Resuscitation (2006) 71, 283—292

ELSEVIER

RESUSCITATION

www.elsevier.com/locate/resuscitation

CLINICAL PAPER

Quality of out-of-hospital cardiopulmonary resuscitation with real time automated feedback: A prospective interventional study☆

Jo Kramer-Johansen[a,b,c,*], Helge Myklebust[d], Lars Wik[a,c,e], Bob Fellows[f], Leif Svensson[g], Hallstein Sørebø[c], Petter Andreas Steen[a,h]

[a] *Institute for Experimental Medical Research, University of Oslo, N-0407 Oslo, Norway*
[b] *Norwegian Air Ambulance Foundation, Department of Research and Education, N-1341 Drøbak, Norway*
[c] *Division of Prehospital Medicine, Ullevål University Hospital, N-0407 Oslo, Norway*
[d] *Laerdal Medical, Tanke Svilands gt. 30, N-4002 Stavanger, Norway*
[e] *The National Competence Center for Emergency Medicine, Ullevål University Hospital, N-0407 Oslo, Norway*
[f] *London Ambulance Service NHS Trust, Kenton Training Centre, HA3 9TD Kingsbury, United Kingdom*
[g] *Kardiologkliniken, Södersjukhuset, 11883 Stockholm, Sweden*
[h] *Divisions of Surgery and Prehospital Medicine, Ullevål University Hospital, N-0407 Oslo, Norway*

Received 18 April 2006; received in revised form 24 May 2006; accepted 26 May 2006

KEYWORDS
Out-of-hospital CPR;
Chest compression;
Ventilation;
Automated external defibrillator (AED);
Advanced life support (ALS);
Clinical trials

Summary
Aims: To compare quality of CPR during out-of-hospital cardiac arrest with and without automated feedback.
Materials and methods: Consecutive adult, out-of-hospital cardiac arrests of all causes were studied. One hundred and seventy-six episodes (March 2002—October 2003) without feedback were compared to 108 episodes (October 2003—September 2004) where automatic feedback on CPR was given. Automated verbal and visual feedback was based on measured quality with a prototype defibrillator. Quality of CPR was the main outcome measure and survival was reported as specified in the protocol.
Results: Average compression depth increased from (mean ± S.D.) 34 ± 9 to 38 ± 6 mm (mean difference (95% CI) 4 (2, 6), $P < 0.001$), and median percentage of compressions with adequate depth (38—51 mm) increased from 24% to 53% ($P < 0.001$, Mann—Whitney *U*-test) with feedback. Mean compression rate decreased

☆ A Spanish translated version of the summary of this article appears as Appendix in the online version at doi:10.1016/j.resuscitation.2006.05.011.

* Corresponding author. Tel.: +47 23016819; fax: +47 23016799.
E-mail address: jo.kramer-johansen@medisin.uio.no (J. Kramer-Johansen).

doi:10.1016/j.resuscitation.2006.05.011

from 121 ± 18 to 109 ± 12 min^{-1} (difference −12 (−16, −9), $P = 0.001$). There were no changes in the mean number of ventilations per minute; 11 ± 5 min^{-1} versus 11 ± 4 min^{-1} (difference 0 (−1, 1), $P = 0.8$) or the fraction of time without chest compressions; 0.48 ± 0.18 versus 0.45 ± 0.17 (difference −0.03 (−0.08, 0.01), $P = 0.08$). With intention to treat analysis 7/241 control patients were discharged alive (2.9%) versus 5/117 with feedback (4.3%) (OR 1.5 (95% CI; 0.8, 3), $P = 0.2$). In a logistic regression analysis of all cases, witnessed arrest (OR 4.2 (95% CI; 1.6, 11), $P = 0.004$) and average compression depth (per mm increase) (OR 1.05 (95% CI; 1.01, 1.09), $P = 0.02$) were associated with rate of hospital admission.

Conclusions: Automatic feedback improved CPR quality in this prospective non-randomised study of out-of-hospital cardiac arrest. Increased compression depth was associated with increased short-term survival.

Trial registration: ClinicalTrials.gov (NCT00138996), http://www.clinicaltrials.gov/.

Introduction

Cardiopulmonary resuscitation (CPR) defined as chest compressions and ventilations are important for survival after cardiac arrest,[1] and quality of CPR influences outcome during basic life support.[2–4]

We have reported recently that quality of CPR during advanced life support (ALS) out-of-hospital did not adhere to the international guidelines,[5,6] when measured with new technology incorporated in standard defibrillators.[7] Chest compressions were not given 48% of the time without spontaneous circulation, and only mean 28% (median 24%) of the chest compressions had a depth of 38—51 mm (guidelines recommendation).[7]

Others have reported similar findings with the same equipment in in-hospital arrests,[8] and the results correspond with those found with other methods of quality measurements during out-of-hospital arrest in first responders and paramedics.[9–11]

Automated verbal feedback consistently has been reported to improve quality of CPR during manikin training,[12,13] with faster recovery of CPR skills when tested 6 and 12 months later.[14,15] Our recent publication on quality of CPR during ALS was a planned baseline-period before studying similar feedback via the defibrillator during CPR on patients. The hypothesis was that addition of such feedback would reactivate CPR skills and improve quality.

Methods

Except for the feedback system and statistical comparisons between the two groups without and with feedback, all details of the methods in this study have been reported in our recent publication,[7] and a condensed version is therefore presented here.

Study design and recruitment

The study was designed as non-randomised because feedback-induced increased awareness of quality problems could spill over to the cases without feedback, thus potentially improving the quality of CPR also in the control group. The study was approved by the regional ethics committees for Akershus (Norway), Stockholm (Sweden), and London (England). Informed consent for inclusion in the study was waived as decided by these committees in accordance with paragraph 26 in the Helsinki Declaration.[16] In this prospective study registered at ClinicalTrials.gov (NCT00138996, Initial release 29th August 2005) patients older than 18 years suffering from out-of-hospital cardiac arrests of all causes were included. From March 2002 to October 2003 quality of CPR was recorded from the defibrillators without feedback to the rescuers. From October 2003 to September 2004 feedback on CPR quality via the defibrillators was activated. No information of the results from the first period was given to the rescuers in any of the involved services or presented in any professional forum until the second period was terminated.

Equipment

Prototype defibrillators based on standard Heartstart 4000 (Philips Medical Systems, Andover, MA, USA) defibrillators were deployed in six ambulances at each site. The defibrillators were approved for investigational use in Europe by DNV (CE-mark; 2002-OSL-MDD-0009) and in the US by FDA (IDE# G020121). The defibrillators had an extra chest pad to be mounted on the lower part of the sternum with double adhesive tape. The pad was fitted with an accelerometer (ADXL202e, Analog Devices, USA) and a pressure sensor (22PCCFBG6, Honeywell, USA). The heel of the rescuer's hand should

be placed on top of the chest pad, and its movement was considered equal to that of the sternum during chest compressions with compression force >2 kg. A second accelerometer within the defibrillator allowed cancelling out vertical motion of the whole patient or supporting surface. The method has been validated in a manikin model.[17] Trans-thoracic impedance was measured by applying a near constant sinusoidal current across the standard defibrillation pads and accelerometer and impedance signals were stored in an extra data card in the defibrillators.[18]

CPR feedback was based on measured quality compared to set standards (Table 1) based on international guidelines and the expert opinion of the researchers. Feedback consisted of both verbal messages in the national language and waveforms on an extra LCD display on the defibrillator (visual feedback). All feedback prompts with time codes were saved to the extra data card. The paramedics were encouraged to fill out written evaluation forms after each CPR episode. Halfway through the intervention period we modified the feedback rules. Findings from the baseline period made us at that time aware that inactivity seemed to be a major problem. A verbal prompt was added after 30, 45, and 60 s without CPR. Based on responses from the users a time buffer was added to avoid overwhelming the paramedics with messages, especially reducing repetitive prompts. Incorrect

Table 1 Target values and priority rules for automatic verbal feedback versions 1 and 2

	Prompts given when	Priority version 1	Priority version 2	Comment
Compressions				
Leaning	>4 kg pad pressure between compressions	1	1	Incomplete release of chest compression
Depth	<38 mm	2	2	
Rate	<90 min^{-1}	3	3	
	>120 min^{-1}	3	3	
Depth	>52 mm	4	4	
Duty cycle	<30%	5	5	
Ventilations				
Impedance change	<0.8 Ω	1	1	Only verbal feedback if <0.8 Ω in version 2
	<1.1 Ω	1	X	
Inspiration time	<0.8 s	2	2	Only before intubation was indicated by pressing a button on the screen
Rate	<6	3	3	
	>16	3	3	
Inspiration time	>2.0 s	4	4	
Pauses/pattern				
Time without compressions	>15 s	1	1	Tonal prompt.
	>30 s	1	1	Changed to verbal prompt in version 2
	>45 s and >60 s	X	1	Additional verbal prompts at 45 and 60 s introduced in version 2
Time without ventilation	>30 s	X	1	Introduced in feedback version 2. If both compressions and ventilations were lacking, a verbal prompt addressing both were given
Change from compressions to ventilation	Only prompts on ventilations and compressions in a 2:15 pattern before intubation was indicated by pressing a button on the screen		X	Removed in feedback version 2

ventilations, which had been a large problem in manikin studies,[13–15] was not so in the baseline period, and was therefore given lower priority. The changes are summarised in Table 1.

Training and treatment protocol

All ambulances were staffed with paramedics trained once a year in ALS according to standard international guidelines.[5,6] Only Stockholm had a two-tiered system with a nurse anaesthetist in the second car attending cardiac arrests. After all personnel had been trained in the use and meaning of the feedback software, concurrent with yearly retraining, feedback was implemented in October 2003 (London and Stockholm) and January 2004 (Akershus). During the whole data collection period ambulance personnel in Akershus used the defibrillator in manual mode and used a modified CPR protocol with 3 min of CPR before the first DC shock and between unsuccessful series of up to three DC shocks.[19] No involved hospital had implemented post-resuscitation therapeutic hypothermia during the data collection period.

Data collection and processing

Data from each episode included scanned patient report forms and locally adapted versions of Utstein style forms. ECG-signals, time signal, events, accelerometer signals, and trans-thoracic impedance were collected from the defibrillators. Each case was viewed and annotated with a custom PC programme designed for this study (Sister Studio, Laerdal Medical, Stavanger, Norway). Annotations were scrutinised and corrected manually if needed by consensus of one of the authors (JK-J) and one engineer with in-depth knowledge of the sampling technology.

Each compression was coded as too deep, too shallow, or acceptable according to the goals in Table 1. Compression part of the duty cycle was defined as the fraction of time with subzero position of the chest pad for each compression. Residual force between compressions exceeding 4 kg was coded as incomplete release. No flow time (NFT) was defined as all pauses between compressions longer than 1.5 s. The sum of such intervals was divided by segment length and represents the fraction of time without circulation (no flow ratio (NFR)). NFR adjusted (NFR_{adj}) for allowable time for rhythm analysis, defibrillator charging and shock delivery, and pulse checks was calculated as described previously.[7] Ventilations were detected by changes in thoracic impedance corrected for compression and blood flow related signals.[18]

The actual number of compressions and ventilations per minute are presented as well as NFR, NFR_{adj} and compression characteristics as described above. Average values can obscure the existence of short time segments with very high or low values; we therefore also report the number of ventilations for all 1 min segments.

Outcome measure

Primary outcome was change in quality variables after introduction of automated verbal feedback. Target values are given in Table 1. Secondary outcomes were rate of hospital admission with spontaneous circulation and survival to hospital discharge with neurological outcome.

Statistical analyses

In the baseline period 176 patients had been included for quality analysis.[7] Sample size for the feedback period was set at a minimum of 100 patients. This was based on calculations with Sample Power 2.0 (SPSS Inc.) with desired power 0.85 and alpha 0.05. An increase in the average chest compression depth to within guidelines from 34 ± 9 to 38 mm required a total of 246 (176 versus 70) patients, and a reduction in no flow ratio from 0.48 ± 0.18 to 0.38 a total of 211 (176 versus 35) patients (equal variances from baseline-period assumed for the two phases).

Data were collected and organised using a spreadsheet program (Excel 2003, Microsoft Corp., Redmond, WA) and statistical analyses performed with SPSS for Windows (SPSS ver. 12.0, Chicago, IL) by one of the authors (JK-J). Confidence intervals for medians were calculated using normal approximation described by Altman.[20] Unless otherwise stated, results were expressed as means $\pm$ standard deviation (S.D.), implying close to normally distributed data. Differences were reported as means with 95% confidence intervals (95% CI). Comparisons of continuous data were done with independent samples *t*-test or Mann–Whitney *U*-test as appropriate, and comparisons of proportions were made with odds ratios (OR) with 95% CI with *P*-values from χ^2-test with continuity correction. Two-sided *P*-values less than 0.05 were considered significant. A logistic regression model assessed admission to hospital as a dependent factor and quality variables as independent factors in addition to factors previously reported to influence survival (time from ambulance dispatch to ambulance crew arrives at patient (response time)), witnessed arrest (yes/no), bystander CPR (yes/no), place of arrest (public/private), and initial rhythm (VF/non-VF).[21] Sex and age were also included in the model.

A forward stepwise approach using likelihood ratios with cut-offs at 0.15 in and 0.20 out were used and *P*-values less than 0.05 were considered significant. The resulting model of logistic regression was then applied to all episodes. The factors previously described were also introduced into the final model. Admission to hospital was chosen as a dependent factor as this was thought to reflect out-of-hospital treatment and not differences in in-hospital treatment.[22]

Results

The annual statistics and demographic data for the three emergency medical service systems have been described previously.[7] Two hundred and forty-three episodes were collected in the baseline period and 120 in the feedback period. Two patients during baseline and three with feedback did not receive CPR and were excluded leaving a total of 241 and 117, respectively. For quality of CPR analysis 65 additional patients were excluded from baseline (27%) and 9 (8%) from feedback. This was due to technical problems in 27 and 5, and failure to apply the extra chest pad in 38 and 4, respectively. The resulting number of episodes with quality data on compressions was 176/241 (73%) in baseline group and 108/117 (92%) in feedback group; OR for completeness of data 4.4 (2.1, 9.2). Due to suboptimal signal quality during parts of the episode ventilation count could only be determined for 163/176 (93%) in baseline group and 98/108 (91%) in feedback group (OR 0.8 (0.3, 1.9)).

There were no differences in demographic and resuscitation episode characteristics between the two periods (Table 2). This also applies to each site analysed separately (data not shown).

Performance of compressions changed with introduction of feedback (Table 3). Average compression depth increased significantly, and the percentage of compressions with correct depth doubled. Average compression rate, which tended to be higher than guidelines recommendations in the baseline period, fell. Average compression part of duty cycle and average number of ventilations per minute was within target during baseline, and did not change. The proportion of 1 min segments with excessively high ventilation rates (>21 min^{-1}) was reduced significantly after introduction of feedback from 396/4109 (10%) to 170/2274 (7.5%) (OR 0.8 (0.6, 0.9), $P = 0.004$). Mean no flow ratio was 0.48 in baseline and 0.44 with feedback ($P = 0.08$).

As described in the Methods section, the feedback rules were changed during the intervention period towards more weight on inactivity of chest compressions (Table 1). Table 4 summarises CPR quality with the two versions of feedback software. The number of compressions per minute increased from 60 to 69 due to reduction of the mean fraction of time without chest compressions from 0.47 to 0.40. This apparently occurred at the expense of compression depth which decreased from mean 39 to 36 mm with the new feedback priorities.

Incomplete release of force on the chest pad between compressions was a minor problem in both phases and was detected at 10,985/373,390 (3%) of the compressions. The median fractions of compressions with incomplete release were still below 1% (Table 3) suggesting that this was a significant problem in some episodes. Indeed episodes with more than 10% incomplete release accounts for more than 50% of these compressions and the number of such episodes were 15/176 and 7/108 in the baseline and feedback group.

The great majority of rescuer comments on the feedback software were positive; 89/103 (86%) indicated that they felt it helped them perform better CPR. Only 3 of 103 (3%) evaluation forms included negative comments from bystanders versus 10/103 (10%) that reported positive comments from bystanders. At the discretion of the ambulance personnel in charge it was possible to turn down the volume of the feedback, switch to tonal prompts only or turn off audible feedback altogether. While 19/108 (18%) chose to turn off audible feedback before the end of the resuscitation episode (mean time from start of monitoring; 7 min), visual feedback on the screen was

Table 2 Demographic and Utstein characteristics of baseline and feedback cohorts

	All sites	
	Baseline	Feedback
Episodes, *N*	241	117
Age	68 ± 15	68 ± 14
Males (%)	172 (71)	80 (68)
Usable (%)	176 (73)	108 (92)
Ambulance witnessed (%)	18 (7)	12 (10)
Bystander witnessed (%)	160 (72)	73 (70)
Bystander CPR (%)	94 (42)	45 (43)
Response time (min)	8 (7, 8)	7 (6, 8)
Shocks per episode	1 (1, 2)	1 (0, 1)
Episodes with ≥1 shock (%)	139 (58)	62 (53)
Shocks per episode in episodes with ≥1 shock	4 (3, 6)	2 (2, 4)

All variables gives as numbers (percentages in parenthesis) except age (mean ± S.D.), response times (minutes, mean with 95% CI) and shocks per episode (median with 95% CI). Ambulance personnel witnessed cases are not included in bystander witnessed, bystander CPR, and response time calculations.

Table 3 Performance of CPR in baseline and feedback groups

	Baseline ($n = 176$)	Feedback ($n = 108$)	Mean difference (95% CI)	*P*-value
No flow				
NFR	0.48 ± 0.18	0.44 ± 0.17	0.04 (−0.01, 0.08)	0.08
NFR_{adj}	0.39 ± 0.17	0.37 ± 0.16	0.02 (−0.02, 0.06)	0.3
Compressions				
Compressions (min^{-1})	64 ± 23	63 ± 21	1 (−4, 7)	0.5
Compression rate (min^{-1})	121 ± 18	109 ± 12	12 (9, 16)	<.001
Depth (mm)	34 ± 9	38 ± 6	−4 (−6, −2)	<.001
Depth 38—51 mm (%)	24 (19, 31)	53 (45, 57)		<.001
Too deep (>51 mm) (%)	0 (0, 1)	1 (0, 3)		0.01
Too shallow (<38 mm) (%)	71 (66, 78)	41 (30, 50)		<.001
Incomplete release (%)	0 (0, 1)	0 (0, 1)		0.08
Compression as part of duty cycle (%)	42 ± 4	41 ± 4	0 (−1, 1)	0.4
Ventilations ($n = 163$ and 98, respectively)				
Ventilations (min^{-1})	11 ± 4.8	11 ± 4.0	0 (−1, 1)	0.8

Values given as mean ± S.D. and differences as mean difference with 95% confidence interval (CI) and *P*-values for difference not equal to 0 from two-sided independent samples *t*-test, except for percentages of compressions with depth 38—51 mm, too deep, too shallow and with incomplete release which are given as median values with 95% CI and *P*-values of difference not equal to 0 with Mann—Whitney *U*-test.

still given. There were no significant differences between these episodes and those with continued feedback for NFR (0.39 ± 0.18 versus 0.46 ± 0.16, difference; 0.07 (−0.01, 0.16), $P = 0.1$), compressions depth (38 ± 8 mm versus 38 ± 6 mm, difference; 0 (−3, 3) mm, $P = 0.9$), or compression rate (113 ± 10 min^{-1} versus 109 ± 13 min^{-1}, difference; −4 (−11, 2), $P = 0.2$), respectively. A tendency towards lower percentage of compressions with correct depth was found in the group that turned verbal feedback off (medians and inter quartile range); 36% (17, 57) versus 55% (32, 69), $P = 0.07$ (two-tailed Mann—Whitney *U*-test).

Outcome measures are presented by an intention to treat analysis (241 versus 117 patients) in Table 5 and include results for subgroups of patients with VF and non-VF (Asystole and PEA) as initial rhythm. Seven patients (2.9%) survived to hospital discharge in the baseline group and five (4.3%) in the feedback group. Neurological status for the survivors was good (Cerebral Performance Category (CPC) 1) for all except for one patient in the baseline group

Table 4 Change in quality between feedback versions 1 and 2

	Feedback 1 ($n = 69$)	Feedback 2 ($n = 39$)	Difference mean (95% CI)	*P*-value
No flow				
NFR	0.47 ± 0.17	0.40 ± 0.16	0.07 (0.01, 0.13)	0.03
NFR_{adj}	0.39 ± 0.17	0.33 ± 0.14	0.06 (−0.00, 0.12)	0.06
Compression				
Compressions (min^{-1})	60 ± 20	69 ± 21	−9 (−17, −1)	0.03
Compression rate (min^{-1})	110 ± 12	108 ± 13	2 (−3, 7)	0.5
Depth per episode (mm)	39 ± 5	36 ± 7	3 (0, 5)	0.04
Depth 38—51 mm (%)	57 (49, 60)	35 (27, 57)		0.01
Too deep (>51 mm) (%)	2 (1, 4)	0 (0, 1)		0.04
Too Shallow (<38 mm) (%)	35 (28, 45)	51 (32, 69)		0.03
Incomplete release (%)	1 (0, 1)	1 (0, 1)		0.6
Ventilations ($n = 62$ and 37, respectively)				
Ventilations (min^{-1})	11 ± 4	11 ± 4	−0 (−2, 1)	0.8

Values given as mean ± S.D. and differences as mean difference with 95% confidence interval (CI) and *P*-values for difference not equal to 0 from two-sided independent samples *t*-test, except for percentages of compressions with depth 38—51 mm, too deep, too shallow and with incomplete release which are given as median values with 95% CI and *P*-values of difference not equal to 0 with Mann—Whitney *U*-test.

Table 5 Outcome according to initial rhythm and intervention

	Baseline	Feedback	OR (95% CI)	*P*-values
All rhythms	241	117		
Admitted alive	42 (17)	27 (23)	1.4 (0.8, 2.4)	0.3
Discharged alive	7 (2.9)	5 (4.3)	1.5 (0.5, 4.8)	0.7
VF as initial rhythm	98	38		
Admitted alive	25 (26)	11 (29)	1.2 (0.5, 2.7)	0.8
Discharged alive	7 (7.1)	2 (5.3)	0.7 (0.1, 3.6)	1
Non-VF as initial rhythm	143	79		
Admitted alive	17 (12)	16 (20)	1.9 (0.9, 4.0)	0.1
Discharged alive	0 (0)	3 (3.8)	Not available	0.04

Number of patients with (%). Odds ratios (OR) with 95% confidence intervals (95% CI). A value above 1 indicates improved outcome in the feedback cohort. *P*-values obtained from two-sided χ^2-test or Fisher's exact test if expected values in any cell were less than 5.

with intact somatic functions but reliance on others for daily life (CPC 3).[23]

The uncorrected odds ratios for the different factors in the logistic regression model are presented in Table 6 for the 284 episodes with complete quality data. Witnessed arrest, average compression depth (Figure 1), and initial rhythm were all significantly related to hospital admission with spontaneous circulation. When factors were combined in one model ($N = 248$) only witnessed arrest (yes/no) (OR 4.2 (1.6, 11), $P = 0.004$) and average compression depth (mm^{-1} increase) (OR 1.05 (1.01, 1.09), $P = 0.020$) were significantly associated with hospital admission with spontaneous circulation. Sex was included in the model with a near significant doubling of odds for admission for women versus men (OR 1.9 (0.97, 3.9), $P = 0.063$). The results from logistic regression analysis of all episodes with these three variables were shown in Table 7. We found no significant relationships with other measures of quality when included into the model (compression rate between 90 and 120 min^{-1} ($P = 0.6$), NFR ($P = 0.6$), or ventilation rate between 6 and 16 min^{-1} ($P = 0.6$)). Forced entry of the other

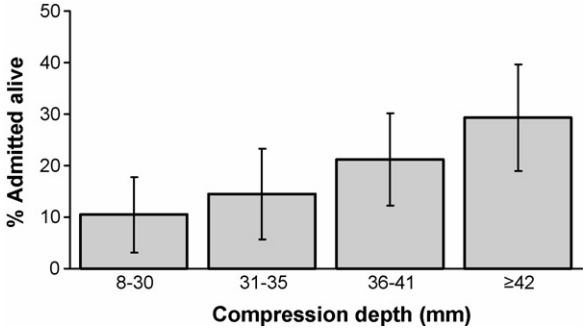

Figure 1 Percentage of patients admitted alive (error bars represent 95% CI) when grouped by increasing compression depth by quartiles.

Table 6 Unadjusted odds ratios with 95% confidence intervals (95% CI) for rate of admittance to hospital

Factor	Number of observations	OR for admittance (exp (β))	95% CI for exp (β)	*P*-value
Witnessed arrest (yes/no)	284 (77/207)	4.6	1.7, 12.0	0.002
Average compression depth (per 1 mm increase)	284	1.05	1.01, 1.09	0.009
Initial rhythm (VF/non-VF)	284 (111/173)	2.0	1.1, 3.6	0.022
Sex (female/male)	282 (76/206)	1.7	0.93, 3.3	0.082
Adrenaline given (yes/no)	272 (233/39)	0.6	0.27, 1.2	0.161
Feedback (yes/no)	284 (108/176)	1.5	0.81, 2.7	0.208
Response time (per 1 min increase)	284	1.04	0.96, 1.1	0.314
Average number of ventilations 6—16 min^{-1} (yes/no)	261 (216/45)	0.7	0.33, 1.5	0.363
Average compression rate 90—120 min^{-1} (yes/no)	284 (167/117)	1.1	0.58, 1.9	0.841
Age (years)	279	1.0	0.98, 1.02	0.956

OR > 1 indicates improved survival compared to the reference group (reference group were absence of feature or lower value of continuous variables).

Table 7 Adjusted odds ratios (adj. OR) with 95% confidence intervals (95% CI) for improved rate of admittance to hospital for the factors included in the regression model

Factor	Adj. OR for admittance (exp (β))	95% CI for exp (β)	*P*-value
Witnessed arrest (yes/no)	4.3	1.6, 11.3	0.003
Average compression depth (per 1 mm increase)	1.05	1.01, 1.09	0.011
Sex (female/male)	1.6	0.84, 3.1	0.153

OR > 1 indicates improved survival compared to the reference group, $N = 282$.

described factors to the model did not alter the overall coefficients and are not shown.

Discussion

In this prospective, clinical trial chest compression depth and rate improved during a period with automatic verbal and visual feedback on CPR quality compared to the time period immediately before feedback was introduced. A limitation is that the study was sequential, not randomised. This was done on purpose to avoid possible spill-over effects to the control group of feedback-induced increased awareness of quality problems. This spill-over was anticipated to be a problem even if cluster randomisation within the same ambulance services was applied. The study was planned prospectively with a formulated hypothesis that the intervention would affect the outcome measure and both positive and negative results would have been of interest. It conforms to the features for studies without internal controls described by Bailar et al.[24] An increased attention to CPR quality was probably present from the baseline period as the participants knew the purpose of the study from the start. The measured effect of feedback is thus not only a Hawthorne effect.[25]

Feedback failed to significantly reduce the fraction of time without chest compressions (no flow ratio). This was the most striking quality problem during the baseline period.[7] When designing the study, we had not anticipated this to be a large problem, and thus feedback was only a tonal signal after 15 and 30 s hands-off, while all other feedback was verbal with higher priority. New priorities in the feedback software significantly improved no flow ratio, apparently at the expense of compression depth. It could be speculated that this was due to an inability to respond to many types of feedback simultaneously. When feedback on hands-off periods increased, the rescuers were less able to concentrate on compression depth, although by nature these two types of error could not appear simultaneously. It is known that increased complexity decreases skill performance during CPR,[26] and from psychological literature there is evidence for reduced attention to new stimuli during attention demanding tasks supporting the popular notion that there is a finite capability for simultaneous tasks.[27]

The poor quality of CPR defined as chest compressions and ventilations can also be due to the long list of interventions included in the guidelines occurring at the expense of chest compressions, as recently suggested in an editorial by Sanders and Ewy.[28] It is illustrative that on 31 of 103 (30%) evaluation forms rescuers commented that they chose not to follow feedback at times when they were concentrating on other tasks such as intubation or placement of an iv. needle. This indicates that to reduce time without chest compressions further might depend on changes in the guidelines and more emphasis on avoiding hands-off intervals in training.[29] Chest compressions have been shown to improve survival,[2–4] while there have been no studies relating increased survival to hospital discharge with tracheal intubation or intravenous drugs.

The evaluation forms from the users of the feedback system helped us identify and improve feedback software and they assured us that the possible annoyance of another source of noise at the scene was outweighed by the perceived benefit from the feedback. However, the lack of systematic debriefing after each episode may introduce bias in a way that only the most positive and negative responses are written down.

In a logistic regression analysis compression depth was significantly associated with short-term survival. This was equally true whether the percentage of compressions with adequate depth or average compression depth in mm were used in the model (data not shown). Of the previously shown determinants of survival[21] only witnessed arrest ended up as significant in our model. Forced entry of the other factors did not change the model. The logistic regression method does not imply causality. Increased chest compression depth improved cardiac perfusion and cardiac output in animal experiments,[30,31] but this has not been possible to study directly in humans. We feel that a ran-

domised, prospective study with different chest compression depths in humans would be unethical as shallow compressions have been detrimental in animals. Clinical studies including measurements of CPR quality would improve our understanding of the effects of CPR quality on survival.

The automated audible feedback system was based on results from basic life support (BLS) manikin studies.[12,13] The advantages over observer feedback are that it gives accurate comments to variables that are hard to judge manually such as compression depth and inflation rate, and it never gets tired or distracted. In the manikin studies the percentage of correct compressions and inflations improved within 3 min with feedback from 46 ± 33 to 87 ± 9 and 18 ± 26 to 62 ± 25, respectively, even 12 months after initial training.[15] The improvements were not as large in this clinical study. This is perhaps not that surprising as the circumstances during out-of-hospital ALS are often difficult with cramped working space and disturbances, very different from the training laboratory. In addition, as mentioned above ALS guidelines include a number of interventions that require focus and physical handling. It is therefore encouraging that feedback did affect the quality of CPR with a parallel trend towards increased ROSC and survival rates.

Based on the present findings further investigations should be encouraged to find optimal feedback; whether it is visual, tonal, and/or voice prompts, and the ideal hierarchy and intensity of feedback. It is also possible that feedback priorities need to be adjusted for different professional cultures and situations. While rapid ventilation rates were observed both in- and out-of-hospital in US studies,[8,32] this was not a problem in this European study.

Future research should also consider measuring the effects of quality awareness. Debriefing, by letting the rescuers review their own performance data just after the arrest, can be one way to generate awareness. Tailored training, where the focus of training is determined by the quality data might also hold potential for continuous improvement.

An incidental finding in the baseline period was three probable unrecognised oesophageal intubations.[33] These were indicated by good thoracic impedance ventilation signals in the pre-intubation period which disappeared completely after the intubation attempt. No oesophageal intubations were indicated in the feedback group. Whether this was due to feedback on disappearance of ventilation signals cannot be determined as the study was not powered to detect any such differences.

Conclusion

Quality of CPR improved with automated feedback. Changing feedback priority caused a parallel change in quality. Among all cases, increased chest compression depth was associated with increased short-term survival in a logistic regression model.

Conflict of interest statements

Authors Kramer-Johansen, Fellows, Svensson, and Sørebø have no conflicts of interest to declare. Myklebust is a full time employee on a fixed salary at Laerdal Medical, Stavanger.

Wik is on an advisory board for Medtronic Medical. Steen is a member of the board of directors for Laerdal Medical and The Norwegian Air Ambulance.

Acknowledgements

We thank all the participating paramedics and nurses for their enthusiasm, criticism and invaluable effort. We especially want to thank the local instructors: Jan Ottem, and Lars Didrik Flingtorp (Division of Prehospital Medicine, Ullevål University Hospital, Oslo, Norway), Helena Borovszky, RN, and Lars Säfsten, RN (Stockholm Ambulance Service, Stockholm, Sweden) Andrew Nord and Allan Bromley (London Ambulance Service NHS Trust, London, UK). The helpfulness and technical skills of Ståle Freyer, MSc (Unversity of Stavanger, Stavanger, Norway), Geir Inge Tellnes, MSc, Linn Sømme, MSc, and Mette Stavland, MSc (all; Laerdal Medical, Research and Development, Stavanger, Norway) are most appreciated, as are the inspiring discussions and inputs from our US collaborators Lance Becker, MD, and Ben Abella, MD, MPhil (University of Chicago Hospitals, Chicago, IL, USA). Bjørn Auestad, PhD (University of Stavanger, Stavanger, Norway) helped with the logistic regression analysis.

Funding: Funding for this project was received from Norwegian Air Ambulance Foundation (JK-J full time PhD scholarship), Regional Health Authority East (unrestricted grants), and Anders Jahre Foundation for Sciences (unrestricted grants). The experimental defibrillators used in this project, and designated server/software for data storage and primary analysis were kindly provided by Laerdal Medical, Stavanger, Norway, and Phillips Medical Systems (Andover, MA, USA). Laerdal Medical also covered the expenses for extra training of involved paramedics, data collection and travel expenses for these purposes. The study sponsors had no further

role in the design of the study and interpretation of the data and had no part in the decision to submit the paper for publication. All authors' (except HM) work were independent of the study sponsors.

References

1. Larsen MP, Eisenberg MS, Cummins RO, Hallstrom AP. Predicting survival from out-of-hospital cardiac arrest: a graphic model. Ann Emerg Med 1993;22:1652—8.
2. Gallagher EJ, Lombardi G, Gennis P. Effectiveness of bystander cardiopulmonary resuscitation and survival following out-of-hospital cardiac arrest. JAMA 1995;274: 1922—5.
3. Van Hoeyweghen RJ, Bossaert LL, Mullie A, et al. Quality and efficiency of bystander CPR. Resuscitation 1993;26:47—52.
4. Wik L, Steen PA, Bircher NG. Quality of bystander cardiopulmonary resuscitation influences outcome after prehospital cardiac arrest. Resuscitation 1994;28:195—203.
5. Guidelines 2000 for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2000; 102(Suppl I): I-1—I-403.
6. International guidelines 2000 for CPR and ECC: a consensus on science. Resuscitation 2000; 46:1—447.
7. Wik L, Kramer-Johansen J, Myklebust H, et al. Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest. JAMA 2005;293:299—304.
8. Abella BS, Alvarado JP, Myklebust H, et al. Quality of cardiopulmonary resuscitation during in-hospital cardiac arrest. JAMA 2005;293:305—10.
9. Sunde K, Eftestol T, Askenberg C, Steen PA. Quality assessment of defribrillation and advanced life support using data from the medical control module of the defibrillator. Resuscitation 1999;41:237—47.
10. Valenzuela TD, Kern KB, Clark LL, et al. Interruptions of chest compressions during emergency medical systems resuscitation. Circulation 2005;112:1259—65.
11. van Alem AP, Sanou BT, Koster RW. Interruption of cardiopulmonary resuscitation with the use of the automated external defibrillator in out-of-hospital cardiac arrest. Ann Emerg Med 2003;42:449—57.
12. Handley AJ, Handley SA. Improving CPR performance using an audible feedback system suitable for incorporation into an automated external defibrillator. Resuscitation 2003;57:57—62.
13. Wik L, Thowsen J, Andreas Steen P. An automated voice advisory manikin system for training in basic life support without an instructor. A novel approach to CPR training. Resuscitation 2001;50:167—72.
14. Wik L, Myklebust H, Auestad BH, Steen PA. Retention of basic life support skills 6 months after training with an automated voice advisory manikin system without instructor involvement. Resuscitation 2002;52:273—9.
15. Wik L, Myklebust H, Auestad BH, Steen PA. Twelve-month retention of CPR skills with automatic correcting verbal feedback. Resuscitation 2005;66:27—30.
16. World Medical Association Declaration of Helsinki, Ethical principles for medical research involving human subjects 2002. http://www.wma.net/e/policy/b3.htm 17.C. [accessed at March 31, 2006].
17. Aase SO, Myklebust H. Compression depth estimation for CPR quality assessment using DSP on accelerometer signals. Biomed Eng, IEEE Trans 2002;49:263—8.
18. Pellis T, Bisera J, Tang WC, Weil MH. Expanding automatic external defibrillators to include automated detection of cardiac, respiratory, and cardiorespiratory arrest. Crit Care Med 2002;30:S176—8.
19. Wik L, Hansen TB, Fylling F, et al. Delaying defibrillation to give basic cardiopulmonary resuscitation to patients with out-of-hospital ventricular fibrillation: a randomized trial. JAMA 2003;289:1389—95.
20. Altman DG. Table B11. In: Altman DG, editor. Practical statistics for medical research. London: Chapman & Hall; 1991. p. 535—7.
21. Holmberg M, Holmberg S, Herlitz J. Low chance of survival among patients requiring adrenaline (epinephrine) or intubation after out-of-hospital cardiac arrest in Sweden. Resuscitation 2002;54:37—45.
22. Langhelle A, Tyvold SS, Lexow K, Hapnes SA, Sunde K, Steen PA. In-hospital factors associated with improved outcome after out-of-hospital cardiac arrest. A comparison between four regions in Norway. Resuscitation 2003;56: 247—63.
23. Chamberlain D, Cummins RO, Abramson N, et al. Recommended guidelines for uniform reporting of data from out-of-hospital cardiac arrest: the 'Utstein style': prepared by a Task Force of Representatives from the European Resuscitation Council, American Heart Association, Heart and Stroke Foundation of Canada, Australian Resuscitation Council. Resuscitation 1991;22:1—26.
24. Bailar III JC, Louis TA, Lavori PW, et al. Studies without internal controls. In: Bailar III JC, Mosteller F, editors. Medical uses of statistics. 2nd ed. Boston, MA: NEJM Books; 1992. p. 105—23.
25. Campbell JP, Maxey VA, Watson WA. Hawthorne effect—implications for prehospital research. Ann Emerg Med 1995; 26:590—4.
26. Rittenberger JC, Guimond G, Platt TE, Hostler D. Quality of BLS decreases with increasing resuscitation complexity. Resuscitation 2005;68:365—9.
27. Most SB, Scholl BJ, Clifford ER, Simons DJ. What you see is what you set: sustained inattentional blindness and the capture of awareness. Psychol Rev 2005;112: 217—42.
28. Sanders AB, Ewy GA. Cardiopulmonary resuscitation in the real world: when will the guidelines get the message? JAMA 2005;293:363—5.
29. European Resuscitation Council Guidelines for Resuscitation 2005. Resuscitation 2005: 67(S1); S1—S190.
30. Babbs CF, Voorhees WD, Fitzgerald KR, Holmes HR, Geddes LA. Relationship of blood-pressure and flow during cpr to chest compression amplitude—evidence for an effective compression threshold. Ann Emerg Med 1983;12: 527—32.
31. Bellamy RF, Deguzman LR, Pedersen DC. Coronary blood-flow during cardiopulmonary resuscitation in swine. Circulation 1984;69:174—80.
32. Aufderheide TP, Sigurdsson G, Pirrallo RG, et al. Hyperventilation-induced hypotension during cardiopulmonary resuscitation. Circulation 2004;109:1960—5.
33. Kramer-Johansen J, Wik L, Steen PA. Advanced cardiac life support before and after tracheal intubation-direct measurements of quality. Resuscitation 2005;68:61—9.