

An in silico framework for integrating epidemiologic and genetic evidence with health care applications: ventilation-related pneumothorax as a case illustration

RECEIVED 31 August 2015
 REVISED 4 February 2016
 ACCEPTED 9 February 2016
 PUBLISHED ONLINE FIRST 23 April 2016



Yelizaveta Torosyan,¹ Yuzhi Hu,^{1,2} Sarah Hoffman,^{1,3} Qianlai Luo,^{1,3} Bruce Carleton,⁴ and Danica Marinac-Dabic¹

ABSTRACT

Objective To illustrate an in silico integration of epidemiologic and genetic evidence that is being developed at the Center for Devices and Radiological Health/US Food and Drug Administration as part of regulatory research on postmarket device performance. In addition to using conventional epidemiologic evidence from registries, this innovative approach explores the vast potential of open-access omics databases for identifying genetic evidence pertaining to devices.

Material and methods A retrospective analysis of Agency for Healthcare Research and Quality (AHRQ)/Healthcare Cost and Utilization Project (HCUPNet) data (2002–2011) was focused on the ventilation-related iatrogenic pneumothorax (Vent-IP) outcome in discharges with mechanical ventilation (MV) and continuous positive airway pressure (CPAP). The derived epidemiologic evidence was analyzed in conjunction with pre-existing genomic data from Gene Expression Omnibus/National Center for Biotechnology Information and other databases.

Results AHRQ/HCUPNet epidemiologic evidence showed that annual occurrence of Vent-IP did not decrease over a decade. While the Vent-IP risk associated with noninvasive CPAP comprised about 0.5%, the Vent-IP risk due to longer-term MV reached 2%. Along with MV posing an independent risk for Vent-IP, female sex and white race were found to be effect modifiers, resulting in the highest Vent-IP risk among mechanically ventilated white females. The Vent-IP risk was also potentiated by comorbidities associated with spontaneous pneumothorax (SP) and fibrosis. Consistent with the epidemiologic evidence, expression profiling in a number of animal models showed that the expression of several collagens and other SP/fibrosis-related genes was modified by ventilation settings.

Conclusion Integration of complementary genetic evidence into epidemiologic analysis can lead to a cost- and time-efficient discovery of the risk predictors and markers and thus can facilitate more efficient marker-based evaluation of medical product performance.

Keywords: In silico evidence integration using big data approaches, Translational Epidemiology, Repurposing and re-utilization of pre-existing genetic data, Medical device safety biomarkers

INTRODUCTION

As part of a public health commitment to advancing the safety and effectiveness of medical products, the US Food and Drug Administration (FDA) makes continuous efforts in developing new approaches to regulatory science. The Center for Devices and Radiological Health (CDRH) has recently initiated a program on the medical device development tools (eg, in vitro tests, markers, etc.) that are expected to strengthen regulatory decision-making by generating more predictive evidence on medical device performance. The CDRH's regulatory science initiatives include Medical Device Epidemiology Network partnerships, which serve the center's vision for US postmarket surveillance that can accurately characterize real-world performance and thereby deliver "the right device to the right person." New methodologies harnessing recent advances in bioinformatics and genomics can be particularly effective in identifying clinical outcomes and molecular markers that can serve as reliable and measurable study endpoints indicative of real-world performance. These new evidentiary approaches can facilitate more individualized risk-benefit assessment, thus aligning with the recently announced Precision Medicine initiative.

This article illustrates an in silico approach to evidence gathering and analysis that is being developed at the Division of Epidemiology of CDRH as part of the regulatory research efforts aimed at strengthening

postmarket surveillance on medical devices. In addition to employing epidemiologic evidence from conventional sources such as registries, this innovative approach explores pre-existing open-source omics data for identifying molecular markers that have a potential for improving device safety in real-world settings. The current illustration is focused on iatrogenic pneumothorax (IP) as a device-related adverse event with considerable health impact.

Pneumothorax is a potentially fatal complication in patients who require mechanical ventilation (MV) support: the incidence of pneumothorax in patients with acute respiratory distress disorder (ARDS) in the ICU setting can be as high as 50%, with a mortality rate of 66%.¹ Since pneumothorax can have an iatrogenic origin, it is particularly important as a patient safety indicator developed by the Agency for Healthcare Research and Quality (AHRQ). IP observed rate (per 100 000) was shown to be slightly higher in females than males (9.03 and 7.30, respectively) and about 24 times higher in patients 75 and older compared to 18-to-39-year-olds (1.61 and 38.05, respectively).² Although an AHRQ-based study showed a decreasing trend from 1998 to 2007 for the risk-adjusted IP rate,³ the risk-adjusted IP rate based on Veterans Health Administration (VHA) data showed an increase from 2001 to 2004.⁴ A review of pneumothorax cases at one VHA medical center (1983–1988) showed that the incidence of IP may exceed that of spontaneous pneumothorax (SP).⁵

Correspondence to Yelizaveta (Lisa) Torosyan, MD, PhD, General Health Scientist, DEPI/OSB/CDRH/FDA, White Oak, Bldg 66, Rm 4102, Silver Spring, MD, USA; yelizaveta.torosyan@fda.hhs.gov; Tel: 301-796-7127 For numbered affiliations see end of article.

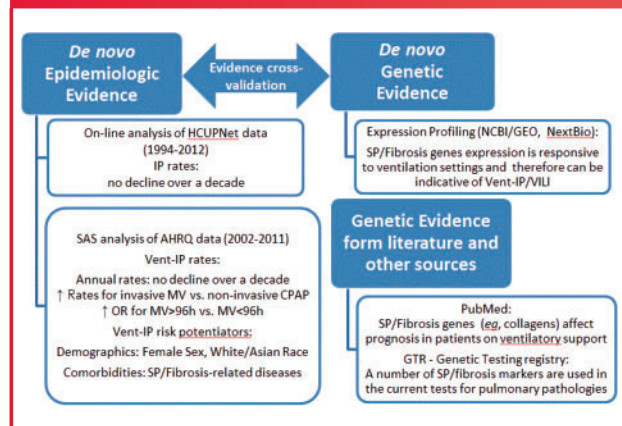
© The Author 2016. Published by Oxford University Press on behalf of the American Medical Informatics Association. All rights reserved. For Permissions, please email: journals.permissions@oup.com.

About 21% of IP cases from the Nationwide Inpatient Sample (NIS) were attributed to respiratory intubation and MV,⁶ pointing at the impact of ventilation-related IP (Vent-IP) on health outcomes around the United States. Vent-IP impact on patient safety was also shown outside of the United States. In a French study from the 1990s,⁷ IP accounted for 38.2% of 217 ICU cases with pneumothorax, and Vent-IP in particular was mostly caused by special ventilatory techniques involving positive end-expiratory pressure (PEEP) and was associated with less favorable outcomes and higher mortality. In a more recent Turkish study,⁸ analysis of 12 010 invasive procedures identified IP in 164 cases, 9.1% of which were considered as Vent-IP caused by barotrauma. A recent large study from Taiwan⁹ showed a relatively low Vent-IP incidence (0.4%, or 124/31 660); however, in-hospital mortality was higher in Vent-IP patients compared to mechanically ventilated patients without pneumothorax (77.4% versus 13.7%, $P < .001$) or to patients with nonventilation pneumothorax (77.4% versus 29.4%, $P < .001$). In a recent Brazilian study,¹⁰ half of the IP cases (9/20 among 645 patients) identified in the pediatric ICU were characterized as Vent-IP.

Although Vent-IP is a relatively rare event in intubated patients with normal lungs, and most of the patients with Vent-IP have underlying lung disease,¹¹ its prompt recognition and treatment is important for the safety of ventilatory support. Vent-IP is mostly caused by high-pressure air as a result of imbalanced ventilation. Optimal settings for PEEP and tidal volume are paramount for avoiding volutrauma and atelectotrauma. While MV was the leading cause of IP in the 1970s, subsequent implementation of new lung-protective ventilation strategies utilizing low tidal volumes and PEEP essentially reduced the Vent-IP occurrence.¹² However, a recent Kaiser Permanente CREST Network study,¹³ which analyzed 1249 patient encounters in US community emergency departments, showed a 16-fold increase (0.1–1.6%, $P < .01$) in IP incidence when positive pressure ventilation was applied. According to a US study on the effectiveness of ventilation modes in real-world settings,¹⁴ use of noninvasive ventilation versus invasive MV can greatly lower Vent-IP risk (0.05% versus 0.5%, $P < .001$); however, the varying rates of noninvasive ventilation among US hospitals impede potential benefits.

Thus, Vent-IP is a device-related adverse event that can still have considerable health impact, despite the implementation of lung-protective ventilation strategies. As a result, we sought to conduct a pilot on the development of a new in silico framework for gathering and analyzing evidence which would (1) allow more comprehensive evaluation of Vent-IP, and (2) outline possibilities for new preventive measures that could refine the risk-benefit assessment of ventilatory strategies. Our search strategy for Vent-IP-related evidence was based on a hypothesis that the factors that are known to contribute to SP may also contribute to Vent-IP. Subsequently, SP/fibrosis-related diseases¹⁵ and corresponding genes (as shown in Supplement 1) were used for deriving and integrating both epidemiologic and genetic evidence pertaining to Vent-IP. As shown in Figure 1, which outlines the evidence streams in this study, the epidemiological evidence showed that continuing occurrence of Vent-IP among ventilation-involving NIS/AHRQ discharges was substantial and occurrence of Vent-IP was related to factors with potential genetic background, such as sex/ethnicity and SP/fibrosis comorbidities. Consistent with the study hypothesis, the cross-species genetic evidence from a number of in vivo models demonstrated that SP/fibrosis-related genes are modified by ventilation-related settings and therefore can reflect ventilation-related responses and adverse events. As a result, the combined evidence from both epidemiologic and genetic streams indicated potential risk modifiers and markers that can be used for the prevention or mitigation of Vent-IP and other ventilation-related injuries.

Figure 1: In silico framework on epidemiological and genetic evidence on Vent-IP.



METHODS

Epidemiological study sample, variables, and outcomes

Data collected by the NIS (called the Nationwide Inpatient Sample prior to 2012, and the National Inpatient Sample thereafter) from the AHRQ from 2002–2011 were used to analyze all available records on ventilation-related procedures with regard to IP. The IP population was first classified by using a discharge diagnosis of IP, as defined by the International Classification of Diseases, 9th Revision (ICD-9) code 512.1. Since this ICD-9 code in isolation is not sufficiently specific for the detection of ventilation-related complications, it was used in conjunction with the ICD-9 codes for ventilatory procedures. Thus, the primary outcome of interest was Vent-IP and the primary exposures of interest were the following ventilatory procedures, used separately or in combination: noninvasive MV or continuous positive air pressure (CPAP) (ICD-9 93.90), continuous invasive MV for less than 96 consecutive hours (MV < 96 h) (ICD-9 96.71), and continuous invasive MV for 96 consecutive hours or more (MV > 96 h) (ICD-9 96.72).

The main demographics-related covariates of interest were sex (male, female) and race/ethnicity (White, Black, Hispanic, Asian, Native American, or other). As a result of the stated hypothesis, the main comorbidity-related covariates of interest were postinflammatory and idiopathic forms of pulmonary fibrosis (ICD-9 515 and 516.31, respectively) as well as genetic diseases associated with SP: cystic fibrosis (ICD-9 277.02, 277.09), Ehlers-Danlos syndrome (ICD-9 756.83), Marfan syndrome (ICD-9 759.82), homocystinuria (ICD-9 270.4), alpha-1-antitrypsin deficiency (ICD-9 273.4), and Birt-Hogg-Dube syndrome (ICD-9 759.89). In order to better characterize the study sample and control for potential confounders, other available patient-related variables (eg, age, obesity, and obesity-related comorbidities) were used (see Supplement 2 for more details).

Statistical data analysis

SAS 9.3 (SAS Institute, Cary, NC, USA) was used to determine the risk distribution of IP by ventilation type, sex, and race/ethnicity. Unweighted frequency tables were used to obtain the prevalence of IP among each subgroup; contingency tables were used to obtain corresponding pairwise odds ratios. Multivariable logistic regression was used to determine the association between IP and ventilation type, with or without the covariates of interest in the unweighted data.

Gene expression profiling

Gene Expression Omnibus (GEO), an open-access genomics data repository from the National Center for Biotechnology Information (NCBI), was used as the main source for in silico discovery of potential

molecular markers for Vent-IP and other ventilation-associated adverse events. The GEO Datasets and Profiles were queried using search terms with SP/fibrosis-related gene names (eg, serpinA1, collagen, etc.) and ventilation-related keywords (eg, ventilat*). After conducting queries and identifying a number (x5) of animal studies mimicking the effects of ventilation, GEO tools such as Cluster Heatmaps, Find Genes, Profile Charts, Profile Neighbors, and T-test Sample Comparison were used for the expression analysis of curated datasets. The GEO2R was additionally used for identifying the lists of differentially expressed genes for further analysis. Some of the noncurated GEO datasets (with no available precomputed gene profile characteristics) were downloaded and analyzed using open-access GenePattern software from the Broad Institute at Harvard and MIT. In addition to GEO/NCBI, the gene expression profiling involved studies available from NextBio (Illumina). The lists of the genes differentially expressed across experimental conditions were ordered by significance and further analyzed using Ingenuity Pathway Analysis (IPA, Qiagen), including the knowledge base and tools such as Compare and Venn diagram, Core/Tox Analyses, Biomarker Filters, and Species/Tissue Expression Overlays.

RESULTS

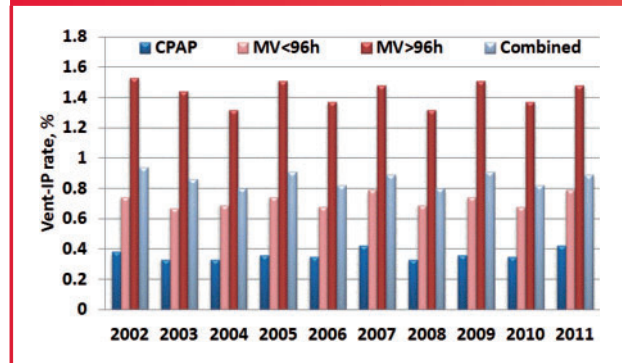
Epidemiologic evidence

Our preliminary analysis of the AHRQ data (1994–2012) using online tools from the Healthcare Cost and Utilization Project (complete results of the HCUPNet-based analysis can be found in Supplements 3 and 4) showed that the IP rate of discharges per 100 000 based on all listed diagnoses varied from 29.2 in 1996 to 18.2 in 2012, mostly fluctuating around the level of 21 per 100 000 since the year 2000. The IP rate based on principal diagnosis was expectedly lower, but remained steady and even slightly increased, reaching levels of 3.7–3.4 per 100 000 in 2009–2012. The in-hospital mortality percentage associated with IP as a principal diagnosis decreased from 2.65% in 1997 to 1.49% in 2005, and fluctuated around 1–1.7% up to 2012.

Our further SAS-based analysis of AHRQ data (2002–2011) focused on the Vent-IP population. On average, Vent-IP was found in 0.88% (22 930/2 596 533) of all discharges involving single or combined ventilation procedures. (The AHRQ-based frequencies for each ventilation type and their combinations as well as the corresponding Vent-IP frequencies can be found in Supplement 5.) However, Vent-IP rates varied substantially depending on ventilation type. Annual Vent-IP occurrence remained expectedly low in noninvasive CPAP, especially compared to longer-term invasive MV > 96 h (Figure 2). The increased odds ratios (OR, 95% CI) for MV-associated Vent-IP were 5.40 (5.11–5.71) and 2.72 (2.57–2.88) for MV > 96 h and MV < 96 h, respectively. Although discharges with a single CPAP or MV procedure comprised the majority of Vent-IP cases, the Vent-IP frequency was higher among discharges with the combined ventilation procedures involving longer-term MV > 96 h (Figure 3).

In addition to the ventilation type, the Vent-IP risk was affected by patient demographics. Vent-IP occurred more frequently in females than in males (0.95% versus 0.82%, respectively) and in Whites than in Hispanics or Blacks (0.99% versus 0.72% and 0.68%, respectively). The effects of ventilation as a primary exposure and the modifier role of demographic factors were further explored using multivariable logistic regression (Figure 4A). Regardless of sex or race/ethnicity, long-lasting invasive ventilation (MV > 96 h) posed the highest risk of Vent-IP when compared to shorter-term MV and noninvasive CPAP. With the amplified effects of sex and race/ethnicity, Vent-IP occurrence reached 2% among White females on MV > 96 h. White females in general had the highest Vent-IP risk, compared to all other subgroups (1.08% versus 0.82%, OR = 1.33, 95% CI, 1.29–1.36, $\chi^2 = 386.41$,

Figure 2: Annual Vent-IP rates (%) among discharges (AHRQ, 2002–2011) with noninvasive ventilation (CPAP), invasive mechanical ventilation (MV < 96 h and MV > 96 h), or combined ventilation procedures (Combined).



$P < .0001$) and especially Hispanic males, who had the lowest Vent-IP risk (1.08% versus 0.66%, OR = 1.66, 95% CI, 1.54–1.78, $\chi^2 = 189.30$, $P < .0001$).

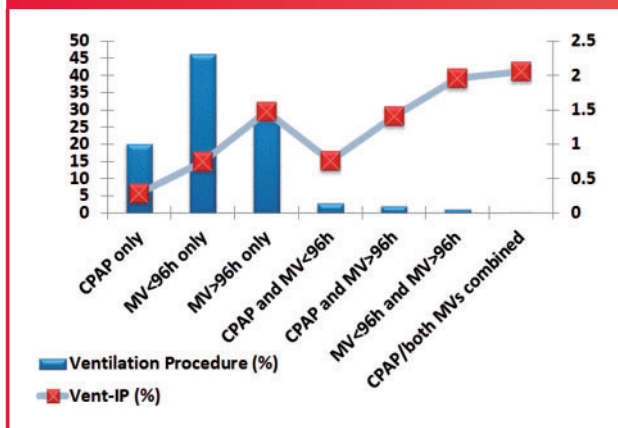
To explore the possible role of comorbidities as additional effect modifiers, the risk of Vent-IP was assessed with regard to pulmonary fibrosis and genetic diseases indicative of SP. (A table showing SP diseases¹⁵ can be found in Supplement 1.) Figure 4B summarizes the main SP/fibrosis comorbidities that essentially increased Vent-IP risk. Even the average risk of Vent-IP among all discharges with SP/fibrosis was 2.5 times higher than that in the entire ventilated population (2.14% and 0.88%, respectively; $P < .0001$). Among a total of 107 discharges listing 1 of the SP diseases, discharges with both Vent-IP and Birt-Hogg-Dube syndrome were most frequent. However, the highest risks for Vent-IP were posed by Marfan syndrome (3.45%, OR = 3.88; $\chi^2 = 45.40$, $P < .0001$) and cystic fibrosis (CF), based on 2 ICD-9 codes (1.85%, OR = 2.11, $\chi^2 = 16.07$, $P < .0001$ and 2.17%, OR = 2.44, $\chi^2 = 4.16$, $P < .0414$).

Pro-fibrotic changes were shown to contribute to ventilation-induced or ventilated-associated lung injury,¹⁶ (VILI/VALI, used here interchangeably to follow the terms used in original studies). Consistent with that, both idiopathic and postinflammatory forms of pulmonary fibrosis (IPF and PPF), which in some cases may have a genetic background,¹⁷ also enhanced the risk of Vent-IP (OR = 4.01, $\chi^2 = 13.08$, $P < .0003$ and OR = 2.57, $\chi^2 = 819.95$, $P < .0001$, respectively).

Although IPF/PPF-Vent-IP, with a total of 912 cases, was more prevalent than SP-Vent-IP, it did not show the White female predominance that was found in the total Vent-IP population. On the other hand, CF-Vent-IP was more frequent among White females, and this predominance was slightly higher among White CF patients with Vent-IP versus their counterparts without Vent-IP (F/M ratios 1.7 and 1.3, respectively; NS). Unlike CF-Vent-IP, Marfan-Vent-IP showed a predominance of White males.

As a result of additional analyses for controlling potential confounders (Supplement 2), MV was confirmed to independently increase the risk of Vent-IP: the odds of having Vent-IP were 3.6 times higher among patients on MV than on CPAP. After adjusting for all potential confounders used, the female sex and the presence of pulmonary fibrosis or SP-related genetic disease was still associated with increased Vent-IP risk, and non-White ethnicities were associated with decreased Vent-IP risk. In addition, age was shown to increase Vent-IP risk. Consistent with the demonstrated Vent-IP effect modification by demographic factors and SP/fibrosis comorbidities, SP in neonates was associated with the White race.¹⁸ SP patients were also shown to

Figure 3: Vent-IP frequencies (%) among discharges (AHRQ, 2002–2011) with single or combined ventilation procedures. The left axis shows percentages for the ventilation procedures and the right axis shows percentages for the corresponding Vent-IP rates.



be taller^{19,20} and slenderer,²¹ supporting the plausibility of race- and sex-associated effects in development of both SP and Vent-IP. Physical characteristics such as weight and height are not available in the NIS database. However, we hypothesized that obese patients could have decreased risk of IP. Obesity and some obesity-related comorbidities were added to the covariates of interest and, as expected, their presence was associated with decreased Vent-IP risk (Supplement 2).

Thus, the AHRQ-based epidemiologic evidence showed that Vent-IP is an adverse event with continuous impact on the safety of ventilatory support, and pointed to White females as a subpopulation most susceptible to development of Vent-IP. Taken together, Vent-IP risk modification by race, sex and body build related factors, as well as by SP/fibrosis-related diseases with a genetic background, suggests a possible role of genetically determined factors in the development of Vent-IP.

Genetic evidence

Next, using a GEO/GenePattern/IPA-based analytic algorithm and NCBI/GEO data from a number of animal models mimicking in vivo ventilation effects, we showed that the expression profiles of some SP-associated genes (eg, collagens) were modified by ventilation settings of tidal volume and pressure.

Multiple collagen family members (COL1A1, COL1A2, COL3A1, COL4A1, COL4A2, COL5A2, and COL15A1) were found among the genes significantly ($P < .001$) upregulated by high tidal volume ventilation in VALI-resistant versus VALI-sensitive rats. Consistent with that, several collagens (COL1A1, COL1A2, COL3A1, COL5A1, and COL5A2) were significantly ($P < .01$) downregulated in VILI-rats versus control rats. Based on these cross-validating results from 2 independent studies (GEO: GDS2709 and GDS1241), COL1A1, COL1A2, COL3A1, and COL5A2 were identified as the most plausible VILI-marker candidates among the collagen family. Remarkably, 3 of these collagens, COL1A1, COL1A2, and COL3A1, are involved in SP-associated Ehlers-Danlos and atypical Marfan syndromes. In addition to a similar downregulation of these 3 collagens, VILI was characterized by upregulation of another SP-gene, SerpinA1 (Figure 5A). A low baseline expression of the 3 SP-associated collagens in VALI-sensitive rats was further reduced by high tidal volume ventilation. Consistent with that, a higher baseline expression of the collagens in VALI-resistant rats was further increased by high tidal volume ventilation (Figure 5B). In addition to the collagens, high tidal volume ventilation modified the expression of

SerpinA1 (alpha-antitrypsin deficiency), Fbn1 (Marfan syndrome), and Flcn (Birt-Hogg-Dubé syndrome).

Among other ventilation-related alterations (Figure 5C), SerpinA1 (which was induced in VILI- and concordantly reduced in VALI-sensitive rats) was shown to be upregulated in rats ventilated with 100% oxygen. In ventilated mice, Flcn was slightly increased by MV and decreased by MV + LPS (a treatment simulating ventilation-associated pneumonia), and the causative gene for CF, CFTR, was induced by MV and further upregulated by MV + LPS.

Additional NextBio-based profiling (data shown in Supplement 6) on Flcn (the causative gene for Birt-Hogg-Dubé syndrome) showed that human Flcn was increased in patients with IPF, and murine Flcn was reduced in a model of hyperoxia-induced bronchopulmonary dysplasia, also known as “ventilator lung in newborn” (<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0022889/>; accessed on March 15, 2015).

The GEO/NCBI-based expression profiling on IPF patients (data shown in Supplements 6 and 7) showed an upregulation trend for FBN1 and multiple collagens (COL6A1, COL6A2, COL6A3, COL12A1, COL14A1, and COL16A1), including COL15A1 (Figure 5D), which was upregulated by the high tidal volume in VALI-resistant rats (data not shown). The collagen alterations in IPF included a slight upregulation of antiangiogenic COL18A1 (data not shown), whose upregulation was also shown in postmortem lung samples from ventilated preterm infants and was considered part of the angiogenic shift and deficient alveolarization in infants with bronchopulmonary dysplasia.²² Further confirming the ventilation-related collagen alterations, elevated Col1A1 was shown to be part of the pro-fibrogenic profile in murine VILI,²³ and elevated ratios of type 1/3 collagens were shown in the postmortem lung samples from ventilated neonates with respiratory distress.²⁴ Fibrotic alterations in general are indicative of collagen synthesis activation and collagenase inhibition, which are considered key events in VILI-related lung remodeling.¹⁶

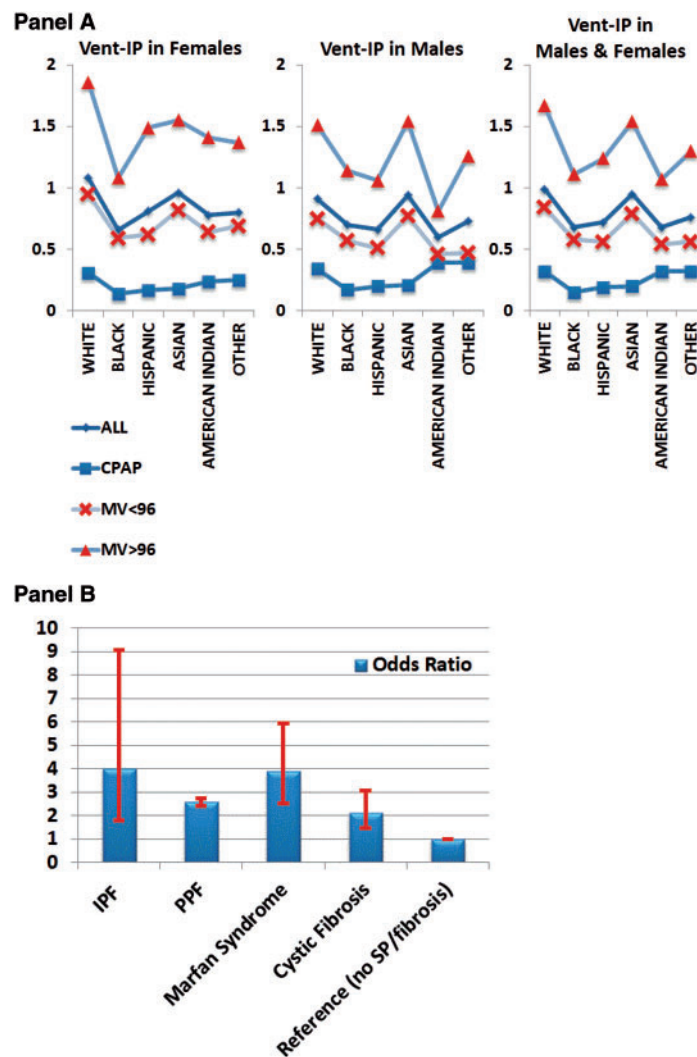
To explore the biological relevance of the identified marker candidates and their potential utility in a clinical setting, we also examined their links to lung pathophysiology and assessed their detectability in human biofluids. The IPA network of SP/fibrosis-genes was associated with respiratory pathways and biofunctions such as the NRF2-mediated oxidative stress response (FLCN) and the morphology (including abnormal) of respiratory system and lung development (FBN1, CFTR, and COL3A1). According to IPA biofluid expression overlay, most of the collagens as well as FBN1, SERPINA1, and CFTR can be detected in blood, plasma/serum, urine, and/or sputum.

Thus, the genetic evidence compiled from a number of independent (human and animal) in vivo models showed that numerous SP/fibrosis-genes (FBN1, FLCN, SERPINA1, and collagens such as COL1A1, COL1A2, and COL3A1) are responsive to ventilation settings (tidal volume and pressure) and therefore can be indicative of ventilation-related lung injuries.

DISCUSSION

This pilot encompasses a translational continuum that is aimed at moving discoveries on Vent-IP risk factors and markers from the starting phase of in silico research to the end phase of new ventilation safety-related health applications. The following discussion provides evidence-based assessment of the projected scope of Vent-IP health impact and the value of Vent-IP marker/predictor-based applications for health practices. As a necessary intermediate step in the translational continuum, this article attempts to disseminate current research results for further development and implementation of evidence-based guidelines on Vent-IP management. This article also outlines next steps (eg, implementation of a new Vent-IP-specific ICD code) that are needed for assessing the real-world impact of Vent-IP and the

Figure 4: Patient demographics and comorbidities as Vent-IP effect modifiers. (A) Vent-IP occurrence (%) among discharges stratified by ventilation type, sex, and race/ethnicity (AHRQ, 2002–2011). (B) Vent-IP occurrence (OR, with 95% CI as red bars) among discharges with SP/fibrosis comorbidities (AHRQ, 2002–2011).



expected mitigating effects from clinical implementation of Vent-IP risk predictors and markers. Of particular use here would be active surveillance methodology to carefully characterize Vent-IP cases and matched patients with ventilator use without IP, which would allow validation of the current study findings and evaluation of real-world outcomes in patients stratified by putative Vent-IP risk predictors and markers.

Since the NIS represents only a 20% stratified sample of US community hospitals (sample size can be verified at <http://www.hcup-us.ahrq.gov/nisoverview.jsp>; accessed on March 13, 2015), the overall Vent-IP frequency found in this study, 22 930 cases in 10 years (more details in Supplement 5) is representative of approximately one-fifth of the all ventilation-related discharges. Without formal weighting, the number of in-hospital Vent-IP cases around the United States can be roughly estimated at 11 500 per year. Meanwhile, the total incidence of Vent-IP may be even higher due to possible cases resulting from home use of ventilatory devices. It should also be taken into account that baseline Vent-IP rates can be further exacerbated under circumstances requiring mass ventilatory support, eg, respiratory epidemics. The

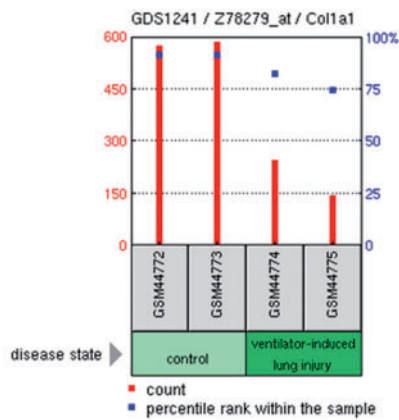
sizeable scope of this projected Vent-IP health impact warrants implementation of a new ICD code that would be specific to ventilation-associated IP (similar to ventilator-associated pneumonia, ICD-9 997.31 and ICD-10 J95.851) and that would allow more effective surveillance of Vent-IP compared to the currently used ICD-9 code (512.1), which incorporates all IP cases regardless of the causative procedure.

In addition to the impact of Vent-IP per se, a comprehensive risk assessment of ventilation-related adverse events should include other possible lung injuries. Positive pressure can contribute to the development of Vent-IP in 2 ways.¹³ First, positive pressure barotrauma can directly cause Vent-IP (this occurs mostly in patients with an underlying chronic pulmonary disease). Second, positive pressure ventilation can enlarge the size of existing pneumothoraces, or even transform some occult abnormalities into a Vent-IP. In any event, Vent-IP occurrence carries a greater risk of serious, even life-threatening barotrauma, indicating the risk for additional adverse outcomes such as ventilator-associated pneumonia or VILI. Thus, Vent-IP risk predictors and markers could have larger implications for ventilation-related adverse outcomes in general.

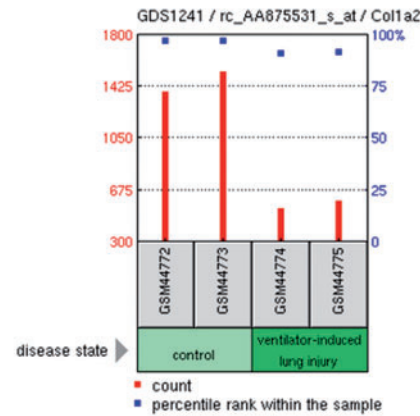
Figure 5: Ventilation-related expression profiling of genes associated with spontaneous pneumothorax and pulmonary fibrosis. **A)** Gene expression of collagen family members (Col1A1, Col1A2, and Col3A1) and SerpinA1 in lung tissue samples from ventilation-induced/associated lung injury model (GEO:GDS1241, rat). **B)** Gene expression of collagen family members (Col1A1, Col1A2, and Col3A1), SerpinA1, Fbn1, and Flcn in lung tissue samples from ventilation-induced/associated lung injury resistant versus sensitive rat strains (GEO:GDS2709). **(C)** Gene expression of SerpinA1, Flcn, and CFTR in lung tissue samples from other ventilation-related animal models (GEO:GDS64, top; GDS:1276, middle and bottom). **D)** Gene expression of FBN1 and collagen COL15A1 in lung biopsy samples from patients with idiopathic pulmonary fibrosis versus controls (GEO:GDS1252).

Panel A

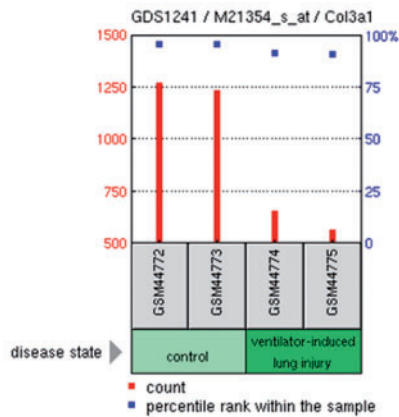
Profile GDS1241 / Z78279_at / Col1a1
Title Ventilator-induced lung injury (rat)
Organism Rattus norvegicus



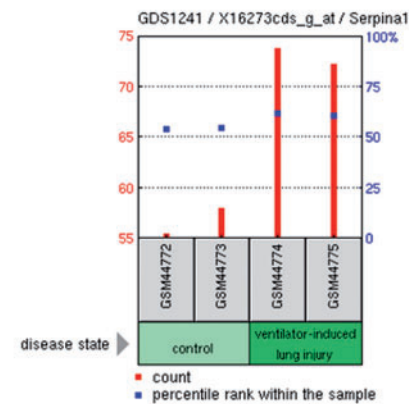
Profile GDS1241 / rc_AA875531_s_at / Col1a2
Title Ventilator-induced lung injury (rat)
Organism Rattus norvegicus



Profile GDS1241 / M21354_s_at / Col3a1
Title Ventilator-induced lung injury (rat)
Organism Rattus norvegicus



Profile GDS1241 / X16273cds_g_at / Serpina1
Title Ventilator-induced lung injury (rat)
Organism Rattus norvegicus



The collagen markers in particular can be applicable to preterm infants, who, when exposed to oxygen and ventilation, are at risk for developing ventilation-related complications such as bronchopulmonary dysplasia, with subsequent chronic lung disease due to arrested alveolar development and dysmorphic angiogenesis.²² In a study on patients with ARDS, evidence of collagen biosynthesis with altered collagen expression was shown to be part of the early-ARDS proteomic profile that differentiated survivors from nonsurvivors.²⁵ Despite the implementation of lung-protective ventilation strategies, ARDS, which is usually indistinguishable from VILI/VALI,²⁶ continues to be a major health care problem, affecting more than 190 000 people in the

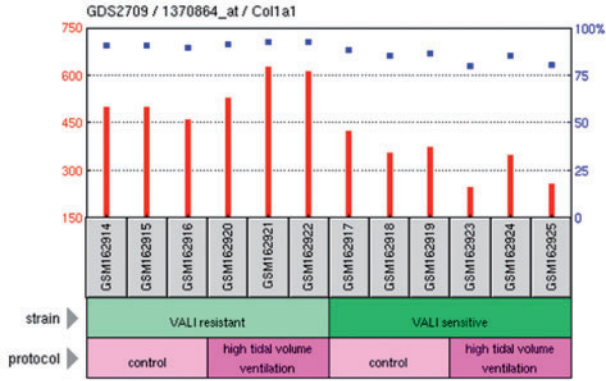
United States annually and causing a mortality rate of 27–45%.²⁷ Since severe fibroproliferative lung disease has been associated with a poor prognosis, with high mortality or prolonged ventilator dependence, Vent-IP/VILI markers enabling the mechanistic understanding and monitoring of ventilation-caused fibroproliferation can be particularly useful in developing preventive/attenuating tactics aimed at improving ventilation-related health outcomes.

Knowing the exact health care impact of ventilation-related complications can help with making more informed clinical and regulatory decisions pertaining to ventilatory support. Implementation of validated Vent-IP risk predictors can individualize risk-benefit assessment (eg,

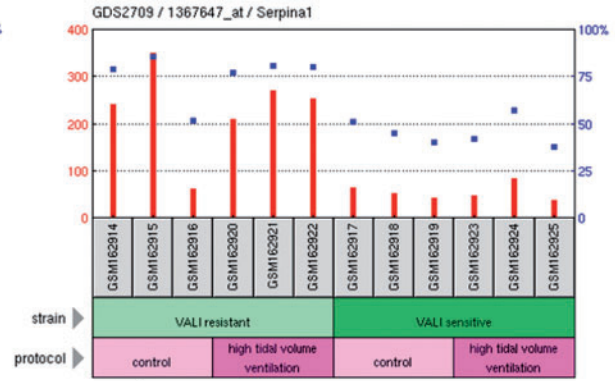
Figure 5: Continued

Panel B

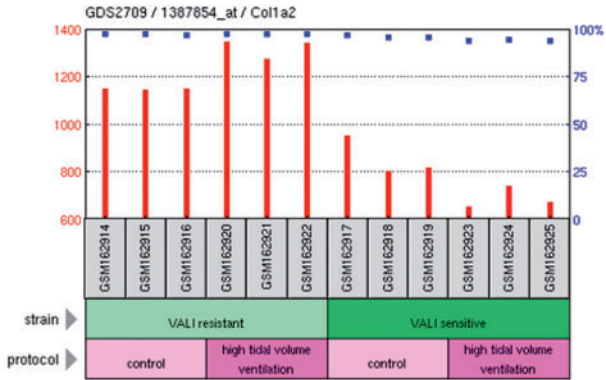
Profile GDS2709 / 1370864_at / Col1a1
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



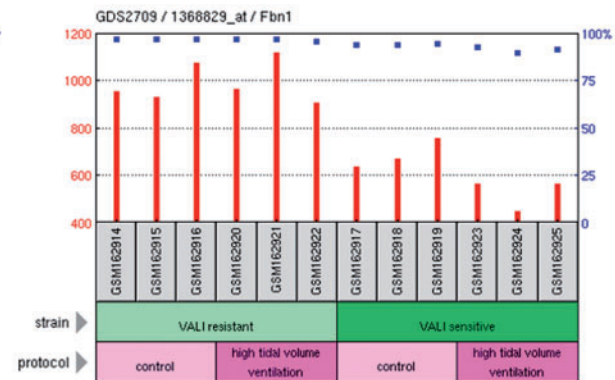
Profile GDS2709 / 1367647_at / Serpina1
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



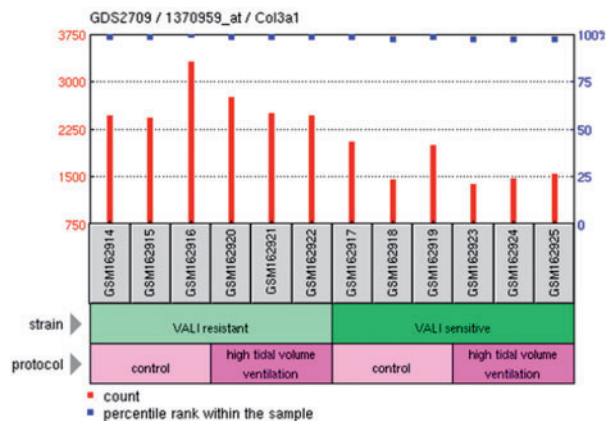
Profile GDS2709 / 1387854_at / Col1a2
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



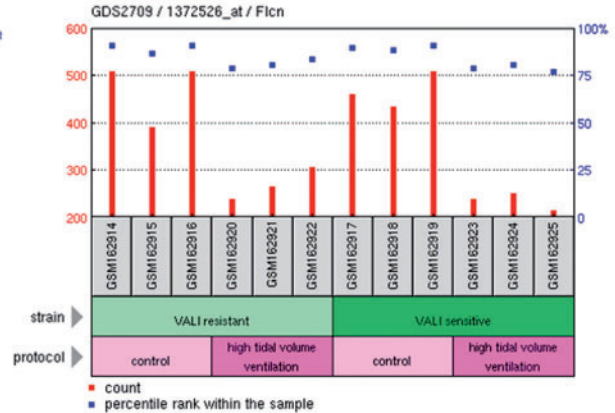
Profile GDS2709 / 1368829_at / Fbn1
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



Profile GDS2709 / 1370959_at / Col3a1
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



Profile GDS2709 / 1372526_at / Flcn
Title Ventilator-associated lung induced injury resistant and sensitive strains
Organism Rattus norvegicus



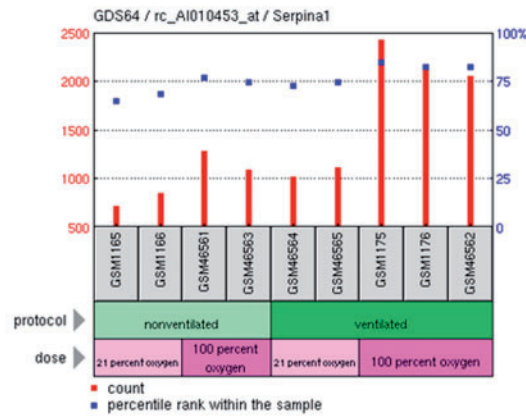
■ count
 ■ percentile rank within the sample

■ count
 ■ percentile rank within the sample

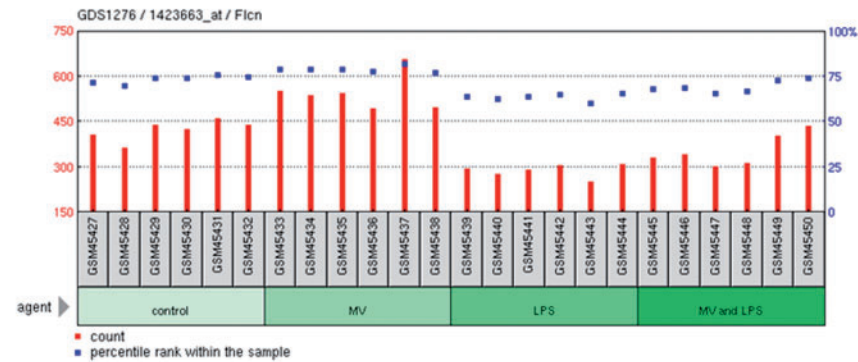
Figure 5: Continued

Panel C

Profile GDS64 / rc_AI010453_at / Serpina1
Title Lung ventilation
Organism Rattus norvegicus



Profile GDS1276 / 1423663_at / Flcn
Title Inflammatory lung injury and mechanical ventilation
Organism Mus musculus



Profile GDS1276 / 1420579_s_at / Cfr
Title Inflammatory lung injury and mechanical ventilation
Organism Mus musculus

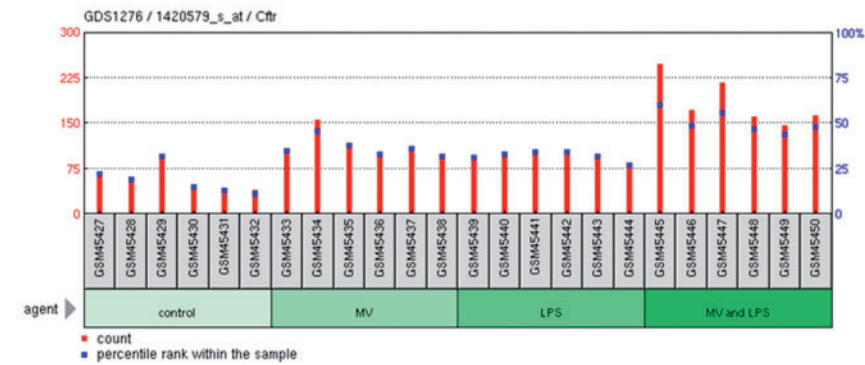
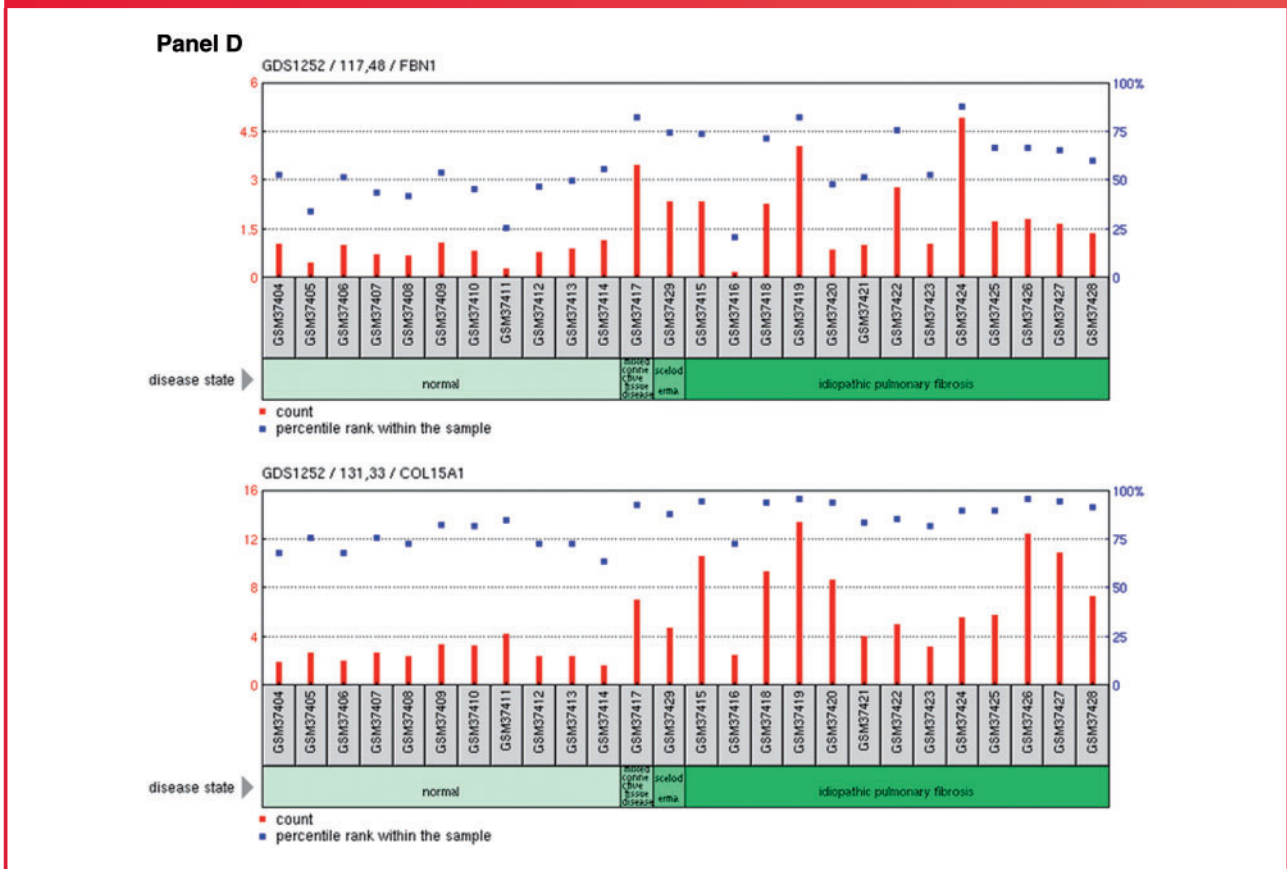


Figure 5: Continued



identifying high-risk patients who are less tolerant to certain ventilation tactics) and can shape the manner in which ventilatory procedures are explained to patients (eg, as part of the informed consent process). Many cases of Vent-IP are preventable, and confirmation of the discriminatory power of Vent-IP risk predictors/markers in real-world settings can be a first step in minimizing Vent-IP incidence. As can be found in the Genetic Testing registry (<http://www.ncbi.nlm.nih.gov/gtr/tests/?term=collagen>; accessed on March 13, 2015), a number of the currently used diagnostic/prognostic tests incorporate collagens and other potential Vent-IP markers and therefore can be leveraged for development of clinically applicable Vent-IP/VILI tests. Implementation of putative Vent-IP/VILI markers that are detectable by routine noninvasive techniques can represent a next benchmark for enhancing the safety of ventilatory support since the implementation of lung-protective ventilatory strategies.

Although our analysis showed a higher association of Vent-IP with pulmonary fibrosis than with SPs, the carrier frequency for CF alone (which, unlike many SP-diseases, is an autosomal-recessive disease) can confer a significant population risk for Vent-IP. About 1 in every 20 Americans is an asymptomatic carrier of the abnormal CFTR gene (http://www.cdc.gov/excite/ScienceAmbassador/ambassador_pgm/lessonplans/high_school/Am%20a%20Carrier%20for%20Cystic%20Fibrosis/Cystic_Fibrosis_Fact_Sheet.pdf). Concordant to the sex/ethnicity trends identified for Vent-IP, CF is more prevalent among Caucasians, and it results in more severe lung disease in females. As a confirmation of possible functional and clinical effects due to CFTR heterozygosity, delta-F508 carriers were shown to have increased risks for diabetes²⁸ and asthma.²⁹ Moreover, PEEP-volume curves measured in Cfr+/- mice showed the presence of a gene dosage

effect that results in greater lung compliance (lower elastance) in Cfr-heterozygotes compared to knockout and wild-type.³⁰ These data strongly suggest that CF carriers may represent an essential subpopulation that, when placed on ventilatory support, may be highly susceptible to development of Vent-IP/VILI and therefore may benefit from genetic testing using Vent-IP/VILI markers. While SP/fibrosis gene expression markers can be more applicable for monitoring the development and resolution of Vent-IP/VILI in patients who are already on ventilatory support, SP/fibrosis gene polymorphisms/mutations can be used for more individualized risk-benefit assessment and preventive action prior to ventilation procedures.

CONCLUSIONS

As a pilot on new evidentiary approaches suitable for health care policy consideration, this study was focused on ventilatory safety as a subject applicable to real-world health care decisions and comprehensible to the wider medical community, including clinicians, researchers, and regulators. However, this proof-of-concept study demonstrates a wider applicability of this in silico approach, which is based on the reutilization of pre-existing data and which thereby paves the way for cost- and time-effective pharmacoepidemiologic and pharmacogenetic solutions that are needed for implementing Precision Medicine initiatives and enhancing medical product safety.

STUDY LIMITATIONS

The current study was limited to Vent-IP as a single ventilation-related event identifiable by a combination of ICD-9 codes. The newly implemented ICD-9 codes for ventilation-associated pneumonia and “ventilation” pneumonitis

(997.31 and 495.7, respectively) did not allow for gathering sufficient data; VILI/VALI was precluded from the study due to a complete lack of ICD-9 codes. The NIS analysis was unweighted.

FUNDING

This work was supported in part by an appointment to the Research Participation Program at the Center for Devices and Radiological Health Office of Surveillance and Biometrics, the US Food and Drug Administration, administered by the Oak Ridge Institute for Science and Education through an inter-agency agreement between the US Department of Energy and the FDA.

COMPETING INTERESTS

The authors have no competing interests to declare.

CONTRIBUTORS

YT: Critical contribution to the conception and in silico design of the work as well as to the data analysis and interpretation.

YH: Substantial contribution to the analysis and interpretation of data.

SH: Substantial contribution to the analysis of data.

QL: Substantial contribution to the analysis of data.

BC: Revision of the manuscript for important content.

DMD: Revision and final approval of the version to be published.

REFERENCES

- Terzi E, Zarogoulidis K, Kougioumtzi I, et al. Acute respiratory distress syndrome and pneumothorax. *J Thorac Dis*. 2014;6(Suppl 4):S435–S442.
- Iatrogenic pneumothorax (provider-level): rate per 1,000 discharges. <http://www.qualitymeasures.ahrq.gov/content.aspx?id=38515&search=pneumothorax>. Accessed February 28, 2015. Source: AHRQ QI. Patient safety indicators #6: technical specifications. *Iatrogenic pneumothorax rate [version 4.4]*. Rockville, MD: Agency for Healthcare Research and Quality; 2012.
- Downey JR, Hernandez-Boussard T, Banka G, Morton JM. Is patient safety improving? National trends in patient safety indicators: 1998–2007. *Health Serv Res*. 2012;47(1 Pt 2):414–430.
- Rosen AK, Zhao S, Rivard P, et al. Tracking rates of patient safety indicators over time: lessons from the Veterans Administration. *Med Care*. 2006;44(9):850–861.
- Despars JA, Sassoon CS, Light RW. Significance of iatrogenic pneumothoraces. *Chest*. 1994;105(4):1147–1150.
- Patient Safety Indicator Comparative Data: Based on the 2009 Nationwide Inpatient Sample (NIS), Version 4.4 (August 2012). http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V44/Comparative_Data_PSI_4.4.pdf. Accessed February 28, 2014.
- Lissac J. [Spontaneous and iatrogenic pneumothorax in the adult. 217 personal cases]. *Bull Acad Natl Med*. 1994;178(2):213–223; discussion 223–225. [Article in French]
- Celik B, Sahin E, Nadir A, Kaptanoglu M. Iatrogenic pneumothorax: etiology, incidence and risk factors. *Thorac Cardiovasc Surg*. 2009;57(5):286–290.
- Hsu CW, Sun SF. Iatrogenic pneumothorax related to mechanical ventilation. *World J Crit Care Med*. 2014;3(1):8–14.
- da Silva PS, de Aguiar VE, Fonseca MC. Iatrogenic pneumothorax in mechanically ventilated children: incidence, risk factors and other outcomes. *Heart Lung*. 2015;44(3):238–242.
- Hsu CW, Sun SF, Lee DL, Chu KA, Lin HS. Clinical characteristics, hospital outcome and prognostic factors of patients with ventilator-related pneumothorax. *Minerva Anestesiol*. 2014;80(1):29–38.
- Baumann MH. Pneumothorax. *Semin Respir Crit Care Med*. 2001;22(6):647–656.
- Vinson DR, Ballard DW, Hance LG, et al., Kaiser Permanente CREST Network Investigators. Pneumothorax is a rare complication of thoracic central venous catheterization in community EDs. *Am J Emerg Med*. 2015;33(1):60–66.
- Tsai CL, Lee WY, Delclos GL, Hanania NA, Camargo CA Jr. Comparative effectiveness of noninvasive ventilation vs invasive mechanical ventilation in chronic obstructive pulmonary disease patients with acute respiratory failure. *J Hosp Med*. 2013;8(4):165–172.
- Chiu HT, Garcia CK. Familial spontaneous pneumothorax. *Curr Opin Pulm Med*. 2006;12(4):268–272.
- Cabrera-Benitez NE, Laffey JG, Parotto M, et al. Mechanical ventilation-associated lung fibrosis in acute respiratory distress syndrome: a significant contributor to poor outcome. *Anesthesiology*. 2014;121(1):189–198.
- Ley B, Collard HR. Epidemiology of idiopathic pulmonary fibrosis. *Clin Epidemiol*. 2013;5:483–492.
- Brenner JS, Karlowicz MG. Nonfatal symptomatic spontaneous pneumothorax in neonates: association with white ethnicity and lack of association with major urinary tract malformations. *Clin Pediatr (Phila)*. 1997;36(4):241–243.
- Sadikot RT, Greene T, Meadows K, Arnold AG. Recurrence of primary spontaneous pneumothorax. *Thorax*. 1997;52(9):805–809.
- Guo Y, Xie C, Rodriguez RM, Light RW. Factors related to recurrence of spontaneous pneumothorax. *Respirology*. 2005;10(3):378–384.
- Kawakami Y, Irie T, Kamishima K. Stature, lung height, and spontaneous pneumothorax. *Respiration*. 1982;43:35–40.
- De Paepe ME, Greco D, Mao Q. Angiogenesis-related gene expression profiling in ventilated preterm human lungs. *Exp Lung Res*. 2010;36(7):399–410.
- Li LF, Liu YY, Kao KC, et al. Mechanical ventilation augments bleomycin-induced epithelial-mesenchymal transition through the Src pathway. *Lab Invest*. 2014;94(9):1017–1029.
- Shoemaker CT, Reiser KM, Goetzman BW, Last JA. Elevated ratios of type I/III collagen in the lungs of chronically ventilated neonates with respiratory distress. *Pediatr Res*. 1984;18(11):1176–1180.
- Bhargava M, Becker TL, Viken KJ, et al. Proteomic profiles in acute respiratory distress syndrome differentiates survivors from non-survivors. *PLoS One*. 2014;9(10):e109713.
- International consensus conferences in intensive care medicine. Ventilator-associated lung injury in ARDS. American Thoracic Society, European Society of Intensive Care Medicine, Société de Réanimation Langue Française. [No authors listed] *Intensive Care Med*. 1999;25(12):1444–1452. Review.
- Burnham EL, Janssen WJ, Riches DW, Moss M, Downey GP. The fibroproliferative response in acute respiratory distress syndrome: mechanisms and clinical significance. *Eur Respir J*. 2014;43(1):276–285.
- Preumont V, Hermans MP, Lebecque P, Buysschaert M. Glucose homeostasis and genotype-phenotype interplay in cystic fibrosis patients with CFTR gene deltaF508 mutation. *Diabetes Care*. 2007;30(5):1187–1192.
- Dahl M, Tybjaerg-Hansen A, Lange P, Nordestgaard BG. Asthma and COPD in cystic fibrosis intron-8 5T carriers. A population-based study. *Respir Res*. 2005;6:113.
- Cohen JC, Lundblad LK, Bates JH, Levitzky M, Larson JE. The “Goldilocks effect” in cystic fibrosis: identification of a lung phenotype in the CFTR knockout and heterozygous mouse. *BMC Genet*. 2004;27(5):21.

AUTHOR AFFILIATIONS

¹Division of Epidemiology, Center for Devices and Radiological Health, CDRH/FDA, Silver Spring, MD, USA

²Columbia University Mailman School of Public Health, New York, NY, USA

³Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁴Pharmaceutical Outcomes Programme, BC Children’s Hospital; Division of Translational Therapeutics, Department of Pediatrics, Child and Family Research Institute, University of British Columbia, Vancouver, Canada