Efficacy of Computerized Decision Support for Mechanical Ventilation: Results of a Prospective Multi-Center Randomized Trial

Thomas D. East, Ph.D., Laura K. Heermann, RN, MS, Richard L. Bradshaw MS, Alejandra Lugo, BS, R. Matthew Sailors, ME, Lewis Ershler MS, C. Jane Wallace, RN,PhD, Alan H. Morris, MD, Bruce McKinley PhD[1], Alicia Marquez RN[1], Allen Tonnesen MD[1], Lee Parmley MD[1],William Shoemaker MD[2], Peter Meade MD[2], Phil Thaut RRT[3], Tracy Hill MD[3], Michael Young MD[4], John Baughman MD[4], Mark Olterman MD[4], Valerie Gooder RN[4], Bill Quinn RRT[5],William Summer MD[5], Vincent Valentine MD[5], Jeff Carlson RRT[6], Bruce Bonnell MD[6], Ben deBoisblanc MD[7], Zeke McClarity RT[7], Jean Cachere RT[7], Kevin Kovitz MD[8], Eleanor Gallagher RN[8], Michael Pinsky MD[9], Derek Angus MD[9], Mark Cohen RRT[9], Len Hudson MD[10], Ken Steinberg MD [10]
Department of Medical Informatics, Cottonwood Hospital, 5770 S. 250 East, Salt Lake City, Utah 84107 and University of Utah, Departments of Medical and Nursing Informatics
1-Hermann Hospital, Houston TX, 2-King-Drew Medical Center, LA CA, 3-Utah Valley Regional Medical Center, Provo UT, 4-McKay Dee Hospital, Ogden UT, 5-Oschner Clinic, New Orleans LA, 6-Butterworth Hospital, Grand Rapids, MI, 7-University/Charity Hospital, New Orleans LA, 8-Tulane, New Orleans LA, 9-Univ of Pittsburgh Med Center, Pittsburgh PA, 10-Harborview Med Ctr, Seattle WA

Tutorial Abstract

200 adult respiratory distress syndrome patients were included in a prospective multicenter randomized trial to determine the efficacy of computerized decision support. The study was done in 10 medical centers across the United States. There was no significant difference in survival between the two treatment groups ($X^2=0.49$ $p=0.49$) or in ICU length of stay between the two treatment groups when controlling for survival ($F(1df)=0.88$, $p=0.37$.) There was a significant reduction in morbidity as measured by multi-organ dysfunction score in the protocol group ($F(1df)=4.1$, $p=0.04$)as well as significantly lower incidence and severity of overdistension lung injury ($F(1df)=45.2$, $p<0.001$). We rejected the null hypothesis. Efficacy was best for the protocol group. Protocols were used for 32,055 hours (15 staff person years, 3.7 patient years or 1335 patient days). Protocols were active 96% of the time. 38,546 instructions were generated. 94% were followed. This study indicates that care using a computerized decision support system for ventilator management can be effectively transferred to many different clinical settings and significantly improve patient morbidity.

Background

The care of critically ill patients is increasingly complex and clinicians frequently suffer from information overload. It is difficult, if not impossible to assess all this information and generate a systematic and reasonable therapy plan. Computerized decision support systems can assist the clinician with many of the tasks such as the iterative management of mechanical ventilation. This decision support not only standardizes care but also may improve the quality of care by reducing mistakes. This standardization of care also makes it possible to thoroughly characterize the current treatment process in order to compare it to a proposed new therapy as part of an ongoing continuous quality improvement (CQI) program.

Several computerized expert systems to assist in the management of ventilation and oxygenation have been described [1-4] . The Miller approach takes a manually entered short medical history, arterial blood gas values, ventilator settings and the new therapy being considered and provides a critique [2]. Such a procedure requires laborious hand entry of data and has not been subjected to a rigorous clinical evaluation. The Menn system is an interactive consultation based program that requires hand entry of patient primary diagnosis, respiratory care charting and most recent arterial blood gas data [1]. Even though the Menn system did not include recommendations on pressure levels (PEEP, CPAP, etc), it was considered to be clinically useful 68% (n=198) of the time, most often (81%) only to organize data. In 14% of the cases, the user stated that the program changed patient care. The VentPlan system is the most comprehensive system [4]. It interpreted physiological data from an on-line post-surgical ICU. Five patient-days worth of data were interpreted retrospectively but the system was not evaluated prospectively or put to routine bedside use. Tong [3] demonstrated, in a small trial, the benefit of a system designed to assist in the weaning of patients from the ventilator. None of these systems covered the whole process of mechanical ventilation from intubation to extubation. No previous systems were used for routine clinical care or evaluated to determine their efficacy. Our overall goal was the creation of a decision support tool for mechanical ventilation that could be widely applied in different clinical settings and to evaluate its efficacy.

A computerized decision support system for the management of mechanical ventilation (respiratory evaluation, oxygenation, ventilation, weaning and extubation) in patients with adult respiratory distress

1091—8280/99/$5.00

syndrome (ARDS) has already been developed and clinically validated at the LDS Hospital [5-9]. The computerized decision support system was used for over 35,000 hours in 111 ARDS patients and controlled decision making 95% of the 24-hour day. The survival rate was 67%, higher than the expected 31-33% from historical data [10, 11], $p < 0.05$. These results demonstrated that computerized decision support for critical care was feasible. This contrasted with the common medical wisdom that protocol control of therapy of such complicated ICU patients is impossible [12].

There has never been a prospective randomized trial of the efficacy of computerized decision support for critical care in general much less for mechanical ventilation specifically. The purpose of this study was to answer two questions: 1) Can a computerized decision support system be exported to other centers and used by clinicians uninvolved with its development? and 2) Does the system have an impact on patient outcome?

Methods

A prospective randomized clinical trial was done at ten different clinical sites *(see author list for site specifics)* 200 ARDS patients were randomly allocated to either the protocol or non-protocol control treatment groups over the four-year accession period. The trial design is summarized in figure 1.

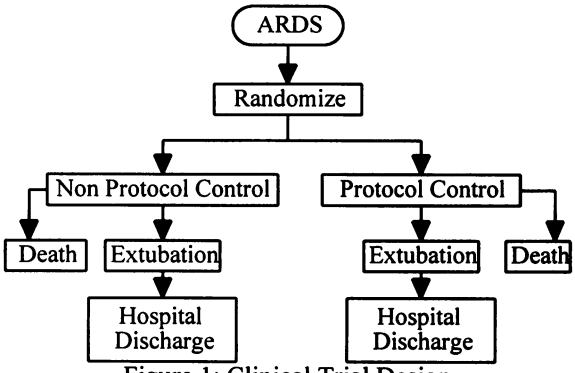

Figure 1: Clinical Trial Design

The clinical endpoint of the trial was either survival at hospital discharge or death.

The ARDS entry criteria were defined using established standards [13, 14 15]:

1. Arterial/alveolar partial pressure of oxygen ratio ≤ 0.3 (or arterial partial pressure of oxygen divided by the fraction of inspired oxygen ≤200)
2. Total static thoracic compliance ≤ 50 ml/cm H_2O
3. No evidence of heart failure or fluid overload (pulmonary artery occlusion pressure ≤ 18)
4. Acute onset of illness accompanied by an ARDS risk factor
5. Radiographic evidence of bilateral diffuse infiltrates

Patients who met any of the following criteria were excluded from the trial:

1. ARDS for > 21 days duration
2. Severe chronic systemic disease or another clinical condition which in itself greatly limits survival. These include but are not limited to:
   a) Irreversible central nervous system
   b) Severe chronic obstructive pulmonary disease
   c) Total body surface burns exceeding 40%
   d) Rapidly fatal malignancy
   e) Chronic left ventricular failure
   f) Chronic renal failure
   g) Chronic liver failure

Stratified randomization with blocking was used. Within each clinical trial site, the patients were initially stratified by age and existence of barotrauma and then randomly assigned to either the protocol or non-protocol treatment groups. Stratification was used to avoid nonuniform distribution of patients with characteristics likely to influence their outcome (i.e. all of the older patients in the protocol group). The treatment groups were balanced in randomly chosen blocks of 2 or 4 to avoid temporal nonuniformity while assuring a lack of prior knowledge of the resulting treatment group. Efficacy was defined as the hierarchy of:

1) **Survival** *(to hospital discharge)*
2) **Costs** *(reflected by length of ICU stay)*
3) **Morbidity** *(a daily multi-organ dysfunction and sepsis score(MODS) was done to assess major organ system function)* [15-17] The MODS score (range=0-30) is a sum of graded scores (0-3) for each of the 8 major organ systems and a six level score (0-6) for sepsis.
4) **Iatrogenic Injury** *(Incidence and severity of overdistension and pressure damage to the lung known as barotrauma that can be a side effect of mechanical ventilation)* There are many chest x-ray scoring systems for cystic fibrosis [18-20]; however, there were none designed to measure barotrauma. Our barotrauma score was the sum of two parts. First, a score from the daily chest radiograph *(if multiple findings-we used highest score)*:

Score	Description
0	No Evidence of Barotrauma
1	Interstitial Air
2	Pneumomediastinum or SubQ emphysema
3	Unilateral Pneumothorax
4	Bilateral Pneumo or massive SubQ emphysema

The second portion of the score was the number of chest tubes (thoracostomy tubes). The barotrauma score not only reflected the incidence of new barotrauma (radiographic score) but also the therapeutic interventions necessary to treat already existing barotrauma (number of chest tubes).

The null hypothesis (H0: There is no difference in efficacy between protocol and non-protocol control) was tested in the four hierarchical levels of efficacy:

Level	Variable	Statistical Test
1	Survival	X^2
2	ICU Length of stay for survivors and non-survivors	3 Way ANOVA
3	Max MODS Score for survivors and non-survivors	3-way ANOVA
4	Max Barotrauma Score for patients with and without barotrauma at randomization	2-way ANOVA

We built into the computerized decision support tools automated monitoring agents that stored in the database a very detailed record of how the system

was used. This allowed the analysis of user acceptance and generalizability. In particular we examined;

1)Percent of total time in the trial during which protocols controlled patient care.
2)Number of protocol instructions that were not followed and why.
3)Variation in performance at different clinical sites.

Results

The first level of the definition of efficacy was survival. The following table summarizes survival for each of the cells in our randomization matrix.

Age	Rx Group	Barotrauma	died	lived	
≥65	Control	No Baro	7	6	13
		Baro	1	1	2
		Baro	1	1	2
	Protocol	No Baro	4	10	14
	Protocol	No Baro	4	10	14
		Baro	4	2	6
		Baro	4	2	6
<65	Control	No Baro	15	35	50
<65	Control	No Baro	15	35	50
		Baro	9	23	32
		Baro	9	23	32
	Protocol	No Baro	16	36	52
	Protocol	No Baro	16	36	52
		Baro	12	19	31
		Baro	12	19	31
			68	132	200
			68	132	200

There was no significant difference in survival between the two treatment groups (Mantel-Haentzel X^2=0.49 p=0.49).

The second level of our definition of efficacy was ICU length of stay as a surrogate for costs.

ICU Length of Stay (days)

Rx Group	survival	Mean	SEM	N
Control	died	25.8	3.5	32
	survived	25.2	2.4	65
Protocol	died	25.5	3.6	36
	survived	27.6	2.1	67

There was no significant difference in ICU length of stay between the two treatment groups when controlling for survival (F(1df)=0.88, p=0.37)

The third level definition of efficacy was morbidity as measured by the maximum daily MODS score.

MODS Score

Group	Survival	Age	Mean	SEM	N
Control	died	<65	15.13	1.37	8
		≥65	16.59	0.70	22
	survived	<65	12.00	0.93	6
		≥65	11.58	0.44	57
Protocol	died	<65	14.88	1.63	8
		≥65	15.12	0.68	26
	survived	<65	11.42	1.30	12
		≥65	11.94	0.48	53

There was a significantly reduced MODS score in the protocol group (F(1df)=4.1, p=0.04)

The efficacy fourth level was iatrogenic lung injury as reflected by a maximum daily barotrauma score.

Barotrauma Score

Group	Baro	Mean	SEM	N
Control	non baro	0.73	0.20	59
	baro	3.56	0.36	32
Protocol	non baro	0.97	0.24	61
	baro	2.69	0.31	35

There was a significantly lower barotrauma score in the protocol group (F(1df)=45.2, p<0.001)

Evaluating the overall null hypothesis; level 1 and 2 were not significantly different, levels 3 and 4 were significantly different. We rejected the null hypothesis that there is no significant difference in efficacy between the protocol and the control group. Efficacy was best for the protocol group.

The results of this study represent over 32,000 hours of experience using the computerized decision support tool. None of this time is in the institution where the tool was constructed. This is the equivalent of 15 staff person years of use and close to four patient years of care. Figure 2 illustrates that despite the variety of clinical settings and clinicians the protocols were used 96% of the time they were intended to be used. This is the time from randomization to extubation excluding times where the family may have elected to withdraw support. Our automated monitoring agents allowed us to examine every single instruction of more than 38,000 generated.

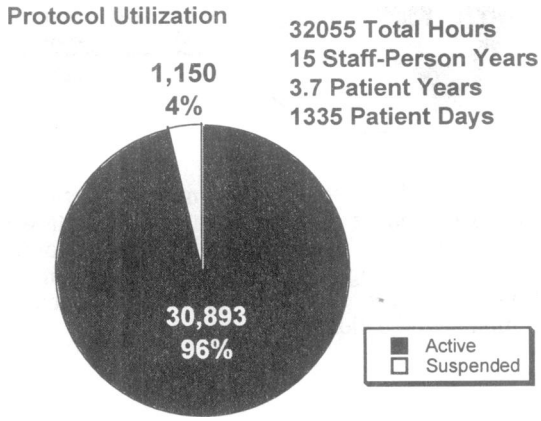

Figure 2: Protocol Utilization (Hours)

Only 6% of instructions were not followed (figure 3).

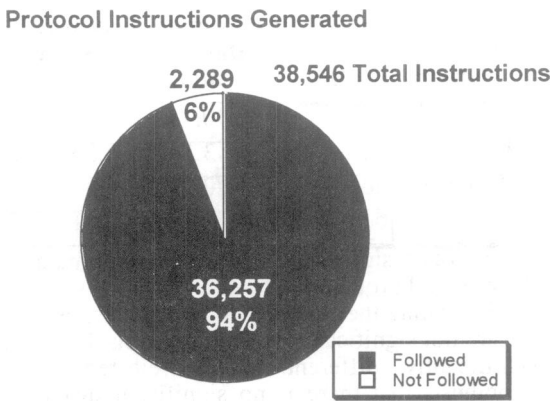

Figure 3: Fraction of instructions followed

Clinicians rarely (0.3%) found substantial clinical reason to not follow the instructions (figure 4). The most frequent problem (37% of declined instructions) was incorrect or missing data charted in the patients record.

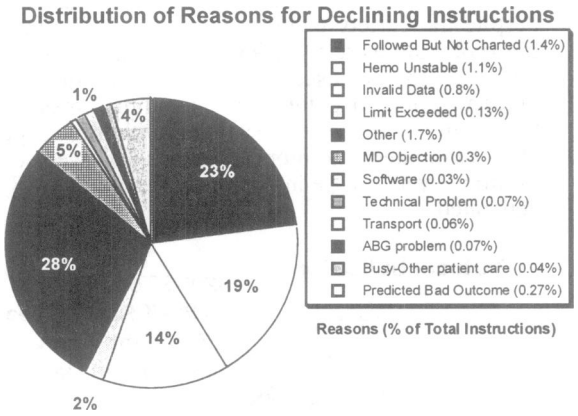

Figure 4: Reasons for declining instructions

The performance was good at all sites; however, there was significant variability in performance between sites (figure 5). The best performance was from the sites that had prior experience with computerized patient records.

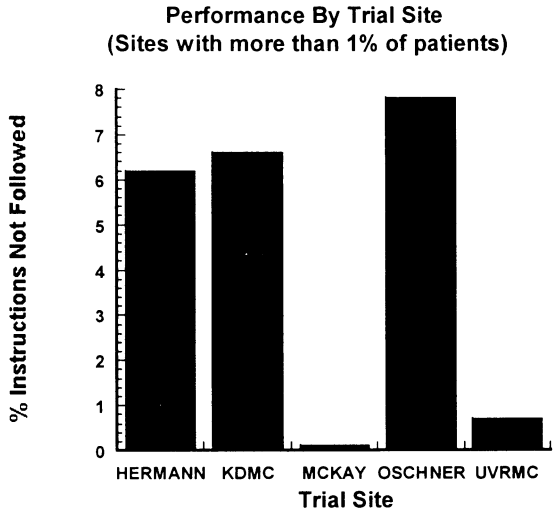

Figure 5: Performance at different clinical trial sites

Discussion

This is one of the first large multi-center prospective randomized trials of the efficacy of computerized decision support. The study demonstrates that although survival and length of stay do not appear to be different there is strong evidence that the number of organs failing and the amount of iatrogenic lung damage were significantly reduced. The success experienced in this study proves that a decision support tool can be created at one site and effectively distributed to other clinical sites uninvolved in its original creation. These results indicate that computers and computerized decision support can be an effective tool for dissemination of standards of care as well as study protocols.

References

1. Menn SJ, Barnett GO, Schnechel D, Owens WD, Pontoppidan H. A computer program to assist in the care of acute respiratory failure. JAMA 1973;223:308-312.
2. Miller PL. Goal oriented critiquing by computer for ventilatory management. Comp Biom Res 1985;18:422-438.
3. Tong D. Weaning patients from mechanical ventilation: a knowledge based approach. In: Miller RA, editor. Proceedings Symposium on Computer Applications in Medical Care (SCAMC); 1990; Washington, DC: IEEE Computer Soc Press; 1990. p. 79-85.
4. Rutledge G, Thomsen G, Farr B, Tovar M, Sheiner L, Fagan L. VentPlan: a ventilator-management advisor. In: Clayton PD, editor. Symposium on Computer Applications in Medice Care; 1991; Washington, DC: IEEE Computer Society Press; 1991. p. 869-871.
5. East TD, Morris AH, Wallace CJ, et al. A strategy for development of computerized critical care decision support systems. Int J Clin Monit Comput 1991;8(4):263-9.
6. East T. Computerized Management of Mechanical

Ventilation. In: Grenvik A, Ayres SM, editors. Textbook of Critical Care. 3 ed. New York: W.B. Saunders; 1994.
7. Henderson S, Crapo RO, Wallace CJ, East TD, Morris AH, Gardner RM. Performance of computerized protocols for the management of arterial oxygenation in an intensive care unit. Int J Clin Monit Comput 1991;8(4):271-80.
8. Morris A, Wallace C, Menlove R, et al. A computerized protocol-controlled randomized clinical trial of new therapy including PCIRV and extracorporeal CO2 removal for ARDS. Am J Respir Crit Care Med 1994;149:295-305.
9. Thomsen G, Pope D, East T, et al. Clinical performance of a rule-based decision support system for mechanical ventilation of ARDS patients. In: Safran C, editor. Proceedings 17th Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill; 1993. p. 339-343.
10. Artigas A, Carlet J, Le Gall JR, Chastang CL, Blanch L, Fernandez R. Clinical presentation, prognostic factors, and outcome of ARDS in the European Collaborative Study (1985-1987). In: Zapol WM, Lemaire F, editors. Adult Respiratory Distress Syndrome. New York: Marcel Dekker, Inc; 1991. p. 37-63.
11. Zapol WM, Frikker MJ, Pontoppidan H, Wilson RS, Lynch KE. The adult respiratory distress syndrome at Massachusetts General Hospital, Etiology progession and survival rates,1978-1988. In: Zapol WM, Lemair F, editors. Adult Respiratory Distress Syndrome. New York: Marcel Dekker Inc; 1991. p. 367-380.
12. Petty TL. ARDS, refinement of a concept and redefinition. Am Rev Respir Dis 1988;138:724.
13. Montgomery AB, Stager MA, Carrico CJ, Hudson LD. Causes of mortality in patients with the adult respiratory distress syndrome. Am Rev Respir Dis 1985;132(3):485-9.
14. Bell RC, Coalson JJ, Smith JD, Johanson WG, Jr. Multiple organ system failure and infection in adult respiratory distress syndrome. Ann Intern Med 1983;99(3):293-8.
15. Suchyta MR, Clemmer TP, Orme JF, Morris AH, Elliott CG. Increased survival of ARDS patients with severe hypoxemia (ECMO criteria). Chest 1991; 99:951-955.
16. Jacobs S, Zuleika M, Mphansa T. The Multiple Organ Dysfunction Score as a descriptor of patient outcome in septic shock compared with two other scoring systems. Crit Care Med 1999;27(4):741-4.
17. Barie PS, Hydo LJ, Fischer E. A prospective comparison of two multiple organ dysfunction/failure scoring systems for prediction of mortality in critical surgical illness. J Trauma 1994;37(4):660-6.
18. Conway SP, Pond MN, Bowler I, et al. The chest radiograph in cystic fibrosis: a new scoring system compared with the Chrispin-Norman and Brasfield scores [see comments] [published erratum appears in Thorax 1994 Nov;49(11):1184]. Thorax 1994;49(9):860-2.
19. Wong EB, Regnis J, Shnier RC, Bye PT, Stewart ME. The relationship between tests of lung function and three chest radiological scoring systems in patients with cystic fibrosis. Australas Radiol 1993;37(3):265-9.
20. Weatherly MR, Palmer CG, Peters ME, et al. Wisconsin cystic fibrosis chest radiograph scoring system. Pediatrics 1993;91(2):488-95.

Acknowledgements

This work was supported by AHCPR grant #HS06594, Siemens, Emtek, ACT/PC, the Respiratory Distress Syndrome Foundation, the Deseret Foundation (LDS Hospital) and Intermountain Health Care (IHC).

This has been a long six-year adventure that would not have been possible without the help of the many people including:

LDS and Cottonwood Hospital: Shauna Bruun Grant and Contracts Manager, Gloria Poarch Office Manager

Hermann Hospital: Mark Nguyen

Harborview: Donna Davis RN and the CIS staff

University of Pittsburgh Medical Center: Cindy Slivik RN, Daniel Martich MD,Judy Vaglia RN, Edgar Delgado RT, Bob Bickerstaff, Mary Beth Coleman RN, Maggie Soncini

Butterworth Hospital: Carmine Alexander, Bruce Robey, William VandenBerg,Marty Walstra

University/Charity Hospital: Chris Glynn RN (CC),Amy Bourgeois RN ,Carol Ross RN

McKay Dee Hospital: Brenda Grover RN, Karen McBride RN

Utah Valley Regional Med Ctr: Jim Hamilton, Larry Cooper, Karl Ludwig ,John Manwaring, Doug Ross MD, Bill Alward MD, Michael Pearce MD, Clark Bishop MD, Flora Gillman

King Drew Medical Center: Janyth Bolden RRT, Mike Bishop DO, Chuck McCarty RRT

We would like to acknowledge all of the fantastic critical care nurses and therapists who did the vast majority of the work... it would not have been possible without you!