Numerous compounds have shown efficacy in limiting development of pulmonary fibrosis using animal models, yet few of these compounds have replicated these beneficial effects in clinical trials. Given the challenges associated with performing clinical trials in patients with idiopathic pulmonary fibrosis (IPF), it is imperative that preclinical data packages be robust in their analyses and interpretations to have the best chance of selecting promising drug candidates to advance to clinical trials. The American Thoracic Society has convened a group of experts in lung fibrosis to discuss and formalize recommendations for preclinical assessment of anti-fibrotic compounds. The panel considered three major themes (choice of animal, practical considerations of fibrosis modeling, and fibrotic endpoints for evaluation). Recognizing the need for practical considerations, we have taken a pragmatic approach. The consensus view is that use of the murine intratracheal bleomycin model in animals of both genders, using hydroxyproline measurements for collagen accumulation along with histologic assessments, is the best-characterized animal model available for preclinical testing. Testing of antifibrotic compounds in this model is recommended to occur after the acute inflammatory phase has subsided (generally after Day 7). Robust analyses may also include confirmatory studies in human IPF specimens and validation of results in a second system using in vivo or in vitro approaches. The panel also strongly encourages the publication of negative results to inform the lung fibrosis community. These recommendations are for preclinical therapeutic evaluation only and are not intended to dissuade development of emerging technologies to better understand IPF pathogenesis.
Many compounds show efficacy in limiting fibroblast/myofibroblast activation in vitro, or beneficial effects in animal models of lung fibrosis. However, very few showing efficacy in animal models have translated into successful therapies for idiopathic pulmonary fibrosis (IPF). Possible explanations may relate to limitations of animal models and/or flaws in trial design. There is urgent need to increase the potential for successful translation of preclinical models to improved patient care, because unsuccessful clinical trials waste valuable patient and financial resources, undermine confidence, and, ultimately, reduce the chance for future successful trials. Thus, preclinical data packages that robustly predict successful clinical trials are needed to help trialists and pharmaceutical companies decide which drug(s) to develop therapeutically.

IPF is a complex, heterogeneous, and progressive disease of unknown etiology. Animal models don’t fully recapitulate physiologic findings of IPF (1, 2) or histopathologic pattern of usual interstitial pneumonia. However, they do enable mechanistic investigations relevant to fibrogenesis. Our goal was to define the minimum standard practice guidelines for preclinical therapeutic assessment, with a goal of enhancing preclinical data assessment going forward. Overall, committee participants believe that judicious use of bleomycin, and other models, can be effective in determining the utility of potential new therapies.

Materials and Methods

To identify standards for use of animal models in preclinical drug design, the American Thoracic Society convened an expert panel to define optimal experimental protocols to ensure that in vivo animal modeling studies have the highest chance of discriminating between potentially effective and ineffective antifibrotic compounds. U.S. and international experts on animal models of lung fibrosis participated. Members of the writing committee submitted conflict of interest statements before the workshop. No important conflicts were identified or became apparent during the workshop.

The panel considered three major themes (choice of animal, practical considerations of fibrosis modeling, and fibrotic endpoints for evaluation) as outlined below. After viewing expert presentations, participants discussed key questions and needs. Participants were encouraged to express opinions and recommendations. Additional recommendations were formulated during teleconferences among writing committee members after the workshop. Disagreement was resolved by discussion and consensus. All workshop attendees reviewed and revised the manuscript before submission.

Recommendations were also informed by the Animal Research: Reporting of In Vivo Experiments guidelines (online at https://www.nc3rs.org.uk/arrive-guidelines [3]) and the In Vivo Experiments guidelines (online at https://www.nc3rs.org.uk/arrive-guidelines [3]) with the aim of minimizing animal experimentation while improving reproducibility and repeatability within scientific research (4, 5).

Animal Use in Fibrosis Models

Species Considerations

A single-model system may never fully recapitulate all aspects of human IPF biology. Prominent IPF features include its progressive and irreversible nature and sex predilection for older males. Similarly, murine models don’t fully recapitulate classical IPF histopathology (6, 7), likely explained by anatomic differences between murine and human lungs (8), temporal homogeneity of animal models, and potentially distinctive pathobiologic mechanisms operating in human disease. Furthermore, there’s considerable strain variation in response to insults used to induce fibrosis (9). However, alternative animal models may not offer better discrimination for pharmacological assessment. Rats may have histopathology that is more reminiscent of IPF, although direct comparisons between rats and mice suggest similar responses to lung injury. Comparative anatomy of the domesticated pig and ferret more closely resemble humans than do mice (10, 11), and both have been used to model cystic fibrosis (12–14), but neither to study IPF. Australian sheep develop fibrosis in response to bleomycin (15), whereas other animals develop spontaneous lung fibrosis, including horses (16, 17), donkeys (18), cats (19), and West Highland white terriers (20). Horses develop fibrosis after experimental yersiniasis infection (21), but none of the other animals have been proven as tractable models of experimental fibrosis. Furthermore, no therapies have been proven to alter the course of fibrosis in these animals, and the cost of purchase and housing of these species makes them difficult for preclinical studies. However, given the potential advantages associated with the comparative anatomy and spontaneous fibrosis in some of these animals, we would encourage further evaluation of these models.

Currently, the panel recommends that mice be considered the first line animal model for preclinical testing, with rats used subsequently if a second species is required, or practical considerations make mice unsuitable.

Age Considerations

IPF is a disease of advanced age; however, most biomedical research is performed in mice 6–8 weeks old. Estimates have been made to correlate the relative age of mice to human age equivalents (Table 1), but few fibrosis studies have taken advantage of aged mice. Studies assessing bleomycin in older mice revealed more exuberant fibrosis, but this remained associated with enhanced inflammatory responses (22, 23). Some studies have demonstrated mechanisms more reminiscent of IPF (6, 22, 24, 25). Hecker and colleagues (26) found that aged mice suffer from impaired resolution or regeneration after injury, consistent with studies showing that scar formation is reduced in fetal wounds (27–29). In addition, three studies have shown that infectious insult with murine yersiniasis can induce worse fibrosis in mice aged 15–24 months compared with younger mice (30–32).

Taken together, aged mice may more closely reflect human pulmonary fibrosis. While this is an understudied area, there is...
currently no evidence that the pathways leading to fibrosis are different in young compared with old mice, although older mice may resolve injury more slowly (26). Furthermore, histological changes associated with aging, such as increased airway thickness (33), may complicate histological assessments necessary to confirm biochemical endpoints of matrix deposition (e.g., hydroxyproline assays).

Given the significant practical hurdles associated with generation of aged mice, (e.g., time and cost), the panel found no currently compelling reason to recommend either standard or prioritized use of aged mice for pharmacological testing, unless there is a particular age-related target (e.g., epithelial stress responses or redox imbalance) that warrants such investigations.

**Sex Considerations**

Although IFP is more common in men than women, the explanation for this is not fully understood. It may reflect the known effect of estrogens on collagen metabolism protecting premenopausal women, or certain X-linked genes (34) preventing development of lung fibrosis. Similarly a study has highlighted that male mice have a greater response to bleomycin than female mice, regardless of age (22). However, the precise mechanism of protection remains unknown, women do develop pulmonary fibrosis, and the National Institutes of Health currently recommends use of both male and female animals in all research studies.

We would recommend that initial studies be performed in male animals; however, it may be appropriate to confirm key findings in female mice (possibly using higher doses of bleomycin) before progressing lead compounds.

**Genetically Modified Animals**

Genetically modified mice have generated great insights into the pathogenesis of pulmonary fibrosis (34–38). Some transgenic models have been proposed as better models of progressive fibrosis due to the spontaneous development of disease over time (39), although replication and independent validation are needed to confirm these findings. Furthermore, caution must be used in interpreting data from genetically modified mice. Transgenic overexpression of a protein allows crude assessment of that protein’s function, but not interpretations of how physiology may change when the protein is expressed at normal levels. Constitutive “knockout” animals often have genetic compensation for missing genes/proteins that may skew outcomes of pharmacological studies. Similarly, doxycycline and tamoxifen, used to induce transgenic expression, may have indirect effects on inflammatory and fibrotic responses that confound other pharmacological agents. Therefore, these models are better for discovery research than preclinical testing.

Within IFP research, recent studies created transgenic mice harboring gene mutations associated with familial forms of IFP. Examples include increased susceptibility of mice expressing mutant surfactant protein C (40) or genetic deletion of the shelterin component, Trf1, of telomerase from epithelial cells (36) that are more susceptible to bleomycin-induced fibrosis, and, recently, prolonged deletion of TRF1 mutant in surfactant protein C-Cre recombinase–expressing cells by exposure to tamoxifen lead to spontaneous fibrosis (41).

The other area in which genetically modified animals are useful is providing more quantitative analysis of drug action or as lineage reporters. For example, fibroblasts from mice expressing luciferase under control of the collagen promoter show a greater dynamic range than is measurable by hydroxyproline assay (42). Cells from these mice may allow analysis of drug action on collagen expression at earlier time points before protein cross-linking. Performing studies on fibroblasts from these mice *ex vivo* provides an opportunity to test drug actions on a sensitive and dynamic variable, such as collagen gene expression. This could allow for determination of half maximal inhibitory concentration (IC₅₀) doses *in vitro* for later testing *in vivo*.

When these observations are taken together, the panel recommends that genetically modified mice may be more suited to discovery research than drug testing, unless the overexpressed or mutant protein is the candidate drug target. However, certain mutant mice may offer opportunities to purify lineage-traced cell populations or provide greater dynamic range for analysis of collagen gene regulation (42) that could be helpful in conjunction with other fibrotic endpoints and approaches.

**Practical Aspects of Fibrosis Models**

**Identify the A Priori Goal of Each *In Vivo* Experiment**

Before performing *in vivo* modeling experiments, there should be clear *ex vivo/in vitro* rationale from relevant cellular experiments. Identification of relevant targets in annotated human tissue specimens, or use of human or animal organ and tissue cultures, can allow the building of a portfolio to support preclinical advancement. Such studies provide proof of concept and build confidence in the likely success, or otherwise, of *in vivo* studies.

Animal models can roughly be divided into those that have a strong inflammatory component and models of less inflammatory injury. Use of bleomycin, radiation, silica, asbestos, FITC, and many cytokine overexpression systems lead to fibrogenesis after a robust inflammatory response, whereas transgenic delivery of transforming growth factor (TGF)-β, TGF-α, targeted depletion of epithelial cells, and models of IFP fibroblasts delivered to immunodeficient mice are less dependent on inflammation (1). Use of models from each category can help determine whether the drug effect is really independent of inflammation, a crucial consideration when assessing potential antifibrotic molecules. Determining the precise goal of experiments before deciding which *in vivo* platform to use is important. For example, experiments aiming to establish potential therapeutic efficacy, pharmacokinetic and pharmacodynamic relationships, toxicity studies, or studies calculating therapeutic windows may require assessments at different time points. Similarly, understanding whether a drug has engaged its biological mechanism of action is a fundamental aspect of evaluating the outcome of clinical trials, and animal models are particularly useful for developing such pharmacodynamic biomarkers for use in early-phase clinical trials. Thus, care should be taken in determining which models may best provide a platform at different stages of the preclinical development pipeline for each unique drug target. Interestingly, both pirfenidone and nintedanib showed efficacy, albeit marginal, in more than one animal model (43) (Table 2).
<table>
<thead>
<tr>
<th>Reference No.</th>
<th>Drug</th>
<th>Species</th>
<th>Insult</th>
<th>Quantified Fibrosis Effect</th>
<th>Histology</th>
<th>Lung Function/Imaging</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>108</td>
<td>Aerosolized pirfenidone D 8</td>
<td>C57Bl/6 mice GNS</td>
<td>IT Bleo</td>
<td>Inhibited Sirius red biochemistry and Ashcroft score D 22, n = 4</td>
<td>H&amp;E Masson’s trichrome</td>
<td>Pirfenidone increased dynamic compliance</td>
<td>NR</td>
</tr>
<tr>
<td>109</td>
<td>Pirfenidone 10 mg/kg PO, D 6</td>
<td>C57Bl/6J female</td>
<td>IT Bleo</td>
<td>Inhibited sircol D 14 No effect on Ashcroft score n = 6</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>110</td>
<td>Pirfenidone 500 mg/kg PO, D −1</td>
<td>SD rats male</td>
<td>IT Bleo</td>
<td>No effect on Col1 mRNA D 28, n = 8</td>
<td>H&amp;E Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>111</td>
<td>Pirfenidone 300 mg/kg PO, D 0</td>
<td>C57Bl/6 female</td>
<td>OM Bleo</td>
<td>Reduced collagen Sircol and fibrosis score D 14, n = 6</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>112</td>
<td>Pirfenidone 400 mg/kg PO, D 10</td>
<td>C75Bl/6 male</td>
<td>IT Bleo</td>
<td>No effect on hydroxyproline, n = 10</td>
<td>Sirius red H&amp;E</td>
<td>NR</td>
<td>40% Mortality with Pirf/Bleo. No significant difference from control, n = 40</td>
</tr>
<tr>
<td>113</td>
<td>Pirfenidone 100 mg/kg BD, D 3–23</td>
<td>Swiss-albino mice male</td>
<td>IT Bleo</td>
<td>Possible reduction hydroxyproline D 21, n = 3, reduced fibrosis score, n = 5</td>
<td>H&amp;E Masson’s trichrome</td>
<td>HRCT</td>
<td>Improved mortality in pirfenidone approx. 50% versus 60%</td>
</tr>
<tr>
<td>114</td>
<td>Pirfenidone 10 μg /D Aerosolized D 10</td>
<td>C57Bl/6 male</td>
<td>IT Bleo</td>
<td>Reduced hydroxyproline, n ≤ 4, time point not clear</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>115</td>
<td>Pirfenidone 500 mg/kg, PO, D −1</td>
<td>ICR mice male</td>
<td>IV Bleo</td>
<td>Reduced Ashcroft score, D 28, n = 8</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>116</td>
<td>Pirfenidone 400 mg/kg, PO D 0, PO D 7, PO D 14</td>
<td>Mice GNS</td>
<td>IT Bleo</td>
<td>Reduced hydroxyproline n = 7, D 0–14 reduced hydroxyproline, n = 12, D 7–14</td>
<td>H&amp;E Masson’s trichrome</td>
<td>Reduced elastance, n = 11, D 0–14, reduced elastance, n = 24, D 7–14</td>
<td>NR</td>
</tr>
<tr>
<td>117</td>
<td>Pirfenidone 100 mg/kg TDS</td>
<td>ICR mice male</td>
<td>IV Bleo</td>
<td>Reduced hydroxyproline n = 10, D 10 and D 28, reduced fibrosis score, n = 10</td>
<td>H&amp;E Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>118</td>
<td>Pirfenidone 400 mg/kg/dy D 14</td>
<td>ICR mice male</td>
<td>IV Bleo</td>
<td>Reduced fibrosis score n = 3 and hydroxyproline (n = 3–5) D 35 and D 49,</td>
<td>H&amp;E</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>119</td>
<td>0.5% Pirfenidone PO D −2</td>
<td>Golden Syrian hamster male</td>
<td>IT Bleo</td>
<td>Significant Reduction in hydroxyproline D 14 &amp; D 21, n = 4</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Reference No.</td>
<td>Drug</td>
<td>Species</td>
<td>Insult</td>
<td>Quantified Fibrosis Effect</td>
<td>Histology</td>
<td>Lung Function/Imaging</td>
<td>Mortality</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>---------</td>
<td>--------</td>
<td>----------------------------</td>
<td>-----------</td>
<td>-----------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>120</td>
<td>Pirfenidone</td>
<td>C57Bl/6 female</td>
<td>IT Bleo</td>
<td>No effect on fibrosis score D 14</td>
<td>H&amp;E Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>400 PO BD D 0</td>
<td>0.5% pirfenidone in diet after second Bleo dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>0.5% pirfenidone in diet after second Bleo dose</td>
<td>Golden Syrian hamster male</td>
<td>IT Bleo x 3</td>
<td>Reduced hydroxyproline, n = 8</td>
<td>H&amp;E</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>122</td>
<td>BIBF1000 (precursor to nintedanib, BIBF 1,120) 50 mg/kg D 0–21 or D 10–21</td>
<td>Wistar rats male</td>
<td>IT Bleo</td>
<td>Reduced procollagen mRNA</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>123</td>
<td>Nintedanib 30 or 60, mg/kg PO D 0–14 or 7–21</td>
<td>C57Bl/6 female</td>
<td>IN Bleo</td>
<td>D 0–7 reduce soluble collagen (Sircol) and fibrosis score, therapeutic delivery reduced fibrosis score, trend to reduced soluble collagen; n = 10</td>
<td>H&amp;E or chromotrope aniline blue</td>
<td>No effect on airway resistance or lung compliance</td>
<td>NR</td>
</tr>
<tr>
<td>124</td>
<td>Nintedanib 30, 50 mg/kg PO D 0–21 or 50 D 10–21</td>
<td>Wistar rats male</td>
<td>IT Bleo</td>
<td>Preventative and therapeutic dosing reduced qualitative fibrosis on histology; n = 10</td>
<td>Masson’s trichrome</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Definition of abbreviations: BD, twice per day; Bleo, bleomycin; D, day; GNS, sex not specified; H&E, hematoxylin and eosin; HRCT, high-resolution computed tomography; ICR, imprinting control region; IN, intranasal; IT, intratracheal; IV, intravenous; NR, not reported; OM, osmotic mini-pump; PO, per oral; SD, Sprague-Dawley; TDS, total dissolved solids.*
Figure 1. Integrating human disease data with preclinical animal studies. Preclinical assessment of drug candidates should make judicious use of animal models, such as the bleomycin model in mice, with strong biochemical and histologic assessments of collagen deposition. In addition, animal models and culture systems can be used for efficacy, safety, and imaging studies. Target validation should involve annotated human specimens. An integrated approach that tests effects of compounds on established fibrosis should help move promising targets to clinical trials. iPSCs, inducible pluripotent stem cells; PD, pharmacodynamics; PK, pharmacokinetics. Figure adapted from a slide shown by Dr. Sheila Violette, who presented this at the American Thoracic Society Conference in May 2015.

The panel’s recommendations are for an integrated approach to drug development using both relevant animal models and appropriate ex vivo/in vitro approaches to provide robust validation of pharmacological mechanisms of action and a dosing framework (Figure 1).

Route of Delivery in Lung Fibrosis Models
When considering bleomycin, most investigators use a single intratracheal administration (1), but repeated intratracheal (44) or repeated intraperitoneal (45) administrations may offer more robust and nonresolving fibrotic pathology. Repetitive intraperitoneal injections may result in profound epithelial cell hyperplasia, although the evidence is insufficient to make firm recommendations.

When considering a single administration of bleomycin, the oropharyngeal route leads to pulmonary fibrosis persisting for up to 6 months (46), while requiring shorter-duration animal sedation, eliminating a surgical procedure and associated postsurgical analgesia, and allowing more rapid post-procedure recovery than the intratracheal route of administration. The fibrotic response that develops after both methods appears similar; however, direct comparisons between these two routes of administration have not been formally assessed. Two different models may provide more robust data for preclinical assessment. Reasonable choices would include adenoviral-driven TGF-β overexpression, FITC, asbestos, or silica, all administered into the lung, or radiation delivered locally to the thorax (reviewed in Refs. 1, 9).

The panel recommends use of a single oropharyngeal administration of bleomycin and that a second model be considered for preclinical testing. In addition, the U.S. Food and Drug Administration will require at least two species for toxicity, but not efficacy studies.

Kinetics of Lung Fibrosis Models
The time course for fibrotic outcomes should be well characterized, and standardized within laboratories. This is important when considering timings of intervention with anti-fibrotic compounds. The bleomycin model is characterized by periods of acute lung injury (Days 0–7), fibroproliferation (Days 3–14), and established fibrosis (generally Days 14–28) that generally resolves over a variable time period (47).

The consensus of the panel is that it is insufficient to only test an anti-fibrotic compound before evidence of histological fibrosis (e.g., giving the drug before Day 7 in the acute bleomycin model). Rather, it is more meaningful to deliver the drug therapeutically after histological evidence of fibrosis (e.g., no sooner than Days 7–10 in the acute bleomycin model) postinstillation after peak of the acute inflammatory phase of the lung injury response.

When considering therapeutic dosing in other models (e.g., adenoviral–TGF-β), results are less likely to be confounded by anti-inflammatory effects. However, it remains important to ensure that therapies are only administered after the resolution of any inflammation, and the precise timing in other models of fibrosis is an important area for future clarification.

It is unclear whether recurrent bleomycin administration offers any advantage over the acute bleomycin model for preclinical assessment of anti-fibrotic therapies. The major advantage of this model is that recurrent injury and fibrosis may better replicate the progressive nature of human disease, although the mechanistic assumption remains the same, namely, that progressive fibrosis is caused by an external insult rather than internal genetic/epigenetic response to injury. However, we would support further investigation to determine whether these models lead to epigenetic changes, which promote fibrosis.

Reproducibility and Statistical Power
A key aspect of improving translatability of preclinical assessment is choosing appropriate, quantitative endpoints (discussed subsequently here) and undertaking a priori power calculations using a predefined primary outcome measure. All other endpoints should be considered secondary, exploratory endpoints requiring further assessment for validation. Care needs to be taken in dealing with missing data, especially in animals that may not complete the study due to death, or breaching animal welfare limits. Although not currently routine, consideration should be given to undertaking sensitivity analysis in circumstances of high levels of missing data (4). Another practice recommendation...
would be to test potential drug candidates in a second laboratory working in the same model, or even a different model.

The panel highly recommends that negative data should be published, and hopes that journal editors will value publishing high-quality negative results to avoid unnecessary duplication of effort and to reduce unnecessary animal experimentation.

Endpoints for Preclinical Assessments

The most accurate analysis for lung fibrosis involves several endpoints, including combinations of biochemical quantification of collagen content, histologic assessment of fibrotic distribution, and other optional parameters, such as gene expression, lung function, and radiographic imaging. Here, we consider common methodologies used for fibrosis endpoint assessments and consider emerging endpoints that may have value in the future.

Biochemical Assessment of Collagen Content

Hydroxyproline content in total lung samples is a surrogate for collagen content (1 μg hydroxyproline = 6.94 μg collagen), and is expressed as micrograms per lung (right, left, or both) (48). Commonly, fibrosis is patchy after introduction of the stimulus by inhalation, often with considerable variation between lungs. Therefore, processing both lungs to powder in liquid nitrogen or homogenized together before analysis of hydroxyproline, or other biochemical assessments, should be considered. If the fibrotic stimulus is deposited with excellent reproducibility, the same lung should be consistently used for this assay to allow comparison between mice and groups. Although this may increase the standard deviation of the data obtained, it permits use of one lung for morphometric analysis, which, overall, could reduce the number of animals required in the experiment. Sirius red is often used as a colorimetric estimate of fibrosis in tissue homogenates; however, the values seem to overestimate absolute values obtained by the gold standard hydroxyproline assay (48, 49). In addition, collagen content can be determined using Sircol collagen assays kits. However, it is important to be aware that this method only allows analysis of “newly formed” acid and pepsin-soluble collagens, and, therefore, is not useful for measuring insoluble collagens integrated into “mature” scar tissue, and Sircol measurements account for only a small fraction of total lung collagen determined by HPLC (46). In addition, care must be taken to use a modified version of the assay that is not adversely impacted by serum proteins present in the samples due to vascular leak. Thus, although this method is suitable for collagen produced in cell culture systems, the panel believes that this assay does not best reflect the effect of an intervention on lung collagen accumulation and fibrosis in vivo.

Given these considerations, the panel recommends hydroxyproline measurement as the optimal primary endpoint for preclinical assessment of novel therapeutic agents.

Gene Expression

Another common practice is reporting extracellular matrix levels by measuring levels of gene expression, generally by quantitative PCR. Although these measurements can be supportive to the overall assessment, they are difficult to interpret if done in isolation. For example, collagen mRNA is extremely stable (50), so small changes in gene expression can correlate with large differences in protein levels.

The panel currently recommends that gene expression studies always be accompanied by biochemical parameters.

Morphological Approach for Histological Assessment

Masson’s trichrome staining is used to assess fibrosis histologically, and, usually, the severity of lung fibrosis is semiquantified in stained tissue sections through an Ashcroft scoring system (51), where the grade of fibrosis is scored from 0 (normal lung) to 8 (total fibrous obliteration of fields) by examining randomly chosen sections. More recently, digital scoring systems using semiautomated image analysis have been used (52, 53) or a combination of semiquantitative scoring by blinded readers and digital analysis (54). The Sirius red stain (55) consists of a dye that binds to the Gly-x-y triple-helix structure found in all collagen fibers. Fibriilar collagens are visualized as birefringent structures under polarized light, but tissues can also be examined by brightfield microscopy and semiquantified by morphometric (visual) scoring systems (56–59). Recently, a digital imaging addition and subtraction method was proposed to determine the relative area of fibrillar collagens and cell content in a semiautomated process using standard software (60). For morphological analysis, the panel recommends that lungs should be fixed at a constant hydrostatic pressure up to a maximum of 30 cm H2O based on normal lung (61, 62). Incomplete perfusion provokes collapse, whereas excessive pressure may provoke alveolar wall rupture, leading to erroneous interpretation. With regard to measures of α-smooth muscle actin to identify myofibroblasts as a surrogate for fibrosis, recent studies suggest that collagen and α-smooth muscle actin do not always correlate with each other (63, 64).

The panel recommends that due to the inherent bias associated with morphological scoring systems, exacerbated by the sporadic nature of histologic assessment, histology should focus on the morphological and molecular characteristics of fibrosis but must never be used to quantify fibrosis without biochemical assessment of lung collagen.

Lung Function

Lung function assessment is helpful in corroborating fibrotic changes and/or therapeutic response of interventions, but presents considerable technical challenges (52, 54, 65). Lung function measures show good correlation with morphological changes at peak of fibrosis, but recover when the lesion extent decreases, even when there is still clear evidence of histologic fibrosis (52). Some functional parameters can be measured through noninvasive methods with whole-body plethysmography, where the animal is conscious and unrestrained; however, this technique evaluates few parameters (e.g., respiratory frequency, tidal volume), and these measures show marginal or no differences between injured and normal lungs (66).

Invasive methods are performed with forced ventilator maneuvers in anesthetized, tracheostomized animals; thus, a disadvantage is that these measures may only be used as single-timepoint outcomes (66). There are two invasive methods, Buxco-force pulmonary maneuvers, and FlexiVent, which allow the measurement of several functional tests of clinical relevance for lung fibrosis, such as pressure–volume curves (66). With forced oscillation techniques, lung impedance can be examined and, through it, several variables, such as lung tissue elastance.
and tissue damping (54, 66, 67). Some studies suggest that values obtained by forced oscillation techniques correlate better with extent of fibrosis than do pressure–volume curves (47, 54). In larger animals (e.g., rats), gas transfer can also be measured. Importantly, lung mechanics measurements with these invasive methods before killing experimental mice should not alter results obtained by histology or collagen determinations performed on the same lungs (54).

In general terms, pulmonary function tests can evaluate response to drugs or antifibrotic mediators (68, 69). Certainly, an important goal in the field is to develop methods that allow monitoring the same mouse over time. As such, repetitive, invasive measurements of lung function may be performed in intubated instead of tracheostomized mice with similar results (70). However, studies so far are limited, and the effect of repeated invasive tests on morphology and collagen content is unknown. Alternatively, the measurement of arterial blood oxygen saturation with pulse oximetry is noninvasive, and can be tested on individual animals in pulmonary fibrosis models, predominantly using in vivo micro-CT of the lungs (73–76). Micro-CT is feasible and allows dynamic visualization of disease progression and response to therapy. However, the degree of resolution is constrained by respiratory and cardiac motion; therefore, implementation of gating for small-animal CT imaging is recommended (77). Abnormal images relevant to interstitial lung diseases, such as ground glass attenuation, reticular opacities, consolidation, traction bronchiectasis, and honeycombing, can be semiquantified by radiologists (blinded to the pathologic results) scoring, for example, on a scale from 0 (absent) to 5 depending on the involved area (78). Micro-CT data can also be digitally processed and the aerated and total lung volumes, lung tissue (including lesions), and mean lung density can be more precisely quantified (73). Importantly, several studies showed that repeated high-resolution lung micro-CT scans for long-term imaging protocols don’t provoke radiotoxicity-induced lung damage (79, 80).

High-resolution micro-CT analysis with spatial resolution of approximately 20 μm has also been employed in murine lungs ex vivo. Fibrotic changes revealed by micro-CT fully matched equivalent histologic sections, and this approach allows evaluation of therapeutic dosing once fibrosis is already established (46).

The panel didn’t feel there was sufficient evidence to recommend routine use of CT scanning for preclinical assessment of antifibrotic therapy. However, the potential advantages associated with CT scanning, especially when measuring evolution of fibrosis over time, merit further investigation.

**Inflammation**

The role of chronic inflammation in IPF remains controversial (81, 82). However, it may be useful to quantify changes, as inflammation does not completely

### Table 3. Summary of Suggested Recommendations to Evaluate Lung Fibrosis in a Preclinical Setting

1. Initial studies should assess bleomycin administered to C57Bl/6 mice ideally via oropharyngeal administration, and consideration should be given to testing both male and female mice for key endpoints.
2. The investigational product should be administered after the acute inflammatory response has subsided, depending on the model kinetics, at least 7–10 d after bleomycin instillation.
3. The study should be powered for a primary endpoint of change in total lung hydroxyproline levels, and key secondary endpoints should include histologic assessment of morphology using either Masson’s trichrome or Sirius red staining. Stitched images of at least a whole lobe should be analyzed and reported. All other endpoints at the current time should be considered exploratory secondary endpoints.
4. Experiments should be reported in line with the ARRIVE guidelines (3), and data should be made available for sharing and meta-analysis. Outcomes of all animals in the study should be described.
5. If the initial assessment is positive, a practice recommendation would be for results to be repeated in a different center. Inclusion of an additional model or different species (e.g., rats) might also increase confidence.
6. The panel encourages researchers to consider peer review and publication (either in open-access journals or institutional websites) of their experimental protocol before undertaking the in vivo experiment in an effort to ensure that controls and experimental power are sound.
7. The panel would encourage further exploration of murine and human lung slice studies to further reduce and replace the reliance on in vivo animal models.
8. The panel would encourage further investigation of novel endpoints with clear translational potential that demonstrate target engagement of mechanism.
9. The panel would encourage further efforts to identify models that show progressive, nonresolving fibrosis for preclinical assessment of novel investigational products.

*Definition of abbreviations: ARRIVE, Animal Research: Reporting of In Vivo Experiments; CT, computed tomography.*
resolve during the evolution of the fibrotic response, and might confound interpretations, especially with agents that impact inflammatory response. Inflammation is usually evaluated in bronchoalveolar lavage (BAL) or lung digest specimens. In the bleomycin model, BAL to obtain lung leukocytes is typically done in the first week postinstillation. Inflammatory cell profiles, including total cell numbers, percentages, and even cell subsets, such as M1/M2 macrophages and CD4/CD8 lymphocytes, can be determined. In addition, cytokines are easily explored in BAL fluids or lung homogenates. The extent of inflammatory lesions may be semiquantified in histology sections from lungs stained with hematoxylin and eosin. However, the panel does not recommend routine assessment of inflammatory cell profiles for the preclinical assessment of antifibrotic drugs, except to a priori confirm presence or absence of effects on inflammatory pathways for a given molecule.

**Emerging Endpoints**

**Novel Imaging Studies**

Imaging modalities are widely used to assess fibrosis; although many are in early development for in vivo modeling, a few warrant commentary as potential preclinical surrogates. Molecular imaging techniques use tracer molecules labeled with contrast reagents for visualization to track cellular events and molecular processes in their native environments in intact living subjects (83). These techniques allow assessment of important molecular targets not shed into extracellular fluids that cannot be sampled in peripheral blood. Molecular imaging probes visualizing recently synthesized collagen, or the TGF-β-activating integrin αvβ6, have already been demonstrated to noninvasively track development of pulmonary fibrosis in the bleomycin mouse model by magnetic resonance imaging or SPECT, respectively (84, 85). High levels of αvβ6 integrins, detected by immunohistochemistry, correlate with poor outcome in IPF (86). Using SPECT scanning, it is possible to identify up-regulated αvβ6 integrins in bleomycin-induced lung fibrosis, predict the development of fibrosis, and demonstrate target engagement using anti-αvβ6 integrin blocking antibodies (85). Use of SPECT and other noninvasive imaging techniques combined with molecular probes for use in murine models of lung fibrosis has recently been reviewed (87). Second harmonic generation microscopy is an exciting, emerging, “label-free” technique for assessing collagen in various tissues, and has been used to score liver fibrosis (88) and to assess collagen structure in IPF (89) and bleomycin-induced lung fibrosis (90). However, its greatest potential is for real-time collagen assessments in vivo, in both animal models and human disease, through microendoscopic second harmonic generation microscopy (91).

**Novel Peripheral Blood Biomarkers**

Several prognostic/diagnostic biomarkers have been identified in IPF using transcriptional or protein profiling (92–96). Another emerging technology is measurement of matrix neo-epitopes created during fibrotic remodeling. Examples include specific fragments of type IV collagen α1 (C4M12a1) and α3 (C4M12a3) chains in serum as indicators of fibrosis when measured by ELISA (97). Recent human IPF studies have also shown that neo-antigens generated by matrix metalloproteinase cleavage are useful biomarkers in IPF (98). Thus, one potential advantage of these techniques is that their use in animal models may be translated to human studies. One note of caution is that neo-epitope panels currently in use for human studies have not been fully characterized for murine tissues. There are several potential reasons: first, cleavage sites on proteins in mice and humans differ; and, second, quantities of cleaved proteins differ significantly at baseline in the two species.

When taken together, the panel is excited about these emerging imaging and biomarker technologies, but does not believe that there is currently sufficient experience to recommend this approach for general preclinical drug testing. The panel recommends that investigators continue to develop specific imaging probes and blood biomarkers to predict disease progression, and to predict responses to new treatments. The panel is also enthusiastic about the high potential translatability of these approaches into clinical studies.

**Ex Vivo Lung Slices for Preclinical Testing**

Ex vivo precision-cut human lung slices for preclinical testing are generally obtained from lung cadaver, surgical resections, or explanted lungs, which are filled with low melting point agarose to allow precision slices of 300–1,000 μm in thickness for use in culture (99, 100). Such lung slices are viable in culture for 5–7 days and preserve three-dimensional architecture (99). By Day 7, such cultures generally show prominent apoptosis, but movement of media over the slices may increase viability. These slices can also be obtained from the lungs of patients with IPF and used to validate findings from in vitro and in vivo studies in human tissue, as well as a dosing framework of novel therapeutic agents (101).

Preliminary studies treating slices with bleomycin, or influenza, have not yielded changes ex vivo that are observed in vivo (29, 73), although encouraging results using TGF-β have been found (102). It is possible to generate lung slices from mice that have been treated with bleomycin in vivo, and these have increased levels of extracellular matrix (103–105); there are encouraging results using nintedanib and pirfenidone in these slices as well (102). An important feature of murine or human lung slice cultures is that they can be used to assess effects of drugs directly on structural cells in the absence of vasculature or inflammatory cell influx. Furthermore, a single mouse lung can generate approximately 30 slices for analysis, making this an emerging alternative to some whole-animal experiments, potentially reducing overall animal numbers used for research.

In conclusion, the panel is excited about the promise that lung slice cultures hold for identification of novel targets, kinetics studies, and ability to compare biology between human and murine systems, and the investigation of slices from IPF lungs. The panel encourages further investigation comparing pharmacological manipulation of fibrotic murine and human lung slices to further develop and improve these assays, and accumulate evidence to use this approach for future preclinical drug testing.

**Conclusions**

The panel recommendations (Table 3) are for comprehensive, preclinical evaluation of potential antifibrotic therapies based on currently available data. They are not to be regarded as a minimum dataset for publication, or as recommendations applicable to discovery biology, although some principles are clearly transferable. ■
This official workshop report was prepared by an ad hoc committee of the Assembly on Respiratory Cell and Molecular Biology.

Members of the subcommittee are as follows:

R. Gisli Jenkins, M.D., Ph.D. (Co-Chair)
Bethany B. Moore, Ph.D. (Co-Chair)
Rachael C. Chambers, Ph.D., FRSB
Oliver Eickelberg, M.D.
Martin Kolb, M.D., Ph.D.
Melanie Konigshoff, M.D., Ph.D.
Geoffrey J. Laurent, Ph.D., F.Med.Sc.
Carmel B. Nathakumar, Ph.D.
Mitchell A. Omlan, M.D., M.A.
Annie Pardo, Ph.D.
Moxes Selman, M.D.
Dean Sheppard, M.D.
Patricia J. Sime, M.D.
Victor J. Thannickal, M.D.
Eric S. White, M.D., M.S.

Author disclosures: R.G.J. received research support from Biogen, Galecto, GlaxoSmithKline, and Novartis Pharma; served on an advisory committee for Biogen, Boehringer Ingelheim, GlaxoSmithKline, InterMune, PharmAkea, and Roche; served as a consultant and as a speaker for InterMune; served as a consultant for Nuformix and Pulmatic; and served on a data and safety monitoring board for Boehringer Ingelheim. B.B.M. received research support from Syntex Biosystems. R.C.C. received research support from GlaxoSmithKline; her spouse is an employee of and has stocks, stock options, or other ownership interests in GlaxoSmithKline. M. Kolb received research support from Actelion Pharmaceuticals, Boehringer Ingelheim, Gilead Sciences, Janssen Pharmaceuticals, and Prometic; served as a consultant for Boehringer Ingelheim, F. Hoffmann-La Roche, Gilead Sciences, GlaxoSmithKline, Janssen Pharmaceuticals, and Prometic; served as a speaker for Boehringer Ingelheim and F. Hoffmann-La Roche; served on an advisory committee for AstraZeneca; and served on a data safety and monitoring board for F. Hoffmann-La Roche. D.S. served as a consultant, had commercialized intellectual property, and had stocks, stock options, or other ownership interests with Plant Therapeutics. P.J.S. served on an advisory committee for GlaxoSmithKline and UCB Biosciences; served as a consultant for Boehringer Ingelheim and Third Rock Ventures; and served on a data and safety monitoring board for InterMune. A.M.T. received research support from Biogen Idec, InterMune, and Boehringer Ingelheim; served on an advisory committee and has stocks, stock options, or other ownership interests with PharmAkea Therapeutics and Plant Therapeutics; and served as a consultant for Blade Therapeutics. O.E., M. Königshoff, G.J.L., C.B.N., M.A.O., A.P., M.S., A.L.T., V.J.T., E.S.W. report no commercial interests.

References

1. Moore BB, Lawson WE, Oury TD, Sisson TH, Raghavendran K, R.G.J. Jenkins, M.D., Ph.D. (Co-Chair). Members of the subcommittee are as follows: R. Gisli Jenkins, M.D., Ph.D. (Co-Chair); Bethany B. Moore, Ph.D. (Co-Chair); Rachael C. Chambers, Ph.D., FRSB; Oliver Eickelberg, M.D.; Martin Kolb, M.D., Ph.D.; Melanie Konigshoff, M.D., Ph.D.; Geoffrey J. Laurent, Ph.D., F.Med.Sc.; Carmel B. Nathakumar, Ph.D.; Mitchell A. Omlan, M.D., M.A.; Annie Pardo, Ph.D.; Moxes Selman, M.D.; Dean Sheppard, M.D.; Patricia J. Sime, M.D.; Victor J. Thannickal, M.D.; Eric S. White, M.D., M.S.; R. Gisli Jenkins, M.D., Ph.D. (Co-Chair); Bethany B. Moore, Ph.D. (Co-Chair); Rachael C. Chambers, Ph.D., FRSB; Oliver Eickelberg, M.D.; Martin Kolb, M.D., Ph.D.; Melanie Konigshoff, M.D., Ph.D.; Geoffrey J. Laurent, Ph.D., F.Med.Sc.; Carmel B. Nathakumar, Ph.D.; Mitchell A. Omlan, M.D., M.A.; Annie Pardo, Ph.D.; Moxes Selman, M.D.; Dean Sheppard, M.D.; Patricia J. Sime, M.D.; Victor J. Thannickal, M.D.; Eric S. White, M.D., M.S.; R. Gisli Jenkins, M.D., Ph.D. (Co-Chair); Bethany B. Moore, Ph.D. (Co-Chair); Rachael C. Chambers, Ph.D., FRSB; Oliver Eickelberg, M.D.; Martin Kolb, M.D., Ph.D.; Melanie Konigshoff, M.D., Ph.D.; Geoffrey J. Laurent, Ph.D., F.Med.Sc.; Carmel B. Nathakumar, Ph.D.; Mitchell A. Omlan, M.D., M.A.; Annie Pardo, Ph.D.; Moxes Selman, M.D.; Dean Sheppard, M.D.; Patricia J. Sime, M.D.; Victor J. Thannickal, M.D.; Eric S. White, M.D., M.S.;


44. Campa JS, McAnulty RJ, Laurent GJ. Application of high-pressure liquid chromatography to studies of collagen production by isolated cells in culture. *Anal Biochem* 1990;186:257–263.


