Computer-Aided Diagnosis for Endotracheal Intubation Confirmation using Video-image Classification

Dror Lederman
Holon Institute of Technology, Holon, Israel
drorl@hit.ac.il

Keywords: Computer-Aided Diagnosis, Intubation Confirmation, Neural Networks.

Abstract: In this paper, a Computer-Aided Diagnosis (CAD) system for endotracheal tube position confirmation, and detection of errors in intubation positioning is presented. Endotracheal intubation is a complex procedure which requires high skills and the use of secondary confirmation devices to ensure correct positioning of the tube. Our novel confirmation approach is based on video images classification and specifically on identification of specific anatomical landmarks, including esophagus, upper trachea and main bifurcation of the trachea into the two primary bronchi ("carina"), as indicators of correct or incorrect tube insertion and positioning. Classification of the images is performed using a neural network classifier. The performance of the proposed approach was evaluated using a dataset of cow-intubation videos and a dataset of human-intubation videos. Each one of the video images was manually (visually) classified by a medical expert into one of three categories: upper tracheal intubation, correct (carina) intubation and esophageal intubation. The image classification algorithm was applied off-line using a leave-one-case-out method. The results show that the system correctly classified 1567 out of 1600 (97.9%) of the cow intubations images, and 349 out of the 358 human intubations images (97.5%).

1 INTRODUCTION

Intubation is a common medical procedure in hospitals as well as in emergency medical units. During intubation, a flexible tube is used to secure passage of air to and from the lungs. The procedure is performed by manually opening the mouth, lifting the tongue using a device called laryngoscope in order to reveal the vocal cords, and inserting an endotracheal tube (ETT) through the vocal cords. The ETT should be positioned between 2 and 5 cm above the bifurcation of the trachea into the two primary bronchi ("carina").

The anatomy of the patient does not always allow easy insertion of the ETT and consequently it might be incorrectly positioned, usually either in the esophagus or in the right main bronchus. Both of these conditions can produce catastrophic results, as the patient might be deprived of oxygen. Unintentional esophageal intubation has been associated with high mortality rate (Silvestri et al., 2005; Timmermann et al., 2007). In cases of right lung intubation (also termed one-lung intubation (OLI)), only one lung is ventilated. Prolonged one lung ventilation might cause serious pulmonary complications such as collapse of the contralateral lung and hyperinflation of the ventilated lung, which might eventually result in hypoxia and pneumothorax, respectively, and has been associated with a significant increase in morbidity (Owen et al., 1987; Zwillich et al., 1974) and Pneumonia (Wang et al., 2009). Both esophageal and OLI may occur after the ETT was positioned correctly ("dislodgement") from many reasons, for example, due to neck flexion during general anesthesia (Vergese et al., 2004; Yap et al., 1994).

Confirmation of correct tube positioning is a challenging task. It requires high skills and the use of secondary objective devices.

Numerous studies, which investigated endotracheal misplacement rates in hospital and pre-hospital settings, reported rates between 0% and 25%, depending among others, on study design (Jacobs et al., 1983; Jemmet et al., 2003; Jones et al., 2004; Katz et al., 2001; Pointer, 1988; Silvestri et al., 2005; Steward et al., 1984; Timmermann et al., 2007; Wang et al., 2009).

In this paper, we present a computer-aided diagnosis (CAD) system for endotracheal intubation confirmation. The system is based on identification of

Lederman, D.
Computer-Aided Diagnosis for Endotracheal Intubation Confirmation using Video-image Classification.
DOI: 10.5220/0006200505340540
In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2017)*, pages 534-540
ISBN: 978-989-758-222-6

specific anatomical landmarks as indicators of correct or incorrect tube positioning.

Based on our previous preliminary work (e.g. (Lederman, 2011)), we further developed and tested our novel approach for automatic endotracheal intubation confirmation. The approach is based on direct visual cues, i.e., identification of specific anatomical landmarks as indicators of correct or incorrect tube positioning. In this study, the system is further developed and evaluated using animal and human tissue model.

The paper is arranged as follows. Section 2 reviews the relevant work in this field. Section 3 presents the proposed confirmation system. The experimental results are presented in Section 3. The results appear in Section 4, followed by discussion. The conclusions appear in Section 5.

2 RELATED WORK

There are various methods and techniques for endotracheal intubation confirmation. The most common technique is auscultation to lung sounds using a stethoscope. This technique requires high attention, and its reliability has been questioned in many studies (Brunel et al., 1989; Howells, 1985; Klepper et al., 1993; Linko et al., 1983; Peterson et al., 1973; Wang et al., 2006; Wodicka et al., 1994). The use of exhaled carbon dioxide detection (CO2) measurements (termed end-tidal CO2 (ETCO2)), has become the gold standard-de-facto for confirming correct tube positioning. However, the method has been found to be unreliable in many emergencies (Bhende et al., 1995; Gravenstein et al., 2004; Li, 2001; Nolan et al., 2005; Webb et al., 1993). In addition, the method can not be used to detect OLI incidents as in such cases the capnogram is generally typical in shape and shows normal ETCO2 values (Gravenstein et al., 2004; Webb et al., 1993). Other techniques have been proposed (e.g., (Lederman, 2006; O'connor et al., 2005; Tejman-Yarden et al., 2006; Tejman-Yarden et al., 2007; Weizman et al., 2008)), but none of them has been proven effective. Therefore, attempts to find the ultimate technique for correct tube position confirmation have been continued.

Our proposed approach is based on direct visual cues, i.e., identification of specific anatomical landmarks as indicators of correct or incorrect tube positioning. In the following, we describe the method and report its performance, evaluated using intubation videos acquired on animals and human beings.

3 MATERIALS AND METHODS

The correct position of an ETT tip is 2-5 cm above the carina. The image of the carina is therefore used as the definitive anatomical landmark for confirming correct endotracheal intubation. Hence, identifying the carina in the acquired video images, and discriminating between the carina and other anatomical structures, is the main idea of the proposed method. The method combines an artificial neural network scheme which is employed in a textural-based feature space. A general block diagram of the proposed system appears in Figure 2.

3.1 The Video-stylet

Intubation is usually performed using an intubating stylet, used to control and guide the ETT. We designed and assembled a designated video-stylet. The tip of the stylet comprises a miniature complementary metal oxide silicon (CMOS) sensor. The inner part of the stylet contains wires to transfer the image and a narrow lumen to spray water or air in order to clear blood and secretions away from the camera sensor (Figure 1).

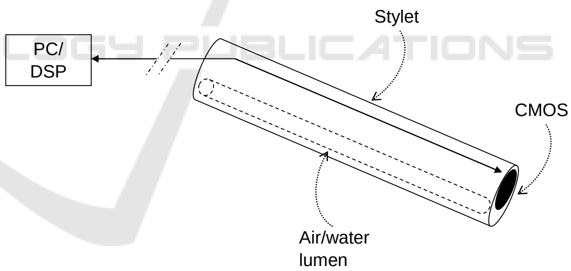

Figure 1: A schematic drawing of the video-stylet which includes the stylet and complementary metal oxide silicon (CMOS) sensor connected to a digital signal processor (DSP).

The image sensor is connected to a processor with an integrated image acquisition component. During intubation, this rigid stylet is inserted into a standard ETT with its camera at the tip. Video signals are continuously acquired and processed by the confirmation algorithm implemented on the processor.

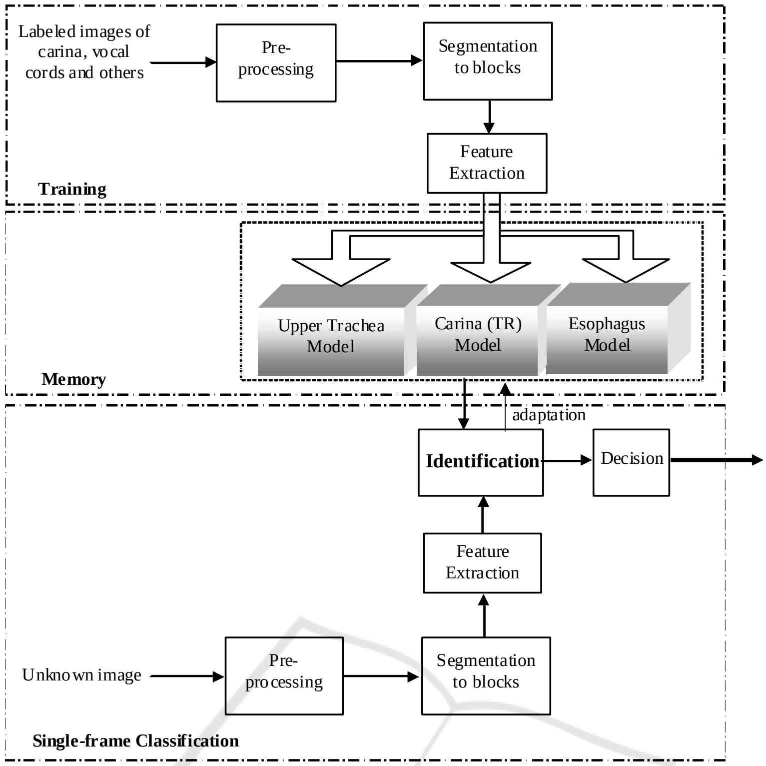

Figure 2: A general scheme of the proposed confirmation system. The system consists of three classes, one representing the upper-trachea, one representing the carina and one representing the esophagus.

3.2 Pre-processing and Features Extraction

The confirmation algorithm is based on classification of specific anatomical landmarks, including the carina, tracheal rings (upper trachea) and esophagus. We use textural features (Haralick et al., 1973) that contain important information about the structural arrangement of surfaces and their relationship to the surrounding environment. In particular, features based on grey level co-occurrence matrices (GLCM) are utilized. These features are based on the assumption that texture information on an image is contained in the overall or "average" spatial relationship, which the grey tones in the image have to one another. More specifically, it is assumed that this texture information is adequately specified by a set of grey tone spatial dependence matrices which are computed for various angular relationships and distances between neighboring resolution cell pairs on the image. One of the advantages of these features is that they are robust to imaging angles and scaling. This property is of great importance to the task in hand, as during intubation the tube may be inserted in different angles and directions, depending on the technique employed by the person performing the procedure. It was therefore hypothesized that textural features will allow reliable classification of the images, independently of the angle at which the tube was inserted.

A brief description of the textural features is now given. Let $f : Lx * Ly \rightarrow I$ be an image with dimensions Lx and Ly, and grey levels $g = 0,1,\ldots,G-1$. Let d be the distance (offset) between two pixel positions (x_1, y_1) and (x_2, y_2). Angles quantized to 45° intervals are considered, such that the neighbors of any pixel can lie on four possible directions: $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135°. A resolution cell is considered to have eight nearest-neighbor resolution cells. The co-occurrence matrix is constructed by observing pairs of image cells at distance d from each other and incrementing the matrix position corresponding to the grey level of both cells. The un-normalized frequencies for direction of 45°, for instance, are defined by:

$$
\begin{aligned}
P(i,j,d,45^\circ) = \#\{&(k,l),(m,n) \\
&\in (L_y, L_x) * (L_y, L_x) | k-m| \\
&= d, |l-n| \\
&= -d \ or \ (k-m = -d, l-n \\
&= d), I(k,l) = i, I(m,n) = j\},
\end{aligned}
\tag{1}
$$

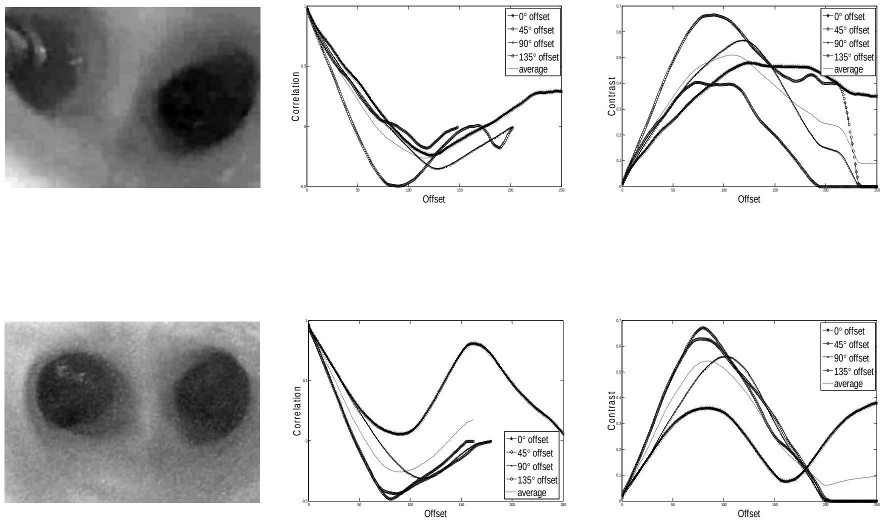

Figure 3: Two examples of carina images (left column) and the calculated textural features: correlation (middle column) and contrast (right column).

where # denotes the number of elements in the set. Measures of the other directions, as well as the normalized measures, can be easily obtained (Haralick et al., 1973).

To construct the feature set utilized in the proposed system, various textural features were extracted from the GLCM. Let $p(i,j)$ denote the $(i,j)^{\text{th}}$ entry in a normalized grey-tone spatial dependence matrix, such that $p(i,j)=P(i,j)/R$, where R is a normalization constant, which was set in this work to the sum of all values of $P(i,j)$, i.e., $R=\sum_{i=1}^{G}\sum_{j=1}^{G}P(i,j)$, and $p_x(i)$ and $p_y(i)$ denote the i^{th} entry in the marginal-probability matrix, obtained by summing the rows and columns of $p(i,j)$, respectively, i.e. $p_x(i)=\sum_{j=1}^{G}P(i,j)$, $p_y(j)=\sum_{i=1}^{G}P(i,j)$. Then, the following features are used to construct the feature set:

Contrast: $f_1=\sum_{n=0}^{G-1}n^2\left(\sum_{i=1}^{G}\sum_{j=1}^{G}p(i,j)\right)_{|i-j|=n}$.

Correlation: $f_2=\left(1/\sigma_x\sigma_y\right)\sum_i\sum_j ijp(i,j)-\mu_x\mu_y$,

where μ_x and μ_y are the means, σ_x and σ_y are the standard deviations of the marginal distributions associated with $p(i,j)$.

Two information measures of correlation: $f_3=\left(HXY-HXY_1\right)/\max\{HX,HY\}$ and $f_4=\left(1-\exp\left[-2.0\left(HXY_2-HXY\right)\right]\right)^{1/2}$, where HX and HY are the entropies of p_x and p_y,

$HXY=-\sum_i\sum_j p(i,j)\log\left\{p(i,j)\right\}$,

$HXY_1=-\sum_i\sum_j p(i,j)\log\left\{p_x(i)p_y(j)\right\}$ and

$HXY_2=-\sum_i\sum_j p_x(i)p_y(j)\log\left\{p_x(i)p_y(j)\right\}$.

Maximal correlation coefficient:

$f_5=\left(\text{second largest eigvenvalue of } Q\right)^{1/2}$, where

$Q(i,j)=\sum_k p(i,k)p(j,k)/p_x(i)p_y(j)$.

The four values that each feature takes on in the four directions are averaged to produce a rotation-invariant feature which is employed by the classification system. Figure 3 shows typical examples of carina images and the corresponding calculated features.

3.3 Classification

In order to classify the video frames, we utilized a feed-forward artificial neural network classifier (ANN) which consists of three layers. The first (input) layer includes neurons that connect to selected features, the second layer includes hidden neurons, and the third (decision) layer includes one neuron that generates a likelihood score of a test case belonging to one of the three categories. To minimize over-fitting and maintain robustness of the ANN performance, a limited number of training iterations (1000), and a large ratio between the momentum (0.9) and learning rate (0.01), is used. The likelihood scores obtained by the ANN classifier in leave-one-subject-out tests are used to make the classification decision.

4 RESULTS

4.1 Classification of Cow Intubation Video Images

In order to perform a preliminary evaluation of the proposed system, we recorded two datasets. The first dataset includes a total of 10 intubation videos that were recorded from animal (cow) models, out of which 1600 images were extracted, visually inspected by a medical expert and classified into one of the following categories: upper-trachea (490 images), carina (550 images) and esophagus (560 images).

The second dataset includes 358 images, extracted from intubations performed on 8 human subjects that were downloaded from various web sites[1]. These images were also categorized into the three categories mentioned above.

Evaluation of the proposed approach was performed using a leave-one-subject-out validation method: in each iteration, the images extracted from all videos (for a particular dataset) but one were used to train the models, i.e. estimate the network parameters, and the images from the remaining video were used to test system performance. This process was repeated such that each image participated once in the testing phase.

The classification results are summarized in Tables 1 and 2, for the two datasets, respectively, where the rows represent the predicted (recognized) classes and the columns represent the actual classes. The system achieved an overall classification rate of 97.9% (1567 out of 1600 images) for the cow intubation database, and 97.5% (349 out of 358 images) for the human intubation database.

Specifically, most of the errors are due to incorrect classification of carina images as upper-trachea (e.g., 12 cases (2.2%), in the cow dataset and 2 cases (2%) in the human dataset), and incorrect classification of upper-trachea images as carina and esophagus (8 cases (1.7%), and 9 cases (1.8%), respectively, for the cow dataset; 2 cases (1.05%), and 2 cases (1.05%), respectively, for the human dataset). For both datasets, in two cases, an esophagus image was mistakenly classified as either upper-tracheal or carina.

[1]University of Florida: http://vam.anest.ufl.edu/airwaydevice/videolibrary/index.html and http://www.youtube.com

Table 1: Summary of classification results for the cow intubations dataset.

Recognized	Actual		
	Upper-trachea	Carina	Esophagus
Upper-trachea	473 (96.5%)	12 (2.2%)	1 (0.2%)
Carina	8 (1.7%)	536 (97.5%)	1 (0.2%)
Esophagus	9 (1.8%)	2 (0.3%)	558 (99.6%)
Total	490	550	560

Table 2: Summary of classification results for the human intubations dataset.

Recognized	Actual		
	Upper-trachea	Carina	Esophagus
Upper-trachea	185 (97.9%)	2 (2%)	1 (1.5%)
Carina	2 (1.05%)	98 (97.0%)	1 (1.5%)
Esophagus	2 (1.05%)	1 (1.0%)	66 (97.0%)
Total	189	101	68

4.2 Discussion

A novel approach for automatic endotracheal intubation confirmation was introduced. According to the approach, direct physical determination of the tube position with respect to the relevant anatomical structures is performed based on image classification. Images are represented using textural features which are utilized by the ANN classifier. The proposed scheme is simple and computationally efficient.

The proposed confirmation method was evaluated using cow and human intubations videos, out of which images were extracted and classified by a medical expert into one of three categories: upper tracheal, carina and esophagus. The method achieved a high precision of 97.9% (1567 out of 1600 images) using the cow intubations dataset, and 97.5% (349 out of the 358 images) using the human intubation dataset.

The method has a number of advantages over existing endotracheal intubation confirmation devices, including reliability in any medical condition, suitability for both esophageal intubation detection and one-lung intubation detection (although not tested in this preliminary study), and the fact that it is fully automatic and may be used, with a designated endotracheal tube, for continuous and long-distance screening of tube misplacement and dislodgment. The method can be easily integrated in all patient monitoring systems. Moreover, the system can be used to improve medical professionals training.

The proposed method is computationally efficient. Specifically, all of the algorithms used in this work were implemented in Matlab R2016a 64bit. Using a conventional PC equipped with Dual Intel Xeon 3.4 GHz with a 16 GBytes of RAM, feature extraction requires less than 1 second for each image. Future improvements are the inclusion of other anatomical landmarks, such as vocal cords, and the development of a video-analysis algorithm, which are expected to improve confirmation performances.

The results are encouraging, but clearly much work is needed to further validate the proposed approach. The available database consists of only 10 cow intubation videos and 8 human intubation videos. A much larger database is required in order to reliably validate system performance. Various factors might challenge the system performance, especially fog and secretions, which could result in poor image quality. In addition, the effect of possible physiological variability between patients on system performance is yet to be evaluated.

Our ultimate goal is to develop a reliable, cost-effective, easy to use and fully automatic device for confirmation of correct tube positioning. For this purpose, we plan to develop an advanced prototype, which will be thoroughly evaluated in pre-clinical trials and, upon receiving the appropriate regulatory approvals, on humans. Based on this preliminary study, we believe that implementation of the proposed method into a real-time confirmation system will lead to a major improvement in the ability to detect intubation incidents as they occur, while the patient is still well oxygenated and stable.

5 CONCLUSIONS

The ANN-based classification system achieved a high precision of 97.9% and 97.5% for the cow and human datasets, respectively. The results are encouraging but as mentioned above, more research is needed in order to reliably validate system performance. With these challenges in mind, successful implementation of the proposed method into a real-time confirmation system can serve as a major contribution to patient safety.

REFERENCES

Bhende, S. M., and Thompson, A. E., 1995. Evaluation of an end-tidal CO2 detector during pediatric cardiopulmonary resuscitation. *Pediat., 95*(3), 395-399.

Brunel, W., Coleman, D. L., and Schwartz, D. E., 1989. Assessment of routine chest reoentgenograms and the physical examination to confirm endotracheal tube position. *Chest, 96*, 1043-1045.

Gravenstein, J. S., Jaffe, M. B., and Paulus, D. A., 2004. *Capnograhpy clinical aspects*: Cambridge University Press.

Haralick, R. M., Shanmugam, M., and Dinstein, I., 1973. Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernatics, SMC-3*(6), 610-621.

Howells, T. H., 1985. Oesophageal misplacement of a tracheal tube. *Anaesthesia, 40*, 398-389.

Jacobs, L. M., Berrizbeitia, L. D., Bernnett, B., and Madigan, C., 1983. Endotracheal intubation in the prehospital phase of emergency medical care. *JAMA, 250*(2175-2177).

Jemmet, M. E., Kendal, K. M., and Fourre, M. W., 2003. Unrecognized misplacement of endotracheal tubes in a mixed urban to rural emergency medical services setting. *Acad. Emerg. Med., 10*, 961-965.

Jones, J. H., Murphy, M. P., and Dickson, R. L., 2004. Emergency physician-verified out-of-hospital intubation: miss rates by paramedics. *Acad. Emerg. Med., 11*, 707-709.

Katz, S. H., and Falk, J. L., 2001. Misplaced endotracheal tubes by paramedics in an urban emergency medical services system. *Ann. Emerg. Med., 37*, 32-37.

Klepper, I. D., Webb, R. K., Walt, J. V. D., Ludbrooks, G. L., and Cockings, J., 1993. The sthethoscope: Application and limintation- an analysis of 2000 incidents reports. *Anaesth. Intens. Care, 21*(5), 575-578.

Lederman, D., 2006. *An energy ratio test for one lung intubation detection.* Paper presented at the 18th Biennial International EURASIP conference, Brno, Czech Republic.

Lederman, D., 2011. Endotracheal intubation confirmation based on video image classification using a parallel GMMs framework- a preliminary evaluation. *Annals of Biomed. Eng., 39(1)*, 508-513.

Li, J., 2001. Capnography alone is imperfect for endotracheal tube placement confirmation during emergency intubation. *J. Emerg. Med., 20*(3), 223-229.

Linko, K., Paloheimo, M., and Tammisto, T., 1983. Capnograhphy for detection of accidental oesophageal intubation. *Acta Anaesthesiol. Scand., 27*, 199-202.

Nolan, J. P., Deakin, C. D., and Soar, J., 2005. European resuscitation council guidelines for resuscitation. *Resuscitation, 67*(1), S39-S86.

O'connor, C. J., Mansy, H., Balk, R. A., Tuman, K. J., and Sandler, R. H., 2005. Identification of endotracheal tube malpopsitions using computerized analysis of breath sounds via electronic stethoscopes. *Anesth. Analg., 101*(3), 735-739.

Owen, R. L., and Cheney, F. W., 1987. Endobronchial Intubation: a Preventable Complication.

Anesthesiology, 67, 255-257.
Peterson, A. W., and Jacker, L. M., 1973. Death following inadvertent esophageal intubation: a case report. *Anesth. Analg., 52*, 398-401.
Pointer, J. E., 1988. Clinical characteristics of paramedics' performance of endotracheal intubation. *J. Emerg. Med., 6*(505-509).
Silvestri, S., Ralls, G. A., and Krauss, B., 2005. The effectiveness of out-of-hospital use of continuous end-tidal carbon dioxide monitoring on the rate of unrecognized misplaced intubation within a regional emergency medical services system. *Ann. Emerg. Med., 45*, 497-503.
Steward, R. D., Paris, P. M., and Winter, P. M., 1984. Field endotracheal intubation by paramedical personnel: success rates and complications. *Chest, 85*(341-345).
Tejman-Yarden, S., Lederman, D., Weksler, N., and Gurman, G., 2006. Acoustic monitoring of double lumen ventilated lungs for the detection of selective unilateral lung ventilation. *Anesth. Analg., 103*, 1489-1492.
Tejman-Yarden, S., Zlotnik, A., Weizman, L., Tabrikian, J., Cohen, A., Weksler, N., and Gurman, G. M., 2007. Acoustic monitoring of lung sounds for the detection of one-lung intubation. *Anesth. Analg., 105*(2), 397-404.
Timmermann, A., Russo, S. G., Eich, C., Roessler, M., Braun, U., Rosenblatt, W. H., and Quintel, M., 2007. The out-of-hospital esophageal and endobronchial intubations performed by emergency physicians. *Crit. Care and Trauma, 104*(3), 619-623.
Vergese, S. T., Hannallah, R. S., Slack, M. C., Cross, R. R., and Patel, K. M., 2004. Auscultation of bilateral breath sounds does not rule out endobronchial intubation in children. *Anesth. Analg., 56-58*.
Wang, H. E., Cook, L. J., Chang, C. H., Yealy, D. M., and Lave, J. R., 2009. Outcomes after out-of-hospital endotracheal intubation errors. *Resuscitation, 80*(1), 50-55.
Wang, H. E., Lave, J. R., Sirion, C. A., and Yealy, M., 2006. Paramedic intubation errors: isolated events or symptoms of larger problems? *Health Affairs, 25*(2), 501-509.
Webb, R. K., Walt, J. H. V. D., Runciman, W. B., Williamson, J. A., Cockings, J., Russel, W. J., and Helps, S., 1993. Which monitor? an analysis of 2000 indicent reports. *Anaesth. Intens. Care, 21*(5), 529-542.
Weizman, L., Tabrikian, J., and Cohen, A., 2008. Detection of one-lung intubation incidents. *Annals of Biomed. Eng., 36*(11), 1844-1855.
Wodicka, G. R., DeFrain, P. D., and Kraman, S. S., 1994. Bilateral asymmetry of respiratory acoustic transmission. *Med. Biol. Eng. Comp., 32*(5), 489-494.
Yap, S. J., Morris, R. W., and Pybus, D. A., 1994. Alterations in endotracheal tube position during general Anesthesia. *Anaesth. Crit. Care, 586-588*.
Zwillich, C. W., Pierson, D. J., and Creagh, C. E., 1974. Complications of assisted ventilation, a prospective study of 354 consecutive episodes. *Am. J. Med., 57*, 161-170.