BJA

British Journal of Anaesthesia 115 (1): 89–98 (2015)

doi: 10.1093/bja/aev028
Advance Access Publication Date 3 March 2015
Critical Care

CRITICAL CARE

Multifaceted bench comparative evaluation of latest intensive care unit ventilators

M. Garnier[1,4,*], C. Quesnel[1,4], J.-P. Fulgencio[1], M. Degrain[3], G. Carteaux[7], F. Bonnet[1,4], T. Similowski[2,4,5], and A. Demoule[2,4,6]

[1]Anaesthesiology and Intensive Care Department, Hôpital Tenon, [2]Medical Intensive Care Unit and Respiratory Division, Groupe Hospitalier Pitié-Salpêtrière, [3]Agence Générale des Equipements et Produits de Santé, APHP, Paris, France, [4]Faculté de Médecine Pierre & Marie Curie, [5]ER10, [6]INSERM U974, Université Pierre et Marie Curie, Paris, France, and [7]Medical Intensive Care Unit, Centre Hospitalier Albert Chenevier-Henri Mondor, APHP, Créteil, France

*Corresponding author: Département d'anesthésie et réanimation – Pôle TVAR, Hôpital Tenon, 4 rue de la Chine, 75020 Paris, France. E-mail: marcgarnier@gmail.com

Abstract

Background: Independent bench studies using specific ventilation scenarios allow testing of the performance of ventilators in conditions similar to clinical settings. The aims of this study were to determine the accuracy of the latest generation ventilators to deliver chosen parameters in various typical conditions and to provide clinicians with a comprehensive report on their performance.

Methods: Thirteen modern intensive care unit ventilators were evaluated on the ASL5000 test lung with and without leakage for: (i) accuracy to deliver exact tidal volume (V_T) and PEEP in assist-control ventilation (ACV); (ii) performance of trigger and pressurization in pressure support ventilation (PSV); and (iii) quality of non-invasive ventilation algorithms.

Results: In ACV, only six ventilators delivered an accurate V_T and nine an accurate PEEP. Eleven devices failed to compensate V_T and four the PEEP in leakage conditions. Inspiratory delays differed significantly among ventilators in invasive PSV (range 75–149 ms, P=0.03) and non-invasive PSV (range 78–165 ms, P<0.001). The percentage of the ideal curve (concomitantly evaluating the pressurization speed and the levels of pressure reached) also differed significantly (range 57–86% for invasive PSV, P=0.04; and 60–90% for non-invasive PSV, P<0.001). Non-invasive ventilation algorithms efficiently prevented the decrease in pressurization capacities and PEEP levels induced by leaks in, respectively, 10 and 12 out of the 13 ventilators.

Conclusions: We observed real heterogeneity of performance amongst the latest generation of intensive care unit ventilators. Although non-invasive ventilation algorithms appear to maintain adequate pressurization efficiently in the case of leakage, basic functions, such as delivered V_T in ACV and pressurization in PSV, are often less reliable than the values displayed by the device suggest.

Key words: bench study, mechanical ventilation, non-invasive ventilation, quality improvement, ventilator performance

Accepted: November 25, 2014

Editor's key points

- Inaccurate data displayed by artificial ventilators can affect clinical decision making and patient care in the intensive care unit.
- This bench study evaluated the effect of varying leakage, resistance, and compliance during different ventilatory modes in 13 modern ventilators.
- There were significant inaccuracies, and variations in performance between different modern ventilators and modes.
- Clinicians should be aware of possible inaccuracies in the data displayed by ventilators.

Evaluation of the technical characteristics of intensive care unit (ICU) ventilators is essential to determine the strengths and weaknesses of each ventilator. When manufacturers provide these characteristics, they may be very different from their performance in clinical settings, making independent bench studies mimicking these conditions necessary. In addition, as devices constantly evolve, regular updates are necessary.

During the past 10 years, bench evaluation studies have investigated three major features of ICU ventilators.[1–14] The first is the accuracy of delivered tidal volume (V_T) and PEEP during assist-control ventilation (ACV). This is a critical point because these two settings are the cornerstones of protective ventilation, which is beneficial to the duration of mechanical ventilation and mortality in patients with acute respiratory distress syndrome.[15–17] The second is the performance of trigger and pressurization in spontaneously breathing patients receiving pressure support ventilation (PSV). Poor trigger and pressurization characteristics may increase the work of breathing[18] and induce patient–ventilator asynchronies that are associated with a poorer prognosis.[19–21] The third is the ability of non-invasive ventilation (NIV) algorithms to compensate the deleterious impact of leaks[2 10 22] that could impact NIV tolerance and success.

Most bench studies have focused on only one of these three aspects; however, ICU ventilators are expected to be polyvalent and used indifferently for protective ACV, for PSV during weaning, and for NIV. In the present study, therefore, we provide an updated global appraisal of 13 of the latest versions of ICU ventilators, including several original aspects. Our primary aim was to determine the accuracy of latest generation ventilators to deliver set values in different simulated typical conditions. Our secondary aim was to give clinicians a comprehensive and relevant report on these performances.

Methods

We evaluated 13 latest generation ICU ventilators, recently revised according to the manufacturers' recommendations, as follows: extend and MonnalT75 (Air Liquide Medical System, Antony, France); V500 and Savina300 (Dräger, Lübeck, Germany); Engström (GE Healthcare, Fairfield, CT, USA); C1, C2, and S1 (Hamilton Medical, Rhäzüns, Switzerland); Servo-i and Servo-s (Maquet, Rastatt, Germany); PB840 (Covidien, Dublin, Ireland); and Avea and Vela (Vyasis®, CareFusion, San Diego, CA, USA; see Supplementary material, Table S1, which provides the main manufacturers' specifications and the software version of each ventilator tested).

Each ventilator was tested using an Active Servo Lung 5000 test lung (ASL5000; IngMar Medical, Pittsburg, PA, USA). During ventilation in leakage conditions, the test lung was connected to the manufacturer's leakage generator (Simulator Bypass and Leak Valve Module; IngMar Medical), allowing three levels of non-adjustable continuous leakage (mild, moderate, and severe leakage), and to a home-made calibrated leakage generator, allowing leaks >7 cm H_2O of pressurization, as previously described[2] (see also the Supplementary material, Methods for supplemental methods and a schema of the bench test). The mild and important continuous leakage generated, respectively, a leakage of about 4–5 and 32–33 litres min^{-1} at a mean airway pressure of 10 cm H_2O for the normal compliance and resistance simulated lung condition.

The design of the entire experiment is represented in Fig. 1. The following tree ventilatory modes were evaluated: (i) assist-control ventilation (ACV); (ii) pressure support ventilation without non-invasive ventilation algorithm (I-PSV); and (iii) pressure support ventilation with non-invasive ventilation algorithm (NI-PSV). In ACV, we assessed the accuracy of delivered V_T and PEEP levels. In I-PSV and NI-PSV, we assessed trigger and pressurization quality and the accuracy of delivered pressure support levels and PEEP levels (Fig. 1).

Assist-control ventilation

Tidal volume and PEEP concordance were evaluated with and without mild continuous leakage, simulating a low level of leakage around the tracheal tube cuff. Regarding resistance and compliance, five lung conditions were simulated, mimicking the major typical anomalies observable in patients, such as the low compliance of acute respiratory distress syndrome patients or the high resistance of chronic obstructive pulmonary disease patients (Fig. 1).

Invasive and non-invasive pressure support ventilations

Trigger and pressurization performance were evaluated with 10 and 20 cm H_2O inspiratory efforts, reproducing normal to moderate and moderate to strong inspiratory pressures (Fig. 1). Trigger and pressurization performance were evaluated using the following six indexes (see Supplementary material, Fig. S1 for graphical representations): (i) the triggering delay ($\triangle T$), defined as the time between the beginning of the inspiratory effort and the beginning of ventilator pressurization; (ii) the pressurization delay (PD), defined as the time after the beginning of ventilator pressurization during which the pressure remains below baseline pressure; (iii) the inspiratory delay (ID), defined as the sum of $\triangle T$ and PD; (iv) the pressure–time product at 0.3 s (PTP0.3), defined as the net area under the pressure–time curve throughout the 0.3 s after the onset of the inspiratory effort; (v) the pressure concordance, defined as 'preset pressure minus pressure reached'; and (vi) the percentage of the ideal curve (PIC), defined as the net area under the pressure–time curve divided by the same area calculated for a perfect square-shaped pressure–time curve. A low PIC could be related to low pressure concordance, low PTP0.3, or both (Supplementary material, Fig. S1D).

Pressure concordance was evaluated for four levels of pressure support (5, 10, 15, and 20 cm H_2O), while PEEP concordance was evaluated for four levels of PEEP (0, 5, 10, and 15 cm H_2O). Invasive pressure support ventilation was evaluated for the five lung conditions, while NI-PSV measures were evaluated in normal, moderately obstructive, and moderately restrictive conditions (Fig. 1).

Measurements

Flow was measured with a Fleisch pneumotachograph (Hans Rudolph, Kansas City, MO, USA) and airway pressure by a

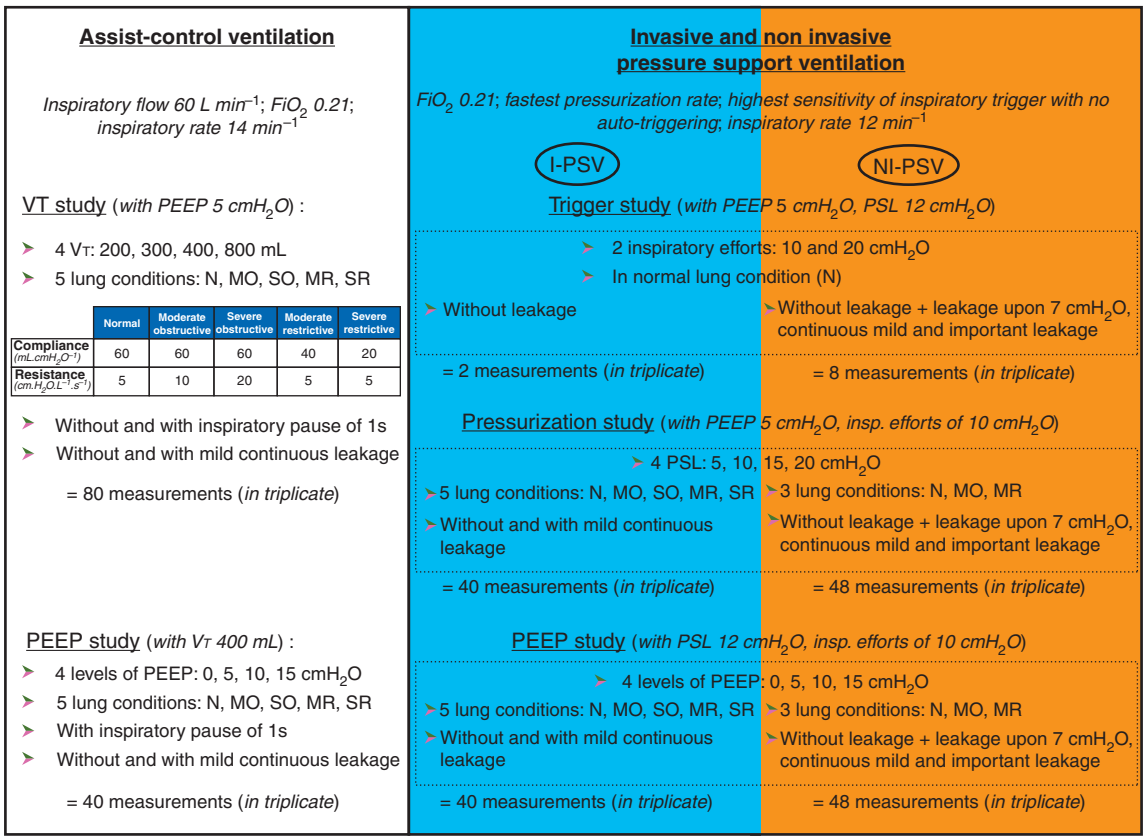

Fig 1 Study protocol. In assist-control ventilation, the study protocol focused on concordance of delivered tidal volumes (V_T) and PEEP. Four values of V_T and four levels of PEEP were tested using five simulated compliance and resistance lung conditions (normal [N], moderate [MO] and severe obstructive [SO], and moderate [MR] and severe restrictive [SR]), with and without an inspiratory pause of 1 s, and with and without mild continuous leakage. In invasive (I-PSV) and non-invasive (NI-PSV) pressure support ventilations, the study protocol focused on trigger and pressurization performance, and delivered pressures and PEEP. Trigger was studied in normal lung conditions for two intensities of inspiratory efforts, without leakage in I-PSV, and with and without three different levels of leakage for NI-PSV. Pressurization capacities and PEEP concordance were studied for, respectively, four levels of pressure support (PSL) and four levels of PEEP, in I-PSV using the five lung conditions with and without continuous mild leakage, and in NI-PSV using the three less severe lung conditions with and without mild and severe continuous leakage, and with leakage upon 7 cm H_2O of pressure. All the measures were recorded and analysed for three breaths, after a minimum of five cycles to reach a steady state.

pressure transducer (DP 15-32; Validyne, Northridge, CA, USA), both inserted between the test lung and the Y-piece of the ventilation circuit. Measures were made in ambient temperature and pressure dry (ATPD) conditions, and then converted into body temperature and pressure saturated (BTPS) conditions as described elsewhere[8] (see also the Supplementary material, Methods for additional methods and details about the conversion formula).

Global appraisal

The main characteristics of each ventilator were summarized using radar charts. Details about the calculation of the proportion of the value from the best ventilator for each characteristic are available in the Supplementary material, Methods.

Statistical analysis

Measures were performed after a minimum of five cycles to reach a steady state and to allow ventilators to react fully to leakage. Each value is the mean of three breaths. As ventilators are intended to supply all types of patients, data obtained for the various lung conditions were pooled for analysis and presented as means (SD).

Delivered values of V_T between efficient and less reliable ventilators were compared using the Mann–Whitney *U*-test. Comparisons of the performance of each ventilator with and without mild continuous leakage in invasive modes were assessed with the paired Wilcoxon test. Finally, comparisons between PSV parameters according to the different ventilators, or the different levels of leakage in NI-PSV, were assessed with the Friedman test with Dunn's correction. A value of $P<0.05$ was considered significant.

Results

Assist-control ventilation

Evaluation of tidal volume

Without leakage, only six ventilators delivered a mean V_T included in the tolerance range 'set V_T [−10% to +10%]', significantly higher than the seven ventilators that exceeded this range

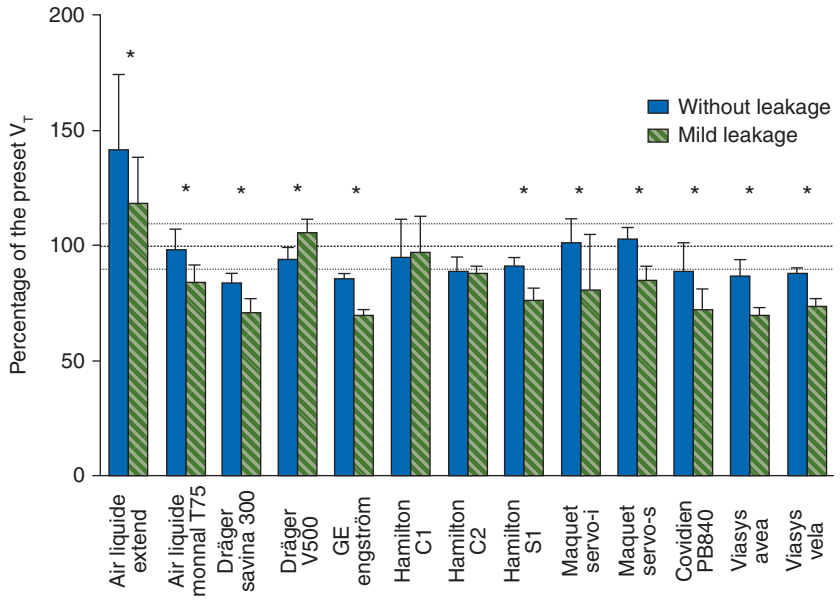

Fig 2 Proportion of the preset tidal volumes (VT) really delivered, without leakage and with mild continuous leakage. For each condition (with and without leakage), data are the means of measured VT in 40 conditions: four preset values of VT (200, 300, 400, and 800 ml) in five lung conditions (normal, moderate and severe obstructive, and moderate and severe restrictive), with and without inspiratory pause. Error bars represent SD of the means. Dotted horizontal lines represent the tolerance range 'set VT [−10% to +10%]'. *P<0.05 between conditions without leakage and with mild continuous leakage.

(P=0.03). The Extend ventilator exhibited the highest discrepancy between the set VT and the VT recorded (142% of the set VT; Fig. 2; see also Supplementary material, Table S2, which provides the numerical results). The addition of mild continuous leakage did not affect the VT only for the C1 and C2 ventilators. On the contrary, the V500 ventilator delivered a higher VT than without leakage, while the 10 remaining ventilators undercompensated the decrease of VT (Fig. 2 and Supplementary material, Table S2).

Evaluation of PEEP

Without leakage, nine ventilators delivered levels of PEEP included in the tolerance range 'set PEEP [−1 to +1 cm H_2O]'. Four ventilators (Extend, S1, PB840, and Vela) delivered PEEP levels outside this range for at least one of the four PEEP levels studied (Fig. 3). The addition of mild continuous leakage led to a proper compensation of PEEP for nine ventilators, while Extend, Servo-I, Servo-s, and Avea failed to maintain the PEEP effectively, which fell by >1 cm H_2O for at least one of the four PEEP levels studied (Fig. 3 and Supplementary material, Table S2).

Pressure support modes (I-PSV and NI-PSV)

Evaluation of trigger function

In I-PSV, the inspiratory delay (ID) differed significantly among ventilators (P=0.03). Mean ID was 100 (20) ms [range 75 (6) to 149 (1) ms; Fig. 4 and Supplementary material, Table S2]. The ID was shorter than 100 ms in eight ventilators and was close to 150 ms for Savina300. The two components of ID also differed among ventilators; mean triggering delay (ΔT) was 79 (10) ms [range 63 (5) to 98 (20) ms, P=0.09], whereas mean pressurization delay (PD) was 20 (19) ms [range 7 (0) to 78 (4) ms, P=0.04]. The ΔT was below 75 ms for only five ventilators (Extend, Savina300, V500, C1, and C2).

In NI-PSV, ID also differed significantly among ventilators (P<0.001). Mean ID was 111 (25) ms [range 78 (10) to 165 (32) ms; Fig. 4 and Supplementary material, Table S2). The ID was <100 ms in only six ventilators and exceeded 150 ms for PB840, which was also unable to detect the 10 cm H_2O inspiratory effort in the presence of a severe continuous leakage. The two components of ID also differed significantly among ventilators; mean ΔT was 93 (22) ms [range 69 (8) to 136 (30) ms, P<0.001], whereas mean PD was 18 (8) ms [range 9 (2) to 37 (16) ms, P<0.001]. The ΔT was below 75 ms for only three ventilators (Savina300, Engström, and Avea) and exceeded 100 ms for four ventilators (S1, Servo-i, Servo-s, and PB840). Nine ventilators maintained similar ID according to the four various conditions of leakage, whereas the four remaining ventilators showed a significantly increased ID with increasing leakage conditions (MonnalT75, PB840, Avea, and Vela; see Supplementary material, Fig. S2, which illustrates the impact of the different levels of leakage on trigger sensibility in NI-PSV).

Evaluation of pressurization capacities

In I-PSV, the percentage of the ideal curve (PIC) differed significantly among ventilators (P=0.04). Mean PIC was 71 (9)% [range 57 (16) to 86 (1)%] and exceeded 75% (most efficient quartile) only for Savina300, Engström, C1, and Servo-i (Fig. 5 and Supplementary material, Table S2). As shown in Fig. 5, a great heterogeneity among ventilators existed regarding the two main PIC components (PTP0.3 and pressure concordance).

In NI-PSV, PIC also differed significantly among ventilators (P<0.001). The mean PIC was 74 (10)% [range 60 (8) to 90 (1)%] and exceeded 75% only for six ventilators (Savina300, V500, Engström, C1, C2, and S1; Fig. 5 and Supplementary material, Table S2). The PIC varied significantly when leaks were applied (P =0.01), with a significant decrease in PIC during important

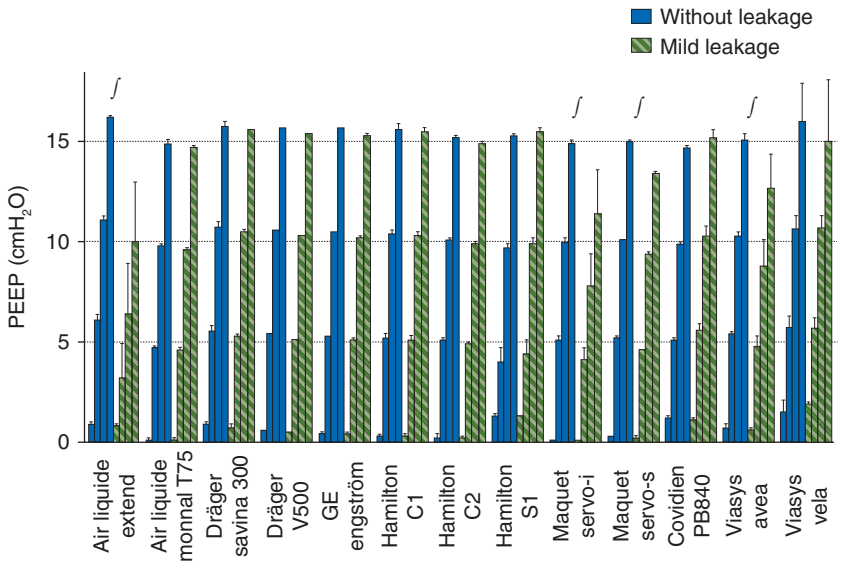

Fig 3 Delivered levels of PEEP in assist-control ventilation without leakage and with mild continuous leakage. For each leakage condition, data are the means of measured PEEP in 20 conditions: four preset values of PEEP (0, 5, 10, and 15 cm H_2O) in five lung conditions (normal, moderate and severe obstructive, and moderate and severe restrictive). Error bars represent SD of the means. ∫: loss of PEEP exceeding 1 cm H_2O between conditions without leakage and with mild continuous leakage for at least one of the four preset values of PEEP.

leakage compared with the condition without leakage. This resulted in particular from Extend, PB840, and Vela, which showed an absolute decrease of PIC above 5% between these two conditions (Supplementary material, Fig. S3).

Evaluation of PEEP

In I-PSV, eight ventilators delivered levels of PEEP included in the tolerance range 'set PEEP [−1 to +1 cm H_2O]'. Five ventilators (Extend, S1, PB840, Avea, and Vela) delivered PEEP outside this range for at least one of the four PEEP levels studied. In the presence of mild continuous leakage, Extend, Servo-i, Servo-s, and Avea failed to maintain the PEEP effectively, which decreased by >1 cm H_2O for at least one of the four PEEP levels studied (Supplementary material, Fig. S4).

In NI-PSV, all ventilators except S1 and Vela delivered levels of PEEP included in the tolerance range 'set PEEP [−1 to +1 cm H_2O]'. All ventilators except Avea efficiently prevented leakage-induced loss of PEEP (Supplementary material, Fig. S4).

Global appraisal

Figure 6 shows radar charts that summarize the main characteristics of each ventilator. It is noticeable that if most ventilators showed satisfactory overall performance, others appeared less efficient. Engström, V500, and C1 appear to be the three most accurate and homogeneous ventilators, exceeding 80% of the value for the best ventilator for all the main characteristics used for construction of the radar chart.

Discussion

Our main findings may be summarized as follows: (i) heterogeneity remains among the performance of the latest generation ICU ventilators; (ii) accuracies of delivered VT in ACV and pressurization in PSV are often less reliable than displayed by the devices; and (iii) NIV algorithms properly attenuate the deleterious impact of leaks on pressurization.

Heterogeneity of recent intesive care unit ventilators and clinical implications

Despite technological progress and the continuous development of ventilators, our multifaceted experiments showed substantial differences among ventilators. Furthermore, the performance of a given device may differ according to ventilatory modes. Given that the purchase of a ventilator is challenging because the cost is high and the device is supposed to last several years, it is important for physicians to be aware of the strengths and weaknesses of each ventilator. Indeed, our study may help physicians to opt for the most appropriate ventilator with regard to the typology of patients and the clinical situations found in their ICU. For example, it seems unsuitable to use certain ventilators for protective ventilation if they deliver a VT exceeding the set VT, sometimes by more than 40%. It also seems inappropriate to deliver NIV or to conduct weaning with a ventilator unable to detect weak inspiratory efforts or characterized by long triggering delays and weak pressurization capacities.

Evolution of ventilator performance over years

Considering PSV trigger performance, no significant progress has been made between 2006[13] and 2013. Indeed, no ventilator could detect inspiratory efforts faster than 50 ms, often exceeding 100 ms when moderate inspiratory efforts were simulated. Considering that Thille and colleagues[13] found results close to those published by Richard and colleagues,[23] trigger performance in PSV has not improved significantly during the last decade, with the exception of devices such as Neurally Adjusted Ventilatory Assist (NAVA). The same observation could be made

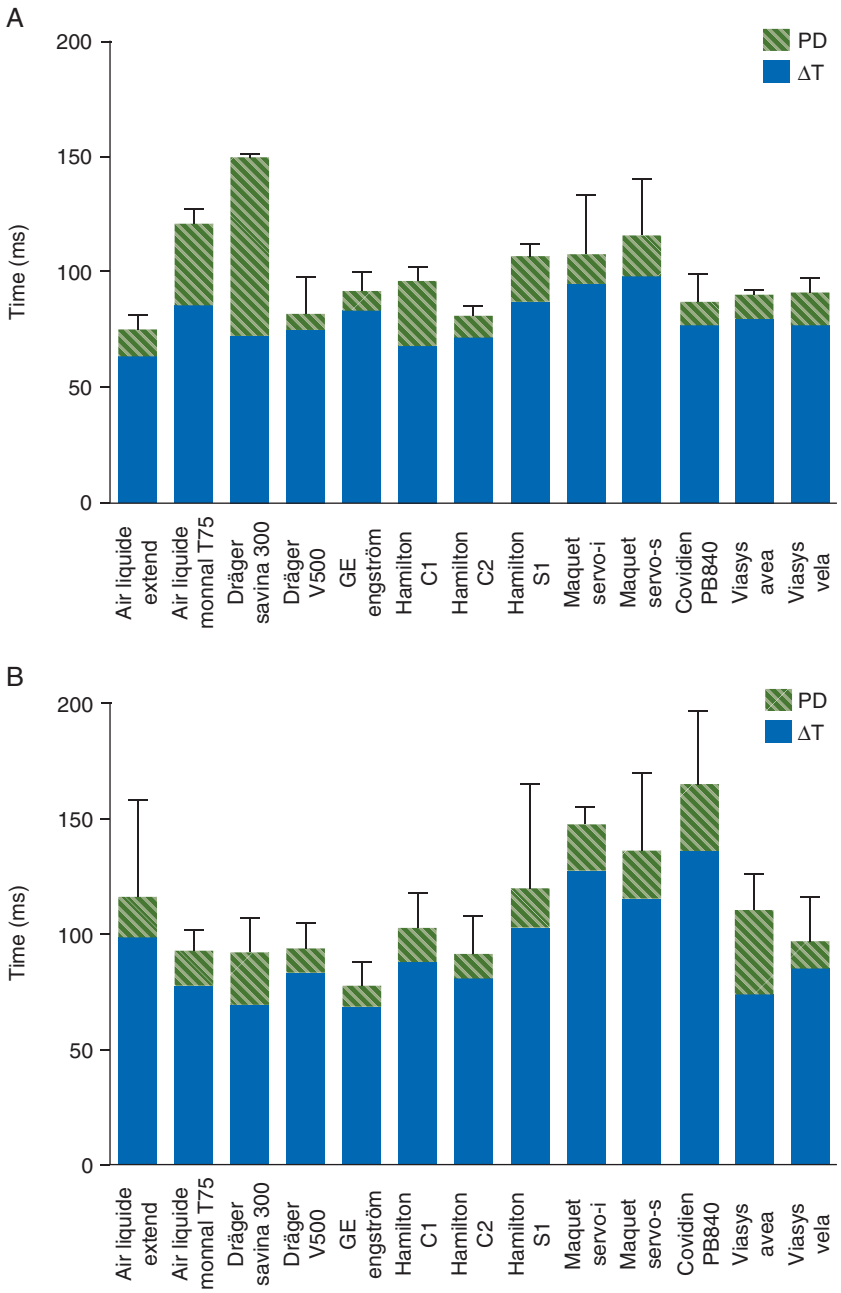

Fig 4 Trigger performance in invasive (I-PSV; A) and non-invasive (NI-PSV; B) pressure support ventilations. (A) Inspiratory delays (ID) in I-PSV in response to moderate and strong inspiratory efforts without leakage. The ID is represented as the sum of its two components, triggering delay (ΔT) and pressure delay (PD). Data are the means, with error bars representing SD. (B) Inspiratory delays, ΔT and PD, in NI-PSV in response to moderate and strong inspiratory efforts, without leakage and with calibrated, mild and severe continuous leakage. Data are the means, with error bars representing SD.

about the difference between preset and delivered V_T, with volume errors often exceeding 10% of the preset V_T in past and present studies.[8] These inaccuracies and their heterogeneity among devices are comparable to those recently found in anaesthetic ventilators,[24] suggesting that technical limits have been reached.

Although inaccuracy of V_T is a critical issue, this may not be the case for response times, which are in fact acceptable. The mean triggering delay in healthy subjects breathing in NIV during 'comfort conditions' is ~300 ms.[25] Moreover, as the time from the initiation of inspiratory effort and apparition of a sensation of dyspnoea in healthy subjects is ~150 ms,[26] it is questionable whether trying to reduce inspiratory triggers below 100 ms is a clinically relevant approach. Indeed, decreasing mean inspiratory delay from 200–220 to 50–60 ms using NAVA has no effect on dyspnoea visual analogue scale values.[27]

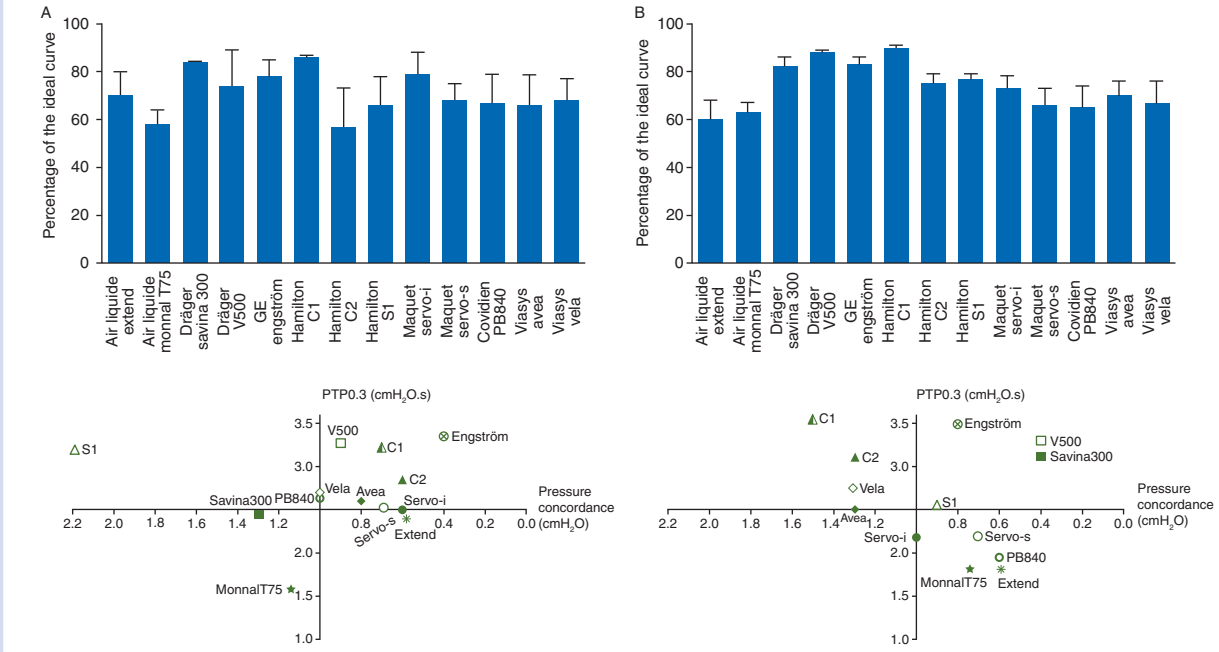

Fig 5 Pressurization performance in invasive (I-PSV; A) and non-invasive (NI-PSV; B) pressure support ventilations. Top panels show the percentage of the ideal curve, defined as the net area under the pressure–time curve divided by the same area calculated for a perfect square-shaped pressure–time curve, for I-PSV (A) and NI-PSV (B). Data are the means, with error bars representing SD. In the bottom panels, each ventilator is represented as a point whose abscissa and ordinate are, respectively, the pressure concordance (defined as preset pressure minus pressure reached) and the pressure–time product at 0.3 s (PTP0.3), for I-PSV (A) and NI-PSV (B). Of note, a low pressure concordance, a low PTP0.3, or both may explain a low percentage of the ideal curve.

In consequence, although some future gains in technical performance, such as accuracy of VT, will be crucial, others will not necessarily be useful for clinical practice and patient care. Performance goals for ventilators, taking into account technical but also physiological aspects combined with patients' experiences, and defined with an active involvement of clinicians, are still lacking.[28] Collaborative definitions of these targets by experts are needed and will be very helpful in future studies to make relevant ventilator rankings. Our study may serve as a starting point for this reflection.

To our knowledge, our study is the first to provide a multifaceted bench evaluation of ICU ventilator performance during simulation of typical clinical conditions. We aimed to present technical results comprehensibly and to highlight that clinicians should be aware that both the international standards set for accuracy of devices and the values delivered by their equipment may not be what they might expect.[28] Hence, we have chosen the PIC as the main descriptor of pressurization performance. This integrative descriptor, previously used in several studies,[14 29–33] simultaneously evaluates the pressurization rate and the level of preset pressure reached. Ideal pressurization capacities mix fast pressurization and inspiratory pressure close to the set pressure, giving a PIC close to 100%. We believe that this is meaningful for physicians and easily comprehensible, because a high percentage means a pressure support close to that delivered by an ideal ventilator, whereas a low percentage means poor ability to support patients' efforts.

We also present a global appraisal of ventilator performance using radar charts, constructed with the most relevant endpoints. Radar charts are useful for displaying multivariate observations. They allow the observer to see at a glance the main performance of a ventilator and to compare ventilators among themselves, easily pointing out outliers. These representations are classically used in the control of quality improvement or to point out strengths and weaknesses, which is exactly what clinicians are entitled to expect of bench studies.

Finally, we tried to reproduce clinical settings as faithfully as possible. For example, we tested ventilators in invasive modes with mild leakage for the first time, because this is a fairly frequent situation in ICU or anaesthesia clinical practice.[34 35]

The main limitation of our study concerns the ability to extrapolate our results obtained from the lung model to real situations. Calibrated inspiratory efforts were generated, which may be different from those observed in real patients. In the same manner, airway leaks generated were continuous and mimicked real conditions only partly. However, this test lung has been deemed appropriate to simulate the characteristics of real patients and has been widely used.[1 4 36–38] It offers the advantage of standardizing a broad spectrum of simulated mechanical characteristics. Besides, it would be ethically questionable to test so wide a range of tidal volumes or pressure levels on the same given patient. Nevertheless, clinical studies conducted in patients are complementary to bench studies, to evaluate patient–ventilator synchrony, comfort, and efficiency of interfaces in NIV. In fact, few studies have focused on both aspects together. Recently, Carteaux and colleagues[2] reported that application of leaks during NIV (with the same model as that used in our study) gave concordant results between bench and patients for the parameters evaluated, suggesting some clinical relevance of bench tests.

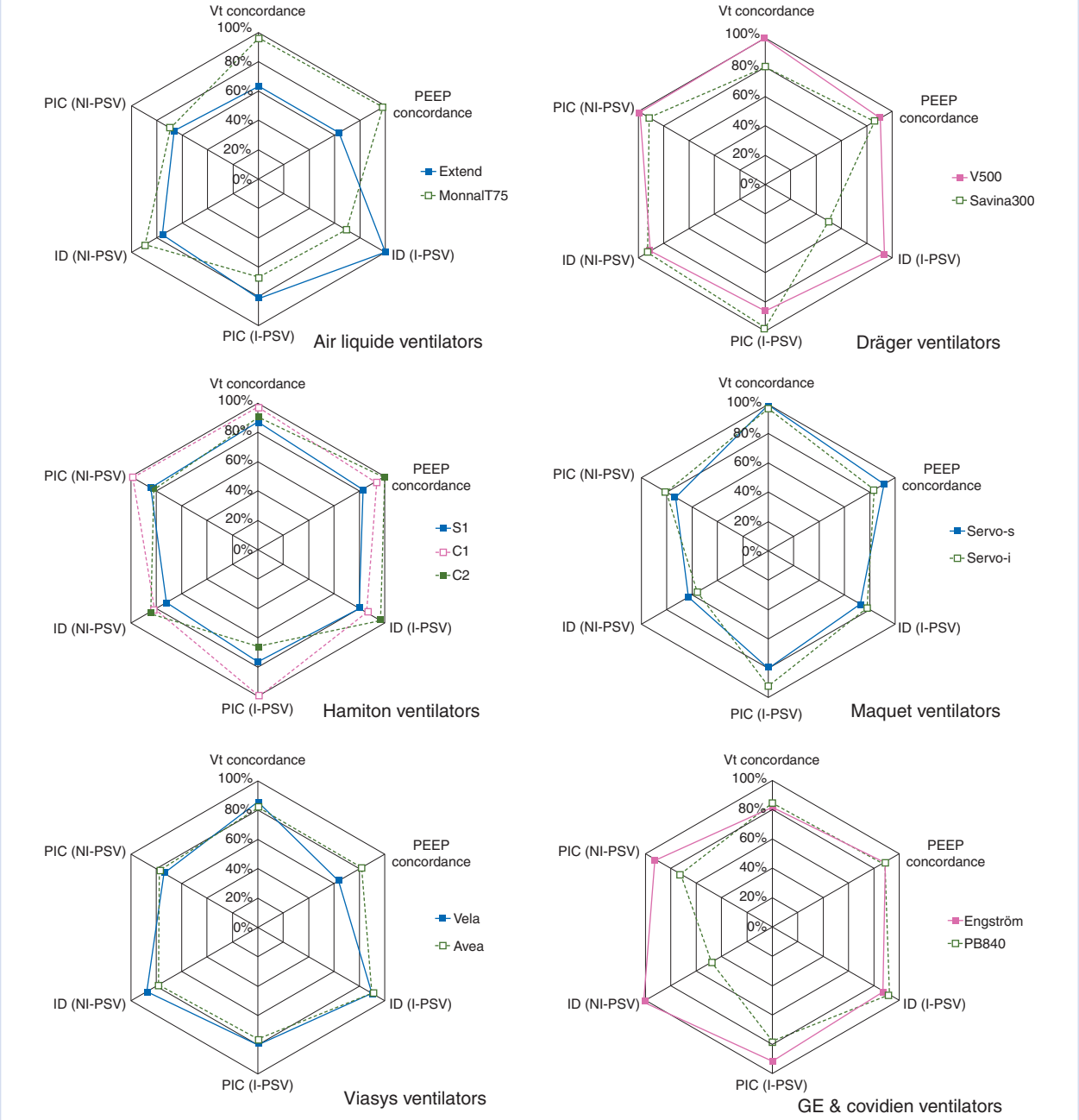

Fig 6 Radar charts summarizing the main characteristics of the ventilators. Data are expressed as a proportion of the value from the best ventilator for the considered characteristics (see the Supplementary material, Methods for details about the calculation of the proportion of the best ventilator for each characteristic). When several ventilators from the same manufacturer where studied, their characteristics are superimposed on the same graph. The following six characteristics are summarized for each ventilator: tidal volume (V_T) concordance (mean of the 80 conditions, with a coefficient of 0.5 for results obtained during mild continuous leakage condition), PEEP concordance [mean of 40 conditions in assist-control ventilation, 40 conditions in invasive (I-PSV) and 48 conditions in non-invasive (NI-PSV) pressure support ventilations, with a coefficient of 0.5 for results in assist-control ventilation and I-PSV obtained during mild continuous leakage condition], inspiratory delays in I-PSV and NI-PSV (ID, mean of two and eight conditions, respectively), and percentage of the ideal curve in I-PSV and NI-PSV (PIC, mean of two and eight conditions, respectively). The tracks of the three most accurate and homogeneous ventilators (exceeding 80% of the value from the best ventilator for all the characteristics) are represented in pink.

In conclusion, bench evaluation of ventilators in conditions mimicking clinical settings is a more informative approach to assess ventilator performance than using basic characteristics provided by manufacturers. Indeed, despite several improvements, the delivered V_T, PEEP, and pressure support levels are often less reliable than the devices suggest. These data suggest that the clinical consequences of discrepancies among modern ventilators cannot be ruled out and need to be considered.

Supplementary material

Supplementary material is available at *British Journal of Anaesthesia* online.

Acknowledgements

The authors acknowledge the 'Agence Générale des Equipements et Produits de Santé (AGEPS)' from the APHP institution (Paris, France) and the ER10 unit of the Pierre & Marie Curie University (Paris, France) for having provided the technical material used in this study. The authors also acknowledge all the members of the 'Groupe de Travail sur les Respirateurs (GTR)' from APHP institution. The authors specially acknowledge Hugues Pernel for his contribution. The authors thank Joanna Shore for English language proofreading.

Authors' contributions

M.G. conceived the protocol, realized the devices' evaluation on the bench test, analysed the collected data, and drafted the manuscript. C.Q. participated in the protocol conception, analysed the collected data, and corrected the manuscript. J.-P.F. participated in the protocol conception and evaluation of the devices on the bench test. M.D. participated in the evaluation of the devices on the bench test. A.D. corrected the protocol and helped M.G. to draft the manuscript. G.C. and F.B. corrected the manuscript. All authors have read and approved the final manuscript.

Declaration of interest

M.G., C.Q., and J.-P.F. report an invitation to a seminar on mechanical ventilation on the animal model in Clermont-Ferrand (France) in 2011, provided by General Electric Healthcare. A.D. is the principal investigator of a clinical study supported by Maquet; he has been a member of a board sponsored by Covidien and has received lecture fees from Maquet and Covidien. M.D., G.C., T.S., and F.B. report no conflict of interest in relation with the subject of the manuscript.

References

1. Borel JC, Sabil A, Janssens J-P, *et al*. Intentional leaks in industrial masks have a significant impact on efficacy of bilevel noninvasive ventilation: a bench test study. *Chest* 2009; **135**: 669–77
2. Carteaux G, Lyazidi A, Cordoba-Izquierdo A, *et al*. Patient-ventilator asynchrony during noninvasive ventilation: a bench and clinical study. *Chest* 2012; **142**: 367–76
3. Fauroux B, Leroux K, Desmarais G, *et al*. Performance of ventilators for noninvasive positive-pressure ventilation in children. *Eur Respir J* 2008; **31**: 1300–7
4. Ferreira JC, Chipman DW, Kacmarek RM. Trigger performance of mid-level ICU mechanical ventilators during assisted ventilation: a bench study. *Intensive Care Med* 2008; **34**: 1669–75
5. Jaber S, Tassaux D, Sebbane M, *et al*. Performance characteristics of five new anesthesia ventilators and four intensive care ventilators in pressure-support mode: a comparative bench study. *Anesthesiology* 2006; **105**: 944–52
6. L'Her E, Roy A. Bench tests of simple, handy ventilators for pandemics: performance, autonomy, and ergonomy. *Respir Care* 2011; **56**: 751–60
7. Luján M, Sogo A, Pomares X, Monsó E, Sales B, Blanch L. Effect of leak and breathing pattern on the accuracy of tidal volume estimation by commercial home ventilators: a bench study. *Respir Care* 2013; **58**: 770–7
8. Lyazidi A, Thille AW, Carteaux G, Galia F, Brochard L, Richard J-CM. Bench test evaluation of volume delivered by modern ICU ventilators during volume-controlled ventilation. *Intensive Care Med* 2010; **36**: 2074–80
9. Oscroft NS, Smith IE. A bench test to confirm the core features of volume-assured non-invasive ventilation. *Respirology* 2010; **15**: 361–4
10. Oto J, Chenelle CT, Marchese AD, Kacmarek RM. A comparison of leak compensation in acute care ventilators during noninvasive and invasive ventilation: a lung model study. *Respir Care* 2013; **58**: 2027–37
11. Park MAJ, Freebairn RC, Gomersall CD. Bench performance of ventilators during simulated paediatric ventilation. *Anaesth Intensive Care* 2013; **41**: 349–58
12. Rabec C, Georges M, Kabeya NK, *et al*. Evaluating noninvasive ventilation using a monitoring system coupled to a ventilator: a bench-to-bedside study. *Eur Respir J* 2009; **34**: 902–13
13. Thille AW, Lyazidi A, Richard J-CM, Galia F, Brochard L. A bench study of intensive-care-unit ventilators: new versus old and turbine-based versus compressed gas-based ventilators. *Intensive Care Med* 2009; **35**: 1368–76
14. Vignaux L, Tassaux D, Jolliet P. Performance of noninvasive ventilation modes on ICU ventilators during pressure support: a bench model study. *Intensive Care Med* 2007; **33**: 1444–51
15. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med* 2000; **342**: 1301–8
16. Briel M, Meade M, Mercat A, *et al*. Higher vs lower positive end-expiratory pressure in patients with acute lung injury and acute respiratory distress syndrome: systematic review and meta-analysis. *JAMA* 2010; **303**: 865–73
17. Mercat A, Richard J-CM, Vielle B, *et al*. Positive end-expiratory pressure setting in adults with acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 2008; **299**: 646–55
18. Chiumello D, Pelosi P, Croci M, Bigatello LM, Gattinoni L. The effects of pressurization rate on breathing pattern, work of breathing, gas exchange and patient comfort in pressure support ventilation. *Eur Respir J* 2001; **18**: 107–14
19. Chao DC, Scheinhorn DJ, Stearn-Hassenpflug M. Patient-ventilator trigger asynchrony in prolonged mechanical ventilation. *Chest* 1997; **112**: 1592–9
20. De Wit M, Miller KB, Green DA, Ostman HE, Gennings C, Epstein SK. Ineffective triggering predicts increased duration of mechanical ventilation. *Crit Care Med* 2009; **37**: 2740–5
21. Thille AW, Rodriguez P, Cabello B, Lellouche F, Brochard L. Patient-ventilator asynchrony during assisted mechanical ventilation. *Intensive Care Med* 2006; **32**: 1515–22
22. Vignaux L, Tassaux D, Carteaux G, *et al*. Performance of noninvasive ventilation algorithms on ICU ventilators during pressure support: a clinical study. *Intensive Care Med* 2010; **36**: 2053–9
23. Richard J-C, Carlucci A, Breton L, *et al*. Bench testing of pressure support ventilation with three different generations of ventilators. *Intensive Care Med* 2002; **28**: 1049–57
24. Wallon G, Bonnet A, Guérin C. Delivery of tidal volume from four anaesthesia ventilators during volume-controlled ventilation: a bench study. *Br J Anaesth* 2013; **110**: 1045–51
25. Raux M, Ray P, Prella M, Duguet A, Demoule A, Similowski T. Cerebral cortex activation during experimentally induced

ventilator fighting in normal humans receiving noninvasive mechanical ventilation. *Anesthesiology* 2007; **107**: 746–55
26. Whitelaw WA, Derenne JP, Milic-Emili J. Occlusion pressure as a measure of respiratory center output in conscious man. *Respir Physiol* 1975; **23**: 181–99
27. Schmidt M, Dres M, Raux M, *et al*. Neurally adjusted ventilatory assist improves patient-ventilator interaction during postextubation prophylactic noninvasive ventilation. *Crit Care Med* 2012; **40**: 1738–44
28. Scott DHT, Drummond GB III. Tidal volume measurement: OK for science, but too difficult for a workstation standard? *Br J Anaesth* 2013; **110**: 891–5
29. Battisti A, Tassaux D, Janssens J-P, Michotte J-B, Jaber S, Jolliet P. Performance characteristics of 10 home mechanical ventilators in pressure-support mode: a comparative bench study. *Chest* 2005; **127**: 1784–92
30. Tassaux D, Strasser S, Fonseca S, Dalmas E, Jolliet P. Comparative bench study of triggering, pressurization, and cycling between the home ventilator VPAP II and three ICU ventilators. *Intensive Care Med* 2002; **28**: 1254–61
31. Bunburaphong T, Imanaka H, Nishimura M, Hess D, Kacmarek RM. Performance characteristics of bilevel pressure ventilators: a lung model study. *Chest* 1997; **111**: 1050–60
32. Chatmongkolchart S, Williams P, Hess DR, Kacmarek RM. Evaluation of inspiratory rise time and inspiration termination criteria in new-generation mechanical ventilators: a lung model study. *Respir Care* 2001; **46**: 666–77
33. Costa R, Navalesi P, Spinazzola G, *et al*. Comparative evaluation of different helmets on patient-ventilator interaction during noninvasive ventilation. *Intensive Care Med* 2008; **34**: 1102–8
34. Efrati S, Leonov Y, Oron A, *et al*. Optimization of endotracheal tube cuff filling by continuous upper airway carbon dioxide monitoring. *Anesth Analg* 2005; **101**: 1081–8
35. Vyas D, Inweregbu K, Pittard A. Measurement of tracheal tube cuff pressure in critical care. *Anaesthesia* 2002; **57**: 275–7
36. Ferreira JC, Chipman DW, Hill NS, Kacmarek RM. Bilevel vs ICU ventilators providing noninvasive ventilation: effect of system leaks: a COPD lung model comparison. *Chest* 2009; **136**: 448–56
37. Marchese AD, Chipman D, de la Oliva P, Kacmarek RM. Adult ICU ventilators to provide neonatal ventilation: a lung simulator study. *Intensive Care Med* 2009; **35**: 631–8
38. Terado M, Ichiba S, Nagano O, Ujike Y. Evaluation of pressure support ventilation with seven different ventilators using Active Servo Lung 5000. *Acta Med Okayama* 2008; **62**: 127–33

Handling editor: J. P. Thompson