Bui Allan (Orcid ID: 0000-0003-3060-8239)

**Title:** Pulmonary function testing for the early detection of drug induced lung disease: a systematic review in adults treated with drugs associated with pulmonary toxicity

**Authors:** Allan Bui[1], Sangjin Han[7], Marliese Alexander[2,3], Guy Toner[4,5], Lou Irving[1,5,6], Renee Manser[1,5,6]
[1] Department of Respiratory and Sleep Medicine, Royal Melbourne Hospital, Parkville, Victoria, 3050, Australia
[2]Pharmacy Department, Peter MacCallum Cancer Centre, Melbourne, Victoria, 3000, Australia.
[3] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, 3000, Australia
[4]Department of Medical Oncology, Sir Peter MacCallum Cancer Centre, Melbourne, Victoria 3000, Australia.
[5]Department of Medicine, St Vincent's Hospital, University of Melbourne, Fitzroy 3065, Australia
[6]Department of Medicine (Royal Melbourne Hospital), University of Melbourne, Parkville, Victoria 3050, Australia.
[7]University Hospital Geelong, Barwon Health, Geelong, Victoria 3220, Australia

**Contact Address:** Department of Respiratory and Sleep Medicine, Royal Melbourne Hospital, 300 Grattan St, Parkville VIC 3050
**Contact email:** allan.bui2@mh.org.au

**Keywords:** pulmonary toxicity; pulmonary function testing; diffusing capacity of the lung for carbon monoxide; Bleomycin; Amiodarone

**Funding**: None
**Disclosure statement:** there are no competing interests to declare from any of the authors.
**Acknowledgements:**
We are grateful to the support of the Department of Respiratory and Sleep Medicine at Royal Melbourne Hospital for supporting costs of printing and retrieving articles.
We are grateful to Smaro Lazarakis, Librarian, Peter MacCallum Cancer Centre, for assisting with the development of our search strategy, running the search and providing us with the results of the search.
We are also grateful for the statistical support from the Melbourne Epicentre – a collaborative research centre between the University of Melbourne and Royal Melbourne Hospital.
**Word count:**
**Abstract** – 249
**Manuscript** - 3407

## Abstract

**Introduction:** Pulmonary function tests (PFT) are sometimes monitored during treatment with known pulmonary toxic drugs to detect asymptomatic drug induced interstitial lung disease (DILD). We conducted a systematic review to assess the accuracy of PFTs, including the diffusing capacity for carbon monoxide (DLCO), for early detection of DILD in a range of drugs.

**Methods:** Using a pre-specified, registered review protocol, OvidMEDLINE and EMBASE were searched from 1946 to February 2018. Two reviewers independently screened abstracts and reviewed full text articles for inclusion. Included studies were assessed for risk of bias using adapted QUADAS-2 domains and primary outcome data was extracted and entered into RevMan5 to estimate sensitivity and specificity with 95% confidence intervals (95% CIs).

**Results:** The search identified 4065 citations and included 42 studies. The most commonly studied drugs were Bleomycin and Amiodarone. Due to clinical heterogeneity between studies, a pooled analysis was not performed. Sensitivity of monitoring with DLCO varied between 0% and 100%, with the majority of studies finding a sensitivity of <80%. Confidence intervals were wide for the majority of studies. Specificity was less than 90% in all studies. Risk of bias was high for the majority of studies for the quality domain of reference standard.

**Conclusion:** The findings of this review do not support routine PFTs for early detection of DILD. Due to methodological limitations, the relatively small number of participants and the low prevalence of DILD in the included studies, there remains significant uncertainty about the sensitivity of PFTs to screen for DILD.

## Introduction

Drug induced interstitial lung disease (DILD) is a common manifestation of drug induced respiratory disease. Epidemiological data on the incidence and mortality of DILD is sparse. One European registry study of interstitial lung disease suggests that 2.5 % to 5% of cases are drug or radiation induced (1). Amongst those hospitalised with drug induced lung disease, the mortality appears to be high, with one study reporting a 17% mortality rate (2). While hundreds of drugs have been associated with respiratory toxicity, common aetiological agents include chemotherapeutic agents (such as Bleomycin), and Amiodarone (3, 4). There is evidence that the diagnosis of interstitial lung disease in the community is often delayed due to both physician and patient factors, with the majority of patients attributing symptoms to other causes and delaying presentation for several months, however it is not clear whether these findings apply to DILD (5). Presenting features of DILD are often non-specific and one recent study found that cough, but not shortness of breath, was associated with the development of pulmonary toxicity with Bleomycin as assessed by CT (6). Monitoring with periodic pulmonary function tests (PFTs) such as spirometry and the diffusing capacity for carbon monoxide (DLCO) has been used for many decades in order to detect pre-symptomatic toxicity and modify treatment appropriately, particularly in the setting of Bleomycin and Amiodarone therapy. However observational studies report conflicting findings about the utility of this approach (7-11).

To our knowledge, few evidence-based clinical practice guidelines have addressed the role of monitoring with PFTs during treatment with drugs associated with DILD. Literature reviews have been limited to a narrative approach. One European-based guideline on the use of Methotrexate in Rheumatoid Arthritis does not recommend routine PFT monitoring (12). The U.S Food and Drug Administration currently recommends that DLCO be monitored monthly during Bleomycin treatment. (13). As a result of rare fatalities associated with the prolonged use of Nitrofurantoin, the British National Formulary has published recommendations that PFTs should be done on patients receiving Nitrofurantoin for longer than 6 months (14). In the absence of a consensus approach to monitoring, current practice appears to be variable (15). Furthermore, monitoring may be associated with false positive results and overdiagnosis, with particular concern with implications of Bleomycin dosing in the management of a germ cell tumours. (16).

The aim of this review was to assess the accuracy of periodic testing with PFTs including total lung capacity, spirometry and DLCO, for the early detection of DILD in people receiving treatment with drugs known to cause pulmonary toxicity.



## Methods

**Study Protocol and Registration**

This systematic review protocol was developed and prospectively registered on PROSPERO (PROSPERO ID: CRD42018090149). The protocol included search strategy, inclusion and exclusion criteria, primary and secondary outcome measures, risk of bias assessment and methods of analysis.

**Literary databases and search strategy**

We searched Ovid MEDLINE* (1946 to February 2018) and EMBASE (1974 to February 2018) using the search strategy outlined in **Appendix A**. The search strategy was designed after an extensive preliminary literature search and in consultation with a librarian. Search results were limited to human studies involving adult participants (age>18), published in the English language and excluded case reports. Reference lists of included studies, relevant book chapters and review articles for additional studies were also searched for relevant papers. Google Scholar was used to search for grey literature (research that is unpublished or published in a non-commercial form) and conference abstracts.

A list of drugs known to be associated with pulmonary toxicity included in the search is available in **Appendix B**. The search was divided in to two components. In the first search, we targeted observational studies reporting pulmonary function tests in participants treated with any of the drugs listed in **Appendix B.** In a second smaller search, we targeted randomised control trials where Bleomycin was used in the treatment of germ cell tumours as pulmonary toxicity is commonly monitored during administration).

**Study selection and eligibility**

We included prospective or retrospective cohort studies or single or both arms of randomised control trials where lung function monitoring was undertaken during treatment with any of the drugs listed in

**Appendix B.** Only studies reporting the incidence of pulmonary toxicity and performing at least one PFT during drug treatment were included.

The primary outcome was sensitivity and specificity of PFTs for the early detection of DILD. PFTs included forced expiratory volume in one second (FEV1), forced vital capacity (FVC), diffusing capacity for carbon monoxide (DLCO) or transfer capacity of the lung for carbon monoxide (TLco) and total lung capacity (TLC) via plethysmography or nitrogen washout. The secondary outcome was change in lung function over time during drug treatment.

We excluded studies with participants aged <18 or with a sample size of less than 10 and those in which data for patients receiving concurrent radiotherapy (chest or mediastinal) or lung surgery had not been analysed separately to patients receiving chemotherapy alone.

Two independent reviewers (AB, SJH) screened titles and abstracts and reviewed full text articles against a pre-defined set of inclusion and exclusion criteria as shown in **Appendix B.** Disagreements were resolved with discussion and consensus with a referral to a third author where necessary (RM).

**Risk of bias analysis and data extraction**

Two independent reviewers reviewed the eligible full text articles for quality assessment and data extraction. Disagreements were resolved by discussion or referral to a third reviewer when necessary. A critical appraisal checklist using the pre-defined adapted tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) domains (patient selection, index test, reference standard and flow and timing) was developed in advance and used to assess risk of bias of the included studies (17) (**Appendix C).** For reference standard, there is currently no gold standard for the diagnosis of DILD however the following diagnostic criteria have been suggested:

1. A history of drug exposure with correct identification of the drug, its dose, and its duration of administration.
2. Clinical, imaging and histopathological patterns which are consistent with earlier observations with the same drug.
3. Exclusion of other lung disease.
4. Improvement following discontinuation of the suspected drug
5. Recurrence of symptoms on re-challenge (although not routinely recommended). (18)

For included studies the reference standard domain was rated as a low risk of bias if each of the first 3 criteria were satisfied.

Data extraction forms (**Appendix D)** were completed independently by each author for each eligible study. Where data was missing, we attempted to contact the authors of the studies for further information.

For studies where sensitivity and specificity were provided only, this data was entered into RevMan5 (19) and forest plots were generated based on the number of participants provided in the study.

**Data synthesis**

Where possible, we calculated sensitivity and specificity with corresponding 95% CIs for individual studies using a threshold of ≥15% decline from baseline lung function as reported in the studies. The results of individual studies were entered into RevMan5 (19) and displayed in a paired forest plot. Pooled analysis within each drug sub-group was planned if there was sufficient data and the data was considered sufficiently clinically and statistically homogeneous. The findings of studies which did not provide sufficient data to calculate sensitivity and specificity were summarised in table format. Changes in lung function during treatment were summarised in table format for studies which reported this outcome.

## Results

We identified 4,065 citations (after the removal of duplicates) and selected 195 articles for full text review after screening abstracts. We included 42 studies in the review which comprised of 3,489 participants (**Figure 1**). Only 19 studies provided sufficient quantitative data that could be extracted for our primary outcome. A further 10 studies provided quantitative data suitable for extraction for our secondary outcome and there were 13 studies which did not provide sufficient quantitative data for extraction for either outcome. Amiodarone and Bleomycin were the two most commonly studied drugs. Characteristics of included studies for which quantitative primary outcome data could be extracted are summarised in **Table 1**. Characteristics of studies which were included in our review but for which we could not extract sufficient primary outcome data are summarised in **Appendix E**.

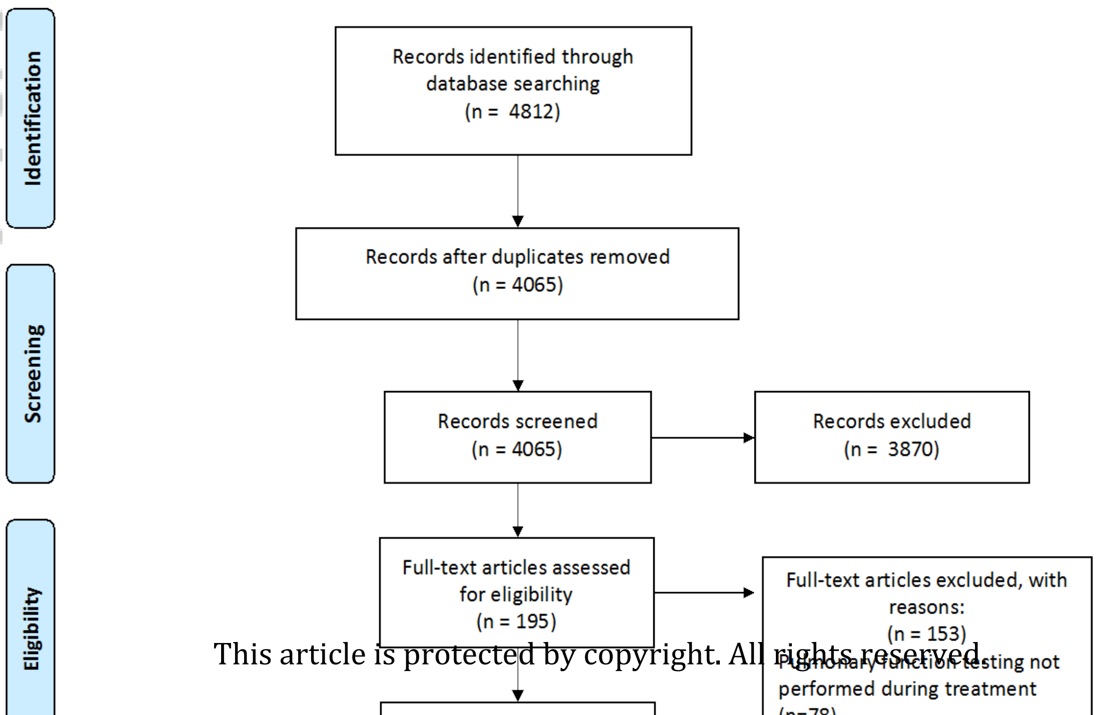


**Figure 1:** PRISMA Flow Diagram

| Author, date, geographical region and study methodology[1] | Drug being evaluated | Condition(s) being treated | Number of participants, age and gender[2] | Types of PFT[3] <br>• Frequency of PFT | Threshold applied for DLCO | Duration of follow up <br>• Loss to follow up[4] |
|---|---|---|---|---|---|---|
| **Chemotherapeutic agents** | | | | | | |
| Castro (20)<br>1996<br>USA<br>Prospective | **Mitomycin,**<br>Vinblastine,<br>Cisplatin | Squamous cell lung carcinoma | 133 patients<br>Mean age of 63 (range 37-85)<br>106 males and 27 females | DLCO, FEV1, FVC and TLC<br>• Baseline and 16 weeks after | 20% | 12 months<br>• 40 patients had both measurements of DLCO |
| Comis (7)<br>1979<br>USA<br>Prospective | **Bleomycin**,<br>Vinblastine,<br>cis-diamminedichloro platinum | Testicular carcinoma and primary mediastinal germ cell tumours | 11 patients<br>Mean age was 25 years (range 19-47)<br>11 Male | DLCO, FEV1 and FVC<br>• Baseline and 1-3 week intervals thereafter | 60% | 7 months<br>• 1 patient |
| Duprat (21)<br>1987<br>Canada<br>Prospective | **Bleomycin,**<br>Cisplatinum,<br>Vincristine | Esophageal carcinoma | 16 patients | DLCO<br>• After 1st and 2nd cycle of chemotherapy | 25% | Not described |
| Friedberg 2003 (22)<br>USA<br>Prospective | **Bleomycin,**<br>Doxorubicin,<br>Vinblastine,<br>Gemcitabine | Hodgkin's disease | 12 patients<br>Median age of 34 years (range, 24-60 years)<br>7 females and 5 males | DLCO<br>• Baseline, after cycle 3 and at the end of therapy | 15% | 12 months |
| Houghton 2010 (16) | **Bleomycin,**<br>Etoposide,<br>Cisplatin | Advanced germ cell tumours | 30 patients<br>Median age of 28 (range 19-39) | DLCO, FVC<br>• Fortnightly | 25% | Not described |
| Lewis (23)<br>1980<br>USA<br>Prospective<br>Consecutive sampling | **Bleomycin,**<br>Cisplatin,<br>Vinblastine | Squamous cell carcinoma of the tonsil, adenocarcinoma of the kidney | 23 patients | DLCO, FEV1 and FVC<br>• Not described | 20% | Not described |

| | | | | | | |
|---|---|---|---|---|---|---|
| Luuresma (24)<br>1983<br>Prospective | **Bleomycin,**<br>Vinblastine,<br>cis-Diammine-dichloroplatinum | Disseminated testicular non-seminomatous tumours | 18 patients<br>Mean age of 32.6 years | DLCO, VC<br>• Before chemotherapy, fixed intervals of 3 weeks until end of treatment then 2 and 4 months after treatment | 20% | 4 months |

[1] Place of study, study methodology and sampling method stated unless not available
[2]Age, gender stated unless not available
[3]DLCO corrected for Haemoglobin unless stated otherwise
[4]Loss to follow up described unless not available
**Abbreviations:** Diffusing capacity of the lungs for carbon monoxide (DLCO), FEV1 (Forced expiratory volume in 1 second), FVC (Forced vital capacity), VC (vital capacity), TLC (Total lung capacity)

| Author, date, geographical region and study methodology[1] | Drug being evaluated | Condition(s) being treated | Number of participants, age and gender[2] | Types of PFT[3]<br>• **Frequency of PFT** | Threshold applied for DLCO | Duration of follow up<br>• **Loss to follow up[4]** |
|---|---|---|---|---|---|---|
| McKeage 1990 (10)<br>New Zealand<br>Retrospective<br>Consecutive sampling | **Bleomycin**, various agents used but not stated | Testicular carcinoma and malignant lymphoma | 81 patients<br>Mean age 39.4 (range 17-69)<br>59 Male 22 Female | DLCO<br>• Not described | 35% | Not described |
| Roncolato 2016 (25)<br>Australia and New Zealand<br>Prospective | **Bleomycin,** Etoposide, Cisplatin | Metastatic germ cell tumours arising in the testis, retroperitoneum, mediastinum or ovary | 43 patients<br>Median age of 28 years (range 18-39) | DLCO, FVC<br>• Baseline, fortnightly thereafter | 25% | 24 months |
| Wolkowicz 1992 (26)<br>Canada<br>Retrospective<br>Consecutive sampling | **Bleomycin,** Vinblastine, cis-diamminedichloro platinum | Non-seminomatous testicular carcinoma | 59 patients<br>Mean age of 27.7 years (range 16-44)<br>59 Male | DLCO, VC, TLC and FEV1<br>• Baseline and 2 months after treatment began | 15% | Follow up not described |
| Adams (27)<br>1988<br>USA<br>Prospective<br>Consecutive sampling | Amiodarone (400-600mg/day) | Refractory arrhythmia | 33 patients<br>Mean age of 60.3 years<br>27 Male and 6 Female | DLCO, FEV1, FVC and TLC<br>• 3-6 monthly | 15% | 22 months<br>• 18 patients |
| Gleadhill (28)<br>1989<br>USA<br>Prospective<br>Consecutive sampling | Amiodarone (136-512mg/day) | Recurrent arrhythmia | 91 patients<br>Mean age of 61.7 years<br>69 Male and 22 Female | DLCO, FVC and FEV1<br>• Baseline, 3,6,12,18 and 24 months<br>• DLCO was not corrected for Hb | 20% | 24 months<br>• 2 patients |

| | | | | | | |
|---|---|---|---|---|---|---|
| Kudenchuk 1984 (29)<br>USA<br>Prospective<br>Consecutive sampling | Amiodarone<br>(100-<br>1200mg/day) | Refractory arrhythmia | 100 patients<br>Mean age of 59 years<br>79 Male and 21 Female | DLCO, FEV1, FVC and TLC<br>• Every 8 months | 15% | 262 days<br>• 31 patients |

[1] Place of study, study methodology and sampling method stated unless not available
[2]Age, gender stated unless not available
[3]DLCO corrected for Haemoglobin unless stated otherwise
[4]Loss to follow up described unless not available
**Abbreviations:** Diffusing capacity of the lungs for carbon monoxide (DLCO), FEV1 (Forced expiratory volume in 1 second), FVC (Forced vital capacity), VC (vital capacity), TLC (Total lung capacity)

| Author, date, geographical region and study methodology[1] | Drug being evaluated | Condition(s) being treated | Number of participants, age and gender[2] | Types of PFT[3]<br>• Frequency of PFT | Threshold applied for DLCO | Duration of follow up<br>• Loss to follow up[4] |
|---|---|---|---|---|---|---|
| Magro (9)<br>1988<br>USA<br>Prospective<br>Consecutive sampling | Amiodarone (200-600mg/day) | Arrhythmias | 89 patients<br>Age ranged from 48 to 70 years | DLCO, FEV1, FVC, FEV1/FVC and TLC<br>• Every 3 months for first year, 6 months thereafter<br>• DLCO not corrected for Hb | 15% | 54 months<br>• 12 patients |
| Ohar (11)<br>1989<br>USA<br>Prospective | Amiodarone (400-800mg/day) | Ventricular arrhythmias unresponsive to conventional therapy | 189 patients<br>127 Male and 62 Female | DLCO, FEV1, FVC, FEV1/FVC and TLC<br>• Every 6 months<br>• DLCO not corrected for Hb | 15% | 43 months<br>• 4 patients |
| Suppli Ulrik 1992 (30)<br>Denmark<br>Prospective<br>Consecutive sampling | Amiodarone (400mg/day) | Refractory tachyarrhythmia | 24 patients<br>Mean age of 53.8 years<br>10 Male and 14 Female | DLCO, FEV1, FVC and FEV1/FVC<br>• Every 12 months<br>• DLCO not corrected for Hb | 20% | 48 months<br>• 10 patients |
| Yamada 2007 (31)<br>Japan<br>Retrospective<br>Consecutive sampling | Amiodarone (50-200mg/day) | Ventricular or supraventricular tachyarrhythmia with structural heart disease or heart failure | 500 patients<br>Mean age of 53 years<br>409 Male and 91 Female | DLCO<br>• Ever 3 months for the first year and 3-6 months thereafter<br>• DLCO not corrected for Hb | 20% | 48 months<br>• 84 patients not included in analysis |
| Cottin (8)<br>1996<br>France<br>Prospective<br>Consecutive sampling | Methotrexate (7.5-15mg/week) | Rheumatoid Arthritis<br>Psoriatic Rheumatism<br>Systemic Lupus Erythematosus | 124 patients<br>Mean age of 56 years<br>36 Male and 88 Female | DLCO, FEV1, FVC and TLC<br>• Baseline, 3 months and 6 monthly intervals thereafter | 20% | 85 months<br>• 2 patients |
| Franzen (32)<br>2016<br>Switzerland<br>Prospective<br>Consecutive sampling | Rituximab (1000mg) | Rheumatoid Arthritis | 33 patients<br>Mean age of 58.8 years<br>9 Male and 24 Female | DLCO, FEV1 and FVC<br>• 2,4,8,26 weeks<br>• DLCO not corrected for Hb | 15% | 26 weeks<br>• 6 patients |

[1] Place of study, study methodology and sampling method stated unless not available
[2] Age, gender stated unless not available
[3] DLCO corrected for Haemoglobin unless stated otherwise
[4] Loss to follow up described unless not available
**Abbreviations:** Diffusing capacity of the lungs for carbon monoxide (DLCO), FEV1 (Forced expiratory volume in 1 second), FVC (Forced vital capacity), VC (vital capacity), TLC (Total lung capacity)

**Table 1:** Characteristics of studies included for primary outcome

**Risk of bias analysis**

The quality of included studies was relatively poor, particularly for the domains of patient selection and reference standard. More than 50% of studies were rated as being at an unclear risk or high risk of bias for patient selection and more than 50% of included studies were at high risk of bias for the reference standard domain. Many studies provided limited information on the diagnostic criteria used to establish a diagnosis of DILD. **Figure 2** displays the risk of bias of the individual studies for the 4 quality domains assessed. **Appendix F** summarises the risk of bias analysis for all studies and for individual studies.

## Primary outcome

*Diagnostic accuracy of DLCO*

**Sensitivity and specificity of DLCO during treatment.**

We identified 19 studies (with a total of 1,494 participants) that included monitoring of DLCO during drug treatment for which we were able to extract sensitivity and specificity. There were 9 studies of monitoring treatment with Bleomycin, 7 monitoring Amiodarone therapy, 1 monitoring Mitomycin C, 1 monitoring low dose Methotrexate and 1 monitoring Rituximab. 4 of the studies included did not have any cases of DILD thus sensitivity was not estimable.

**Figure 3** shows the paired forest plot estimates for sensitivity and specificity of DLCO monitoring during Bleomycin treatment. The point estimate for sensitivity varied from 0% (at a threshold of 15% and 20%) to 100% (at a threshold of 25%). However, the confidence intervals are wide reflecting significant uncertainty for the true estimate of sensitivity due to the small number of participants and the low incidence of pulmonary toxicity. Furthermore, there is no relationship between sensitivity and increasing threshold for decline in DLCO as we would expect sensitivity to increase with increasing threshold. The point estimate for specificity varied from 66% (at a threshold of 15%) to 89% (at a threshold of 60%). However, in one study, the point estimate for specificity was 0%.

**Figure 3** also shows the paired forest plots for sensitivity and specificity for DLCO monitoring during treatment with Amiodarone. The point estimate for sensitivity varied from 20% to 100% (at a threshold of 15%). The point estimate for specificity varied from 54% (at a threshold of 15%) to 83% (at a threshold of 20%)

For studies that utilised Mitomycin C, Methotrexate and Rituximab, the data has been summarised in Appendix G.

**Sensitivity and specificity of spirometry measurements.**

We identified 3 studies that used spirometry measurements during the treatment of patients with Bleomycin and 2 studies with Amiodarone as displayed in **Figure 4**.



## Bleomycin

### >15% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Friedburg 2003 | 0 | 7 | 5 | 0 | 0.00 [0.00, 0.52] | 0.00 [0.00, 0.41] |
| Wolkowicz 1992 | 4 | 17 | 5 | 33 | 0.44 [0.14, 0.79] | 0.66 [0.51, 0.79] |

### >20% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Lewis 1980 | 0 | 7 | 1 | 15 | 0.00 [0.00, 0.97] | 0.68 [0.45, 0.86] |
| Luuresma 1983 | 0 | 2 | 3 | 13 | 0.00 [0.00, 0.71] | 0.87 [0.60, 0.98] |

### >25% decline

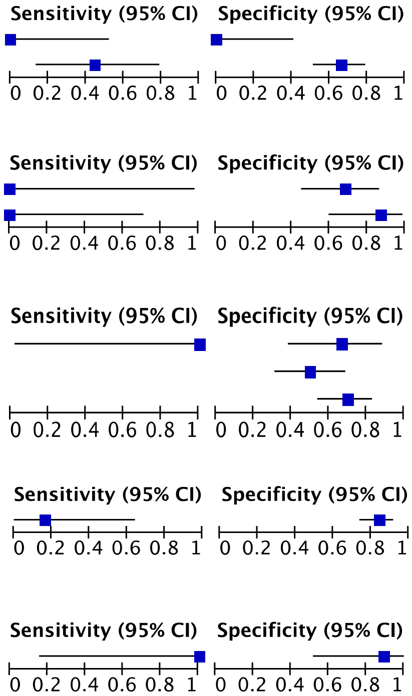
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Duprat 1987 | 1 | 5 | 0 | 10 | 1.00 [0.03, 1.00] | 0.67 [0.38, 0.88] |
| Houghton 2010 | 0 | 15 | 0 | 15 | Not estimable | 0.50 [0.31, 0.69] |
| Roncolato 2016 | 0 | 13 | 0 | 30 | Not estimable | 0.70 [0.54, 0.83] |

### >35% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| McKeage et al – 1990 | 1 | 12 | 5 | 63 | 0.17 [0.00, 0.64] | 0.84 [0.74, 0.91] |

### >60% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Comis 1979 | 2 | 1 | 0 | 8 | 1.00 [0.16, 1.00] | 0.89 [0.52, 1.00] |




## Amiodarone

### >15% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Adams et al. 1988 | 6 | 11 | 3 | 13 | 0.67 [0.30, 0.93] | 0.54 [0.33, 0.74] |
| Kudenchuck et al. 1984 | 7 | 12 | 2 | 48 | 0.78 [0.40, 0.97] | 0.80 [0.68, 0.89] |
| Magro et al. 1988 | 8 | 22 | 0 | 37 | 1.00 [0.63, 1.00] | 0.63 [0.49, 0.75] |
| Ohar et al. 1989 | 2 | 12 | 8 | 19 | 0.20 [0.03, 0.56] | 0.61 [0.42, 0.78] |
| Yamada et al. 2007 | 27 | 115 | 13 | 256 | 0.68 [0.51, 0.81] | 0.69 [0.64, 0.74] |

### >20% decline

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Gleadhill et al. 1989 | 4 | 15 | 0 | 72 | 1.00 [0.40, 1.00] | 0.83 [0.73, 0.90] |
| Suppli Ulrik et al. 1992 | 0 | 7 | 0 | 17 | Not estimable | 0.71 [0.49, 0.87] |
| Yamada et al. 2007 | 23 | 98 | 16 | 279 | 0.59 [0.42, 0.74] | 0.74 [0.69, 0.78] |

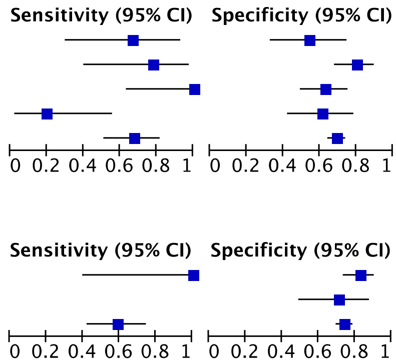
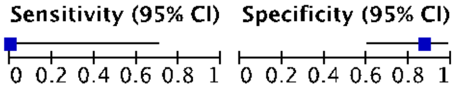



**Figure 3**: Top - paired forest plots for sensitivity and specificity of DLCO for monitoring for detection of DILD in patients receiving Bleomycin
Bottom – paired forest plots for sensitivity and specificity of DLCO for monitoring for detection of DILD in patients receiving Amiodarone

# Bleomycin

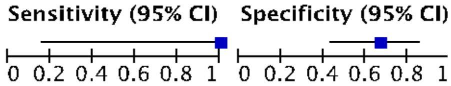
## >20% decline in VC

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Luuresma 1983 | 0 | 2 | 3 | 13 | 0.00 [0.00, 0.71] | 0.87 [0.60, 0.98] |



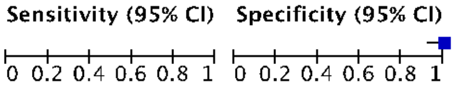

## >20% decline in FVC

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Lewis 1980 | 2 | 7 | 0 | 14 | 1.00 [0.16, 1.00] | 0.67 [0.43, 0.85] |




## >25% decline in FVC

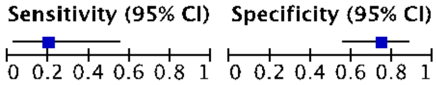
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Roncolato 2016 | 0 | 0 | 0 | 43 | Not estimable | 1.00 [0.92, 1.00] |




# Amiodarone

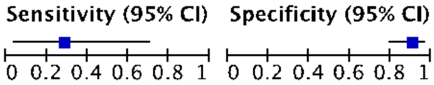
## >15% decline in FVC

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Ohar et al. 1989 | 2 | 8 | 8 | 23 | 0.20 [0.03, 0.56] | 0.74 [0.55, 0.88] |




## >15% decline in TLC

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kudenchuck et al. 1984 | 2 | 5 | 5 | 47 | 0.29 [0.04, 0.71] | 0.90 [0.79, 0.97] |




## >15% decline in FEV1/FVC

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Ohar et al. 1989 | 0 | 5 | 10 | 16 | 0.00 [0.00, 0.31] | 0.76 [0.53, 0.92] |

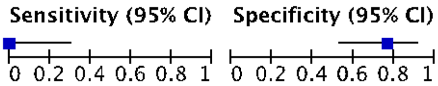



**Figure 4**: Top - paired forest plots for sensitivity and specificity of spirometry measurements for monitoring for detection of DILD in patients receiving Bleomycin
Bottom – paired forest plots for sensitivity and specificity of spirometry measurements for monitoring for detection of DILD in patients receiving Amiodarone

**Qualitative summary of included studies with insufficient quantitative data**

The findings of 23 studies included in the review for which we were unable to extract sufficient quantitative data are outlined in **Appendix H**. Ten of these studies included data which was extracted for our secondary outcome. Of note, several studies reported cases of DILD which were not detected prior to clinical presentation by monitoring with DLCO or spirometry (33, 34)

## Secondary outcome

**Change in DLCO during treatment**

Eighteen studies reported change in DLCO and/or spirometry during treatment. In general, DLCO was found to decrease from baseline during the treatment periods. See **Appendix I**.

**Heterogeneity**

We elected not to perform a meta-analysis due to the lack of homogeneity between included studies, even within drug and test threshold subgroups. We did not perform tests of statistical heterogeneity; however, it should be noted that within some drug and threshold subgroups confidence intervals for some studies did not overlap suggesting significant heterogeneity.

## Discussion

To our knowledge, this is the first systematic literature review to examine the diagnostic test accuracy of PFTs for the early detection of DILD. The review included 42 studies and synthesised a substantial body of literature, however conclusions are limited by the quality and size of the studies included. The most frequently studied drugs were Bleomycin and Amiodarone. The largest study included in the review was undertaken in patients being treated with Amiodarone, in which monitoring with DLCO (threshold change >15%) was found to have a sensitivity of 0.68 (95% CI: 0.51 to 0.81) and specificity of 0.69 (95% CI: 0.64 to 0.74). Specificity improved only marginally when the threshold was increased to > 20%; 0.74 (95% CI: 0.69 to 0.78). This level of test accuracy suggests that monitoring with DLCO periodically is not an ideal screening test for the early detection of DILD, however it is not clear whether the experience with Amiodarone can be extrapolated to other drugs due to differences in mechanism of action, pharmacokinetics and the heterogeneity of DILD (35)

Due to the relatively low prevalence of pulmonary toxicity, the small sample size of included studies and the early stopping rules for Bleomycin based on PFTs employed in some studies, there remains considerable uncertainty about the sensitivity of spirometry and DLCO for the early detection of DILD during treatment with Bleomycin. Despite a specific search of RCTs which included Bleomycin in any treatment arm, we did not identify any studies that had examined the impact of PFT monitoring on mortality related to DILD or disease related mortality. We also did not consider anaesthetic-related or potential long-term pulmonary toxicity from Bleomycin. Of note, at least 3 studies included in the review reported fatal cases of pulmonary toxicity with Bleomycin despite prospective monitoring with pulmonary function tests. (33, 34, 36) The U.S Food and Drug Administration (FDA) currently recommends regular chest x-rays and that DLCO is monitored monthly during treatment with Bleomycin and discontinuation of treatment should occur when DLCO falls below 30- 35% of the pre-treatment value (13). In some institutions, it is common practice to cease treatment if DLCO values drop more than 15-25% from pre-treatment values (25). In our review, at all thresholds reported, in the studies for which we were able to extract primary outcome data, the specificity of DLCO was less than 90%. This finding supports concerns that monitoring may lead to the early cessation of Bleomycin with the potential to reduce the chance of cure (or increase toxicity by substituting additional other chemotherapy) in the management of germ cell tumours (37). At present, we do not have sufficient understanding of the relationship between mild abnormalities of lung function in asymptomatic individuals and the development of severe DILD. We also lack an understanding of the time course of these changes; for example, a 35% fall in DLCO after 2 doses of Bleomycin is likely to be more significant than the same fall after 10 doses. Recent studies have reported that more than 60% of patients with germ cell tumours being treated with Bleomycin, Etoposide and Cisplatin have evidence of interstitial lung abnormalities on computed tomography imaging performed during, or at the end of treatment, and in the majority of cases these are asymptomatic and resolved spontaneously (6). Given the poor sensitivity and specificity, if monitoring of pulmonary function is undertaken during treatment with Bleomycin, a decline in DLCO adjusted for Hb of ≥30% should prompt further clinical evaluation and investigation (such as high-resolution CT chest) before decisions are made about omitting or reducing the dose of Bleomycin. Furthermore, the development of cough or breathlessness or new pulmonary physical signs should be further investigated promptly even when lung function remains stable. In addition, patients should be encouraged to report symptoms early, regardless of whether PFT monitoring is undertaken.

The review was conducted using a published protocol and following appropriate methodology for the conduct of systematic reviews (38) (PROSPERO ID: CRD42018090149), however there are limitations. Most of the included studies were small and many had a high or unclear risk of bias. The reporting of the results of many of the included studies was poor and therefore any quantitative synthesis would have excluded a significant proportion of the available literature. Our primary outcomes, sensitivity and specificity, are useful measures of test characteristics, but have limitations, particularly given the lack of a true gold standard for diagnosis of DILD. Of note, more than half of the studies included in the review used an inadequate standard for diagnosis of DILD or did not adequately describe the standard used. It is recommended that criteria used to evaluate screening tests should be applied to monitoring for adverse drug reactions, however it has yet to be established whether early detection of asymptomatic of DILD will improve mortality and morbidity. (39).

Within and between the different drug subgroups of included studies there were multiple sources of clinical heterogeneity including differences in patient populations, types of co-interventions, frequency of pulmonary function testing and the thresholds applied to test results. In addition, in some studies sampling methods for patient selection were unclear and length of follow up was variable. Furthermore, in some Bleomycin studies, treatment was ceased when lung function declined to a certain threshold (which varied between studies). These factors, may in part, account for the wide variation in point estimates and confidence intervals noted between included studies for the sensitivity and specificity for DLCO monitoring (Figures 3 and 4).

Many of the studies included in this review were conducted prior to the era of precision medicine. Tyrosine kinase inhibitors (TKIs) and immune-checkpoint inhibitors (ICIs) are relatively common causes of DILD in modern oncological practice (40). The risk of pulmonary toxicity with immunotherapy such as ICIs is relatively high compared to standard chemotherapy regimens. (41) In a recent prospective trial utilising PFT for screening with the use of Ipilimumab in melanoma patients, PFT yielded a poor sensitivity and specificity in identifying clinically significant disease. However, changes in DLCO could be a marker of subclinical disease and routine PFT measurements during treatment may aid risk stratification in screening for pulmonary toxicity.(42) Clinical trials with newer agents provide an opportunity to evaluate alternative approaches to early detection such as biomarkers however, preliminary research with currently available biomarkers has been relatively disappointing (43). Emerging evidence for new approaches to testing and monitoring for the early detection of DILD would ideally be evaluated using the same broadly accepted framework used to evaluate screening interventions (39). The findings of this review highlight important methodological issues that could be improved in future research in this field.

## Conclusion

This systematic review did not identify evidence supporting the routine use of PFTs for the early detection of DILD in patients being treated with Amiodarone or Bleomycin. Due to the methodological limitations, the relatively small number of participants and the low prevalence of DILD in the included studies, there remains significant uncertainty about the sensitivity of pulmonary function tests to detect DILD early in patients being treated with pulmonary toxic drugs. Further high-quality research is needed to inform clinical practice in the early detection of DILD in the future. Clinicians choosing to utilise PFTs must be aware of the limitations of this approach, particularly the poor sensitivity and specificity in identifying clinically significant DILD.

## References


1. Thomeer M, Costabel U, Rizzato G, Poletti V, Demedts M. Comparison of registries of interstitial lung diseases in three European countries. European Respiratory Journal. 2001;18(32 suppl):114s-8s.
2. Wu T-Y, Jen M-H, Bottle A, Molokhia M, Aylin P, Bell D, et al. Ten-year trends in hospital admissions for adverse drug reactions in England 1999–2009. Journal of the Royal Society of Medicine. 2010;103(6):239-50.
3. Schwaiblmair M, Berghaus T, Haeckel T, Wagner T, Von Scheidt W. Amiodarone-induced pulmonary toxicity: An under-recognized and severe adverse effect? Clinical Research in Cardiology. 2010;99(11):693-700.
4. Sleijfer S. Bleomycin-induced pneumonitis. Chest. 2001;120(2):617-24.
5. Cosgrove GP, Bianchi P, Danese S, Lederer DJ. Barriers to timely diagnosis of interstitial lung disease in the real world: the INTENSITY survey. BMC Pulmonary Medicine. 2018;18(1):9.
6. Shamash J, Sarker SJ, Huddart R, Harland S, Joffe JK, Mazhar D, et al. A randomized phase III study of 72 h infusional versus bolus bleomycin in BEP (bleomycin, etoposide and cisplatin) chemotherapy to treat IGCCCG good prognosis metastatic germ cell tumours (TE-3). Annals of Oncology. 2017;28(6):1333-8.

7. Comis RL, Kuppinger MS, Ginsberg SJ, Crooke ST, Gilbert R, Auchincloss JH, et al. Role of single-breath carbon monoxide-diffusing capacity in monitoring the pulmonary effects of bleomycin in germ cell tumor patients. Cancer Research. 1979;39(12):5076-80.
8. Cottin V, Tebib J, Massonnet B, Souquet PJ, Bernard JP. Pulmonary function in patients receiving long-term low-dose methotrexate. Chest. 1996;109(4):933-8.
9. Magro SA, Lawrence EC, Wheeler SH, Krafchek J, Lin HT, Wyndham CR. Amiodarone pulmonary toxicity: prospective evaluation of serial pulmonary function tests. Journal of the American College of Cardiology. 1988;12(3):781-8.
10. McKeage MJ, Evans BD, Atkinson C, Perez D, Forgeson GV, Dady PJ. Carbon monoxide diffusing capacity is a poor predictor of clinically significant bleomycin lung. Journal of Clinical Oncology. 1990;8(5):779-83.
11. Ohar JA, Jackson F, Jr., Redd RM, Evans GR, Bedrossian CW. Usefulness of serial pulmonary function testing as an indicator of amiodarone toxicity. American Journal of Cardiology. 1989;64(19):1322-6.
12. Pavy S, Constantin A, Pham T, Gossec L, Maillefert JF, Cantagrel A, et al. Methotrexate therapy for rheumatoid arthritis: clinical practice guidelines based on published evidence and expert opinion. Joint Bone Spine. 2006;73(4):388-95.
13. Administration USFaD. Blenoxane (Bleomycin Sulfate for injection) 2010 [Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2010/050443s036lbl.pdf.
14. Cetti RJ, Venn S, Woodhouse CR. The risks of long-term nitrofurantoin prophylaxis in patients with recurrent urinary tract infection: a recent medico-legal case. BJU international. 2009;103(5):567-9.
15. Cox G, Johnson J, Kinnear WJ, Johnston ID. Amiodarone and the lung: wide variations in clinical practice. Respiratory Medicine. 2000;94(11):1130-1.
16. Houghton B, Grimison PS, Toner GC, Thomson DB, Friedlander M, Gebski V, et al. Bleomycin dosing for germ cell tumours - Are we too cautious? Asia-Pacific Journal of Clinical Oncology. 2010;2):46.
17. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
18. Schwaiblmair M, Behr W, Haeckel T, Markl B, Foerg W, Berghaus T. Drug induced interstitial lung disease. Open Respiratory Medicine Journal. 2012;6(1):63-74.
19. Review Manager (RevMan). Version 53. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2014.
20. Castro M, Veeder MH, Mailliard JA, Tazelaar HD, Jett JR. A prospective study of pulmonary function in patients receiving mitomycin. Chest. 1996;109(4):939-44.
21. Duprat G, Jr., Chalaoui J, Sylvestre J, Robidoux A, Duranceau A. Pulmonary complications of multimodality therapy for esophageal carcinoma. Canadian Association of Radiologists Journal. 1987;38(1):27-31.
22. Friedberg JW, Neuberg D, Kim H, Miyata S, McCauley M, Fisher DC, et al. Gemcitabine added to doxorubicin, bleomycin, and vinblastine for the treatment of De Novo Hodgkin disease: Unacceptable acute pulmonary toxicity. Cancer. 2003;98(5):978-82.
23. Lewis BM, Izbicki R. Routine pulmonary function tests during bleomycin therapy. Tests may be ineffective and potentially misleading. Journal of the American Medical Association. 1980;243(4):347-51.
24. Luursema PB, Star Kroesen MA, Van Der Mark Th W. Bleomycin-induced changes in the carbon monoxide transfer factor of the lungs and its components. American Review of Respiratory Disease. 1983;128(5):880-3.

25. Roncolato FT, Chatfield M, Houghton B, Toner G, Stockler M, Thomson D, et al. The effect of pulmonary function testing on bleomycin dosing in germ cell tumours. Internal Medicine Journal. 2016;46(8):893-8.
26. Wolkowicz J, Sturgeon J, Rawji M, Chan CK. Bleomycin-induced pulmonary function abnormalities. Chest. 1992;101(1):97-101.
27. Adams GD, Kehoe R, Lesch M, Glassroth J. Amiodarone-induced pneumonitis: Assessment of risk factors and possible risk reduction. Chest. 1988;93(2):254-63.
28. Gleadhill IC, Wise RA, Schonfeld SA, Scott PP, Guarnieri T, Levine JH, et al. Serial lung function testing in patients treated with amiodarone: a prospective study. American Journal of Medicine. 1989;86(1):4-10.
29. Kudenchuk PJ, Pierson DJ, Greene HL, Graham EL, Sears GK, Trobaugh GB. Prospective evaluation of amiodarone pulmonary toxicity. Chest. 1984;86(4):541-8.
30. Suppli Ulrik C, Backer V, Aldershvile J, Pietersen AH. Serial pulmonary function tests in patients treated with low-dose amiodarone. American Heart Journal. 1992;123(6):1550-4.
31. Yamada Y, Shiga T, Matsuda N, Hagiwara N, Kasanuki H. Incidence and predictors of pulmonary toxicity in japanese patients receiving low-dose amiodarone. Circulation Journal. 2007;71(10):1610-6.
32. Franzen D, Ciurea A, Bratton DJ, Clarenbach CF, Latshang TD, Russi EW, et al. Effect of rituximab on pulmonary function in patients with rheumatoid arthritis. Pulmonary Pharmacology and Therapeutics. 2016;37:24-9.
33. Bell MR, Meredith DJ, Gill PG. Role of carbon monoxide diffusing capacity in the early detection of major bleomycin-induced pulmonary toxicity. Australian and New Zealand Journal of Medicine. 1985;15(2):235-40.
34. Wheeler RH, Liepman MK, Baker SR, Earhart RH, Bull FE, Ensminger WD. Bleomycin, vincristine, and mitomycin C with or without methotrexate in the treatment of squamous cell carcinoma. Cancer Treatment Reports. 1980;64(8-9):943-9.
35. Adams PC, Gibson GJ, Morley AR, Wright AJ, Corris PA, Reid DS, et al. Amiodarone pulmonary toxicity: clinical and subclinical features. Quarterly Journal of Medicine. 1986;59(229):449-71.
36. Levi JA, Raghavan D, Harvey V, Thompson D, Sandeman T, Gill G, et al. The importance of bleomycin in combination chemotherapy for good-prognosis germ cell carcinoma. Journal of Clinical Oncology. 1993;11(7):1300-5.
37. O'Sullivan JM, Huddart RA, Norman AR, Nicholls J, Dearnaley DP, Horwich A. Predicting the risk of bleomycin lung toxicity in patients with germ-cell tumours. Annals of Oncology. 2003;14(1):91-6.
38. Moher D, Liberati A, Tetzlaff J, Altman DG, and the PG. Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. Annals of Internal Medicine. 2009;151(4):264-9.
39. Pirmohamed M, Ferner RE. Monitoring drug treatment. British Medical Journal Publishing Group; 2003.
40. Nishino M, Hatabu H, Hodi FS, Ramaiya NH. Drug-Related Pneumonitis in the Era of Precision Cancer Therapy. JCO Precision Oncology. 2017;1(1):1-12.
41. Rahouma M, Baudo M, Yahia M, Kamel M, Gray KD, Elmously A, et al. Pneumonitis as a complication of immune system targeting drugs?-a meta-analysis of anti-PD/PD-L1 immunotherapy randomized clinical trials. J Thorac Dis. 2019;11(2):521-34.
42. Franzen D, Schad K, Kowalski B, Clarenbach CF, Stupp R, Dummer R, et al. Ipilimumab and early signs of pulmonary toxicity in patients with metastastic melanoma: a prospective observational study. Cancer immunology, immunotherapy : CII. 2018;67(1):127-34.

43. Ohnishi H, Yokoyama A, Yasuhara Y, Watanabe A, Naka T, Hamada H, et al. Circulating KL-6 levels in patients with drug induced pneumonitis. Thorax. 2003;58(10):872-5.