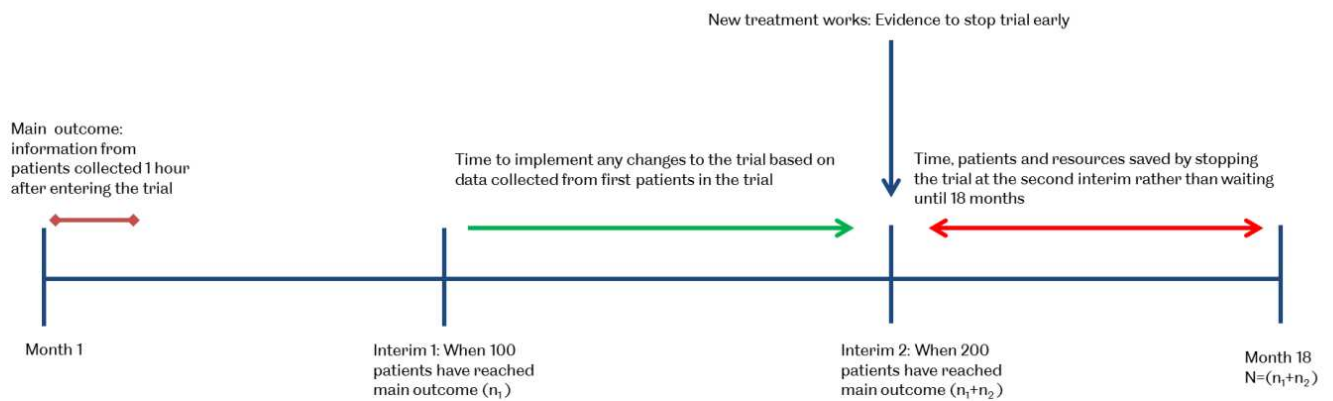
Figure 1 Illustration of the short time to the primary outcome relative to the length of recruitment that allows interim analyses in emergency medicine clinical trials where n_1 and n_2 are the number of patients in each treatment group

The corresponding author was contacted to request any missing information. Where possible we estimated missing data based on other information provided in the article. The database was summarised using trials with complete information only.

To determine which trials could have used a simple adaptive design, the length of recruitment (R) was compared to the time to the primary outcome (t). Trials with either of these variables missing were excluded. A decision rule was adopted from Sully *et al* (2014) (16) who stated there was little benefit to conducting a futility assessment of a clinical trial if $t/R \geq 0.25$. This was based on their proposal that the most desirable time to conduct a futility assessment is when 75% of the participants have been recruited. We believe this rule can be extended to any group sequential design, not just futility assessments, giving the following decision rule:

$$\textbf{(1)} \qquad \text{if } \frac{t}{R} < 0.25 \text{ had the potential to use an adaptive study design}$$

The analysis was mainly descriptive to provide an overview of the current use of adaptive designs in emergency medicine and their potential use. Summary statistics and graphical displays were used to summarise the variables, stratified by whether the trial data was examined early or not. Continuous variables were summarised using their mean, standard deviation (SD), median and inter-quartile range (IQR). Histograms were plotted. Categorical variables were summarised using counts, percentages and visually using bar plots.

Results

The CONSORT style diagram in Figure 2 illustrates the flow of articles through the review. Common reasons for exclusion were reporting the results of a systematic review or meta-analysis, discussion articles and observational studies. In total 216 articles were deemed to meet the inclusion criteria. AnEM contributed the largest number of articles to the review 94/216 (44%), followed by AcEM 80/216 (37%).

A total of 28 trials did not report sufficient information about the length of recruitment or time to the primary outcome. These trials were removed from the analyses as it was not possible to apply the decision rule (1). A total of 188 trials remained in the review.

The categorical variables are summarised in Table 1, stratified by whether the trial examined the data early (pre-planned or unplanned), or there was no examination of the data prior to the end of enrolment. The majority of trials were two armed, definitive studies with at least 80% power. There were some trials with missing power information but as this does not affect the application of the decision rule they remained in the analysis. Only a small number of non-adaptive trials stopped early (7/188). Reasons included resource constraints such as difficulties in obtaining drugs for the trial and slow recruitment.
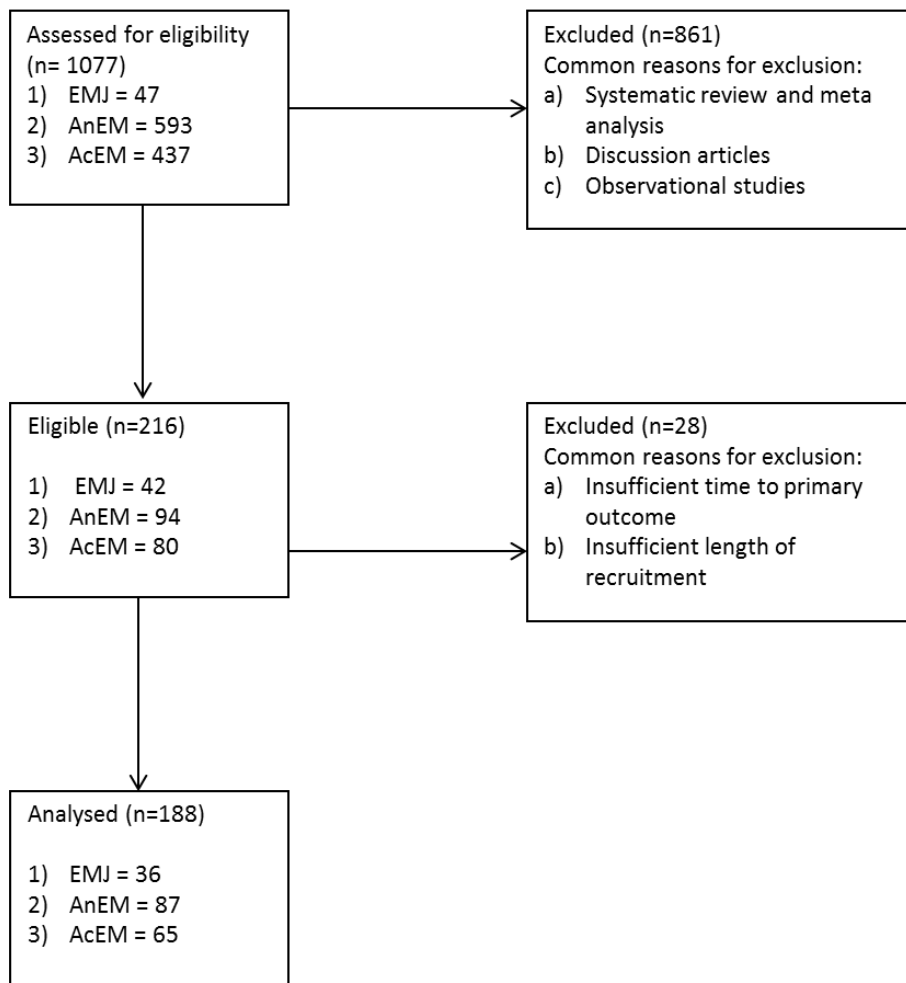
```
┌─────────────────────────────────┐              ┌─────────────────────────────────┐
│ Assessed for eligibility        │              │ Excluded (n=861)                │
│ (n= 1077)                       │─────────────▶│ Common reasons for exclusion:   │
│ 1)  EMJ = 47                    │              │ a)  Systematic review and meta  │
│ 2)  AnEM = 593                  │              │     analysis                    │
│ 3)  AcEM = 437                  │              │ b)  Discussion articles         │
│                                 │              │ c)  Observational studies       │
└─────────────────────────────────┘              └─────────────────────────────────┘
              │
              │
              ▼
┌─────────────────────────────────┐              ┌─────────────────────────────────┐
│ Eligible (n=216)                │              │ Excluded (n=28)                 │
│                                 │              │ Common reasons for exclusion:   │
│ 1)   EMJ = 42                   │─────────────▶│ a)  Insufficient time to primary│
│ 2)  AnEM = 94                   │              │     outcome                     │
│ 3)  AcEM = 80                   │              │ b)  Insufficient length of      │
│                                 │              │     recruitment                 │
└─────────────────────────────────┘              └─────────────────────────────────┘
              │
              │
              ▼
┌─────────────────────────────────┐
│ Analysed (n=188)                │
│                                 │
│ 1)  EMJ = 36                    │
│ 2)  AnEM = 87                   │
│ 3)  AcEM = 65                   │
│                                 │
└─────────────────────────────────┘
```

Figure 2 CONSORT style diagram of articles through the review of three emergency medicine journals

Characteristic		Examined Early (n=23)		Not Examined Early (n=165)		Total (n=188)	
		n	%	n	%	n	%
Power							
	80-89%	19	82.6	118	71.5	137	72.9
	≥90%	4	17.4	31	18.8	35	18.6
	NA	0	0	16	9.7	16	8.5
No. arms	2	23	100	137	83.0	160	85.1
	3	0	0	22	13.3	22	11.7
	4	0	0	6	3.6	6	3.2
Stopped Early	Y	14	60.9	7	4.2	21	11.2
	N	9	39.1	158	95.8	167	88.8
Pilot	Y	1	4.3	5	3.0	6	3.2
	N	22	95.7	160	97.0	182	96.8

Table 1 Summary of the categorical characteristics of the trials stratified by whether there was any early examinations of the data , whether pre-planned or unplanned.

Characteristic		Examined Early (n=23)	Not Examined Early (n=165)	Total (n=188)
Length of Recruitment (Days)	Median	639.0	427.0	474.0
	IQR	(422.5, 958.5)	(212.0, 669.0)	(242.8, 704.5)
	Mean (SD)	711.0 (409.49)	508.8 (373.99)	533.6 (383.15)
Time to Primary Outcome (Days)	Median	1.0	1.0	1.0
	IQR	(0.083, 1.5)	(0.042, 3.0)	(0.042, 3.0)
	Mean (SD)	4.4 (8.41)	24.7 (73.05)	22.2 (68.80)
Target Sample Size *	Median	124.0	115.0	120.0
	IQR	(99.0, 202.0)	(66.0, 247.0)	(73.0, 243.0)
	Mean (SD)	242.0	253.7	252.1

Table 2 Summary of the continuous characteristics of the trials stratified by whether there were any early examinations of the data where SD is the standard deviation and IQR is the interquartile range. *13 trials did not report a target sample size and so it was not possible to estimate the standard deviation.

Length of recruitment varied from 28 days to 1,919 days (Table 2). The histogram in Figure 3 shows the median was 474 days, equivalent to 1.3 years. On average the length of recruitment was much longer than the time it took patients to reach the primary outcome. The time to the primary outcome ranged from 5 minutes to 1 year, however, as the bar plot in Figure 4 demonstrates, the majority of trials 128/188 (68.1%) had a primary outcome measured over less than one day.
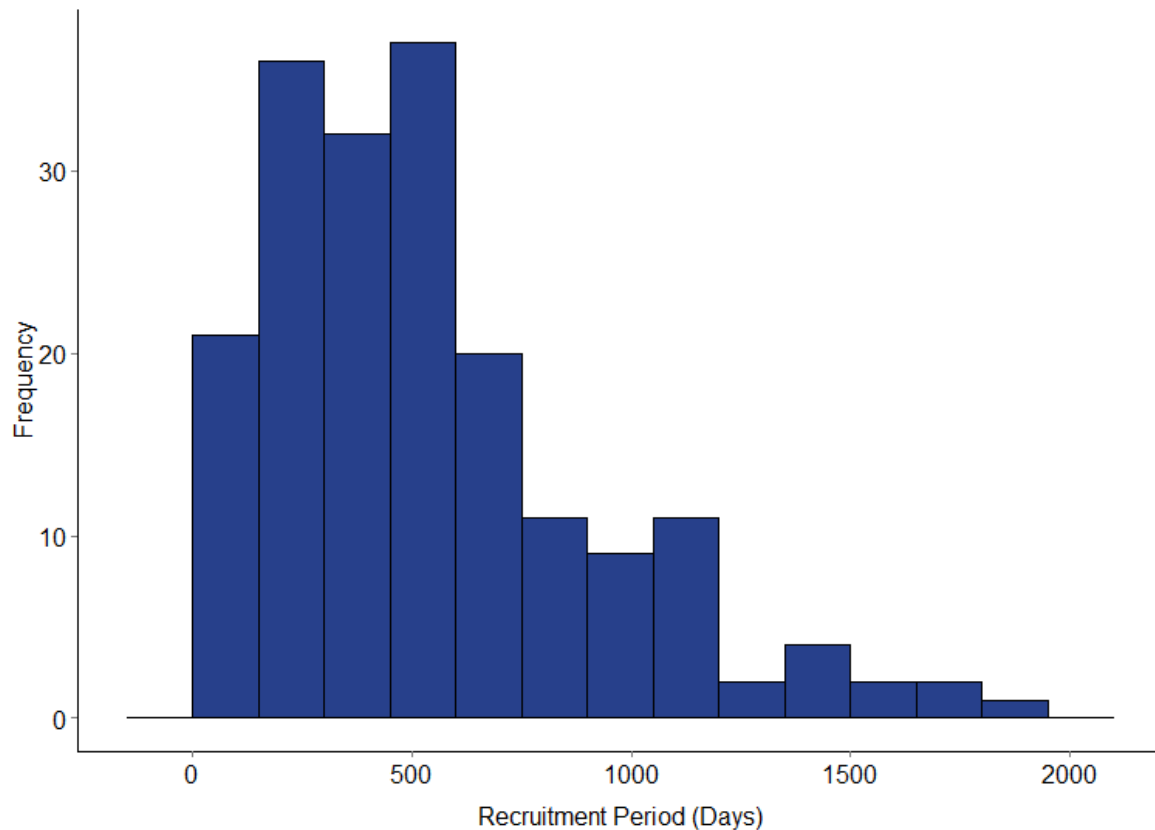
Figure 3 Histogram of the length of recruitment R for all trials in the review measured in days

Of the 188 trials, 23/188 (12.2%) analysed the data before the trial ended. Nineteen discussed pre-planned methodology and so are deemed to be an adaptive design. The four trials conducting unplanned interim analyses were mostly motivated by difficulties with patient recruitment and are not defined as adaptive designs.

A common reason for interim analyses in the studies with adaptive designs included safety checking (n = 6). For example in the trial by Messenger *et al* (2008) patients requiring sedation and pain relief for orthopaedic reduction or abscess drainage were randomised to subdissociative-dose Ketamine or Fentanyl. They conducted an interim safety analysis once 50% of the participants had been recruited (17). They found that their pre-specified stopping criteria were met and the data monitoring committee advised for recruitment to stop.

Ten of the studies used adaptive designs to evaluate futility and efficacy. As an example the trial by Auerbach *et al* (2009) compared Jet Lidocaine against Placebo in Children. They performed a blinded interim analysis for futility after 25% of the participants had been recruited in the first phase of their study. (18) The independent data monitoring committee decided that the trial should continue.
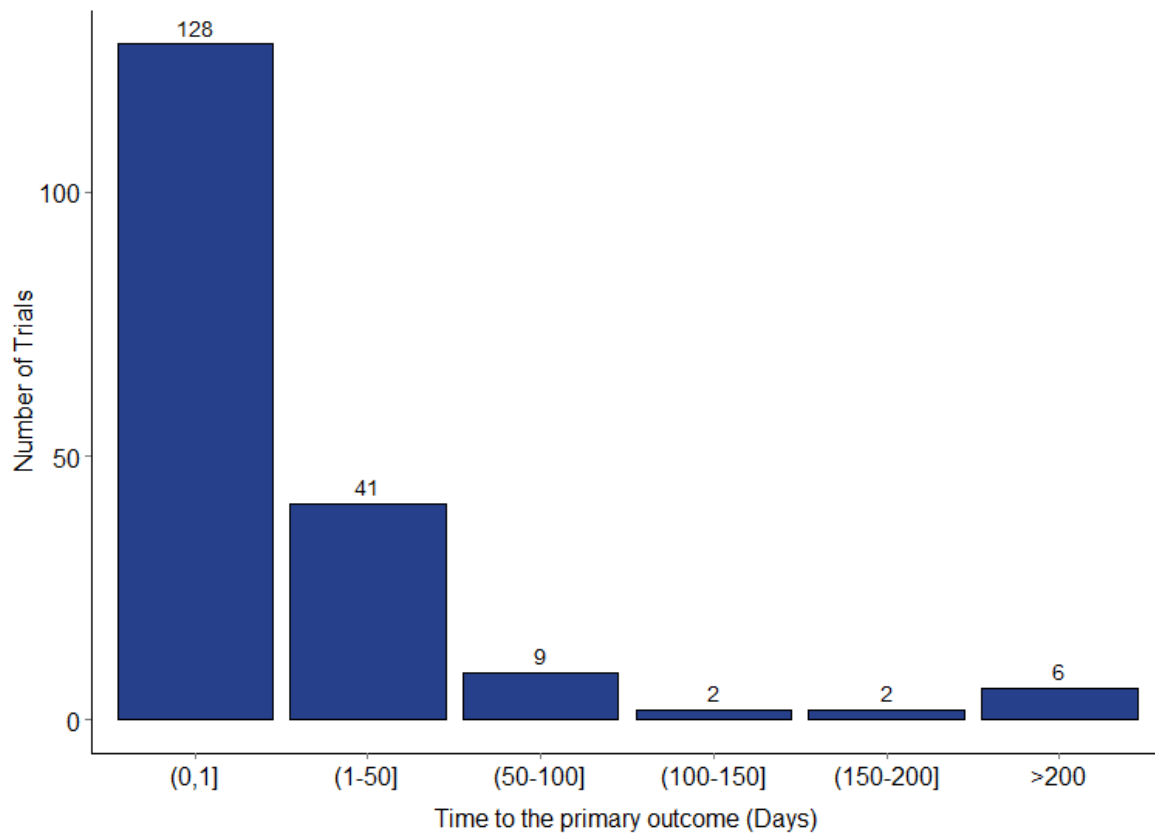
Figure 4 Barplot of the time to the primary outcome categorised into groups to illustrate the spread of the data

Of the remaining 165 fixed trials, we found that 154 (93.3%) could have used an adaptive design when applying the decision rule. (1) For 11 of the fixed trials (6.67%) an adaptive design could not have been used because by the time patients reached the primary outcome, the trial would have almost or completely finished recruiting leaving little to no time to implement any modifications to the trial.

Figure 5 Barplot of the number of trials that, could have been adaptive, could not have been adaptive, were adaptive and had an unplanned interim analysis in the review of three emergency medicine journals.

Discussion

Only 19/188 trials in the review were considered to have used an adaptive design and 154/165 (93.3%) of the fixed design trials had the potential to use an adaptive design based on the decision rule.

The limited uptake of adaptive designs in publicly funded trials has been explored by Dimairo *et al* (2016). (19) Key barriers included a lack of knowledge and expertise, few case studies and limited funding and support for the planning of adaptive designs. These barriers are likely to hinder the use of adaptive designs in emergency medicine.

Reservations towards adaptive designs have focussed on the potential to introduce bias. (5) However, the possible rewards of using adaptive designs in emergency medicine could be great. Research into the practical application of adaptive methods is increasing. This should encourage researchers to pursue adaptive designs and take advantage of the benefits they can offer. Appropriate planning and reporting of adaptive designs should allay any concerns. (7)

Good reporting of the adaptive methods used is important for a reader to determine the appropriateness of the analysis and to fully understand the results as outlined by the FDA draft guidance documents. (3,20) Bauer and Einfalt (2006) (21) make recommendations, which we endorse, that the following points should be included in the reporting of an adaptive design:

- The motivation for using an adaptive trial design
- The type of adaptations used
- A description of the trial that was initially planned
- Adaptations made to the trial including the reasons for this
- Estimates and confidence intervals with appropriate adjustments accounting for the adaptive nature of the trial.

In this review a trial was deemed to have been adaptive only if a pre-planned examination of the data was made before the trial terminated. Arguably when data monitoring committees examine data for safety outcomes, if no formal analysis is conducted and no adaptations made to the trial then this is not an adaptive design. In this review all pre-planned early examinations of the data were judged to be adaptive designs, perhaps as less formal monitoring was not considered important enough to include in the publication.

The PhRMA working group for adaptive designs (3) define an adaptive trials as making adaptations to the trial using data collected so far without undermining the `validity and integrity' of the trial. The FDA (7) state that these adaptations should be `prospectively planned' and hence unplanned, uncontrolled examinations of the data should be avoided as they are likely to impinge on the validity of the trial. Only those trials with pre-planned adaptive analyses are truly adaptive and unplanned interim analyses of data should be avoided.

The use of the decision rule adapted from Sully *et al* (2014) (16) relies solely on the length of recruitment and time to the primary outcome. This is a simplistic rule and, although it is a good indicator of whether a trial could have used adaptive methods, does not consider other factors that will influence whether adaptive methods can be implemented. For example, the anticipated rate of recruitment may influence whether an adaptive design is appropriate. A multi-centre trial where the time to the primary outcome is a quarter of the length of recruitment would be deemed suitable for an adaptive design using this rule. However, we might anticipate a large proportion of patients being recruited in the last quarter of the trial once the study is fully set up and all centres are recruiting.

Recruitment could potentially end before a sufficient number of patients have reached the primary outcome to conduct an interim analysis.

Additionally, the adaptive trial design may have undesirable consequences on other aspects of the trial. For example, a trial that stops early based on evidence from the primary outcome may not provide enough information to consider secondary outcomes such as the cost-effectiveness of the intervention. We recommend a full consideration of logistical as well as statistical factors when selecting the best trial design. There are a broad range of adaptations that can be made to a trial and new methodology is developing quickly. We suggest a simple group sequential design should be considered.

Limitations

In our study, it was not always clear whether adaptive methods had been used. Some authors included in detail the methods and motivations for interim inspections of the data; however, for others the rationale was less clear. If safety monitoring was conducted but not reported we will not have fully captured the extent to which data is monitored in emergency medicine clinical trials. We have made the assumption that when a trial is stopped early and a reason provided that no analysis was conducted to inform this decision. It is likely that analyses are conducted but not reported to inform this decision. The generalisability of the results are limited by the definitions and assumptions used.

The article selection and data extraction was conducted by one statistician (LF). A subjective decision was made about whether the trial should be included and whether the trial was adaptive. The decisions made should be consistent throughout; however, replication of this research might produce slight variations in the results. The main results and conclusions will be representative of the use of adaptive designs in emergency medicine clinical trials. A large amount of missing entries remained that were excluded from this analysis. Summaries of the trials excluded do not indicate any differences between compared to those kept in the review. The impact of excluding these trials on generalizability should be limited.

This review provides a `snap shot' of the use of adaptive designs in emergency medicine trials. By focusing solely on published trials in emergency medicine the results are vulnerable to publication bias. It is possible more emergency medicine trials are using adaptive methodology but these trials have been unsuccessful in publishing their work in one of the three journals.

Conclusions

There seems to be limited uptake in the use of these trial designs in emergency medicine despite almost all of the fixed trial designs being suitable for an adaptive study design. Failing to take advantage of adaptive designs has the potential to be costly to patients and research. It is recommended that where practical and logistical considerations allow adaptive designs should be used for all emergency medicine clinical trials.

What this paper adds

What is already known on this subject

- Adaptive methods are becoming increasingly popular in the design of clinical trials
- Often emergency medicine clinical trials require a short time for patients to reach the primary outcome relative to the amount of time patients are recruited making them well suited to adaptive designs
- It is not clear to what extent adaptive study designs are being used in emergency medicine

What this study adds

- Currently the use of adaptive study designs in the emergency medicine clinical trials sampled is limited
- In the sample of emergency medicine trials reviewed 93% had the potential to use adaptive methods based on our decision rule
- There is great potential for more emergency medicine studies to use an adaptive study design where logistical factors allow

References:

1. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human use. *Statistical principles for clinical trials E9.* 1998;

2. Hatfield I, Allison A, Julious SA, et al. Adaptive Designs undertaken in clinical research:a review of registered clinical trials. *Trials* 2016 19;17(1):150

3. Gallo P, Chuang-Stein C, Dragalin V, et al. Adaptive designs in clinical drug development--an Executive Summary of the PhRMA Working Group. *J Biopharm Stat* [Internet]. 2006;16(3):275–83.

4. Morgan CC, Huyck S, Jenkins M, et al. Adaptive Design: Results of 2012 Survey on Perception and Use. *Ther Innov Regul Sci* [Internet]. 2014;48(4):473–81

5. Chow SC, Chang M. Adaptive design methods in clinical trials–a review. Orphanet journal of rare diseases. 2008 May 2;3(1):1.

6. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. CRC Press; 1999 Sep 15.

7. U.S. Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics. 2010; Available from: http://www.fda.gov/downloads/Drugs/Guidances/ucm201790.pdf

8. Goodacre SW, Bradburn M, Cross E, The Randomised Assessment of Treatment using Panel Assay of Cardiac Markers (RATPAC) trial: a randomised controlled trial of point-of-care cardiac markers in the emergency department. *Heart*. 2011;97(3):190–6.

9. Sutton L, Julious S A, Goodacre SW. Influence of adaptive analysis on unnecessary patient recruitment: reanalysis of the RATPAC trial. *Ann Emerg Med*. Elsevier Inc.; 2012;60(4):442–8.e1.

10. Chow SC, Corey R. Benefits, challenges and obstacles of adaptive clinical trial designs. *Orphanet journal of rare diseases.* 2011 Nov 30;6(1):1.

11. Flight L, Julious SA. Practical guide to sample size calculations: an introduction. *Pharmaceutical Statistic*s. 2016 Jan 1;15(1):68-74.

12. Thadhani R, Appelbaum E, Pritchett Y, Chang Y, Wenger J, Tamez H, Bhan I, Agarwal R, Zoccali C, Wanner C, Lloyd-Jones D. Vitamin D therapy and cardiac structure and function

in patients with chronic kidney disease: the PRIMO randomized controlled trial. *Jama*. 2012 Feb 15;307(7):674-84.

13. Pritchett Y, Jemiai Y, Chang Y, Bhan I, Agarwal R, Zoccali C, Wanner C, Lloyd-Jones D, Cannata-Andía JB, Thompson T, Appelbaum E. The use of group sequential, information-based sample size re-estimation in the design of the PRIMO study of chronic kidney disease. Clinical Trials. 2011 Apr 1;8(2):165-74.

14.   Lewis RJ, Bessen HA. Sequential clinical trials in emergency medicine. *Ann Emerg Med* 1990;19(9):1047–53

15.   Gibb M, Julious SA. An audit of two journals to assess their adherence to CONSORT checklist criteria. School of Mathematics and Statistics, University of Sheffield; 2013.

16.   Sully BG, Julious SA, Nicholl J. An investigation of the impact of futility analysis in publicly funded trials. *Trials*. 2014 Feb 17;15(1):1..

 17.   Messenger DW, Murray HE, Dungey PE, et al. Subdissociative-dose ketamine versus fentanyl for analgesia during propofol procedural sedation: A randomized clinical trial. *Acad Emerg Med.* 2008;15(10):877–86.

18.   Auerbach M, Tunik M, Mojica M. A Randomized, Double-blind Controlled Study of Jet Lidocaine Compared to Jet Placebo for Pain Relief in Children Undergoing Needle Insertion in the Emergency Department. *Acad Emerg Med*. 2009 May;16(5):388–93.

19.   Dimairo M, Boote J, Julious SA, Nicholl JP, Todd S. Missing steps in a staircase: a qualitative study of the perspectives of key stakeholders on the use of adaptive designs in confirmatory trials. *Trials*. 2015 Sep 28;16(1):1.

20.   U.S. Food and Drug Administration. Draft Guidance for industry and Food and Drug Administration Staff: adaptive design clinical trials for dMedical Device Clinical Studies. 2015; Available from: http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446729.pdf19.

21.   Bauer P, Einfalt J. Application of Adaptive Designs – a Review. *Biometrical J* 2006 Aug;48(4):493–506