Reference values for maximal inspiratory pressure: A systematic review

Isabela MB Sclauser Pessoa PT PhD1, Verônica Franco Parreira PT PhD2, Guilherme AF Fregonezi PT PhD3, A William Sheel BPE PhD4, Frank Chung PT MSc5, W Darlene Reid PT PhD6,7,8

BACKGROUND: Maximal inspiratory pressure (MIP) is the most commonly used measure to evaluate inspiratory muscle strength. Normative values for MIP vary significantly among studies, which may reflect differences in participant demographics and technique of MIP measurement.

OBJECTIVE: To perform a systematic review with meta-analyses to synthesize MIP values that represent healthy adults.

METHODS: A systematic literature search was conducted using Medline, Embase, Cochrane, Cumulative Index to Nursing and Allied Health (CINAHL) and Sport Discus databases. Two reviewers identified and selected articles, and abstracted data. Methodological quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool. A random-effects model was used to calculate overall means and 95% CIs.

RESULTS: Of 22 included articles, MIP data were synthesized according to age group and sex from six reports (n=840) in the meta-analyses. The mean QUADAS score was 3.5 of 7. The age range was between 18 and 83 years (426 men, 414 women). MIP began to decrease with age in the 40 to 60 years age range and continued to fall progressively with age. For the same age group, men tended to have higher MIPs than women. Sensitivity analysis of withdrawing studies from the meta-analysis identified one study that contributed more to heterogeneity in some age groups.

DISCUSSION: MIP was higher in men and decreased with age, which was initially apparent in middle age. Several characteristics of participants and MIP technique influence values in healthy individuals.

CONCLUSIONS: The present meta-analysis provides normative MIP values that are reflective of a large sample (n=840) and likely represents the broadest representation of participant characteristics compared with previous reports of normative data.

Key Words: Adults; Muscle strength; Reference values; Respiratory muscles

Les valeurs de référence de la pression inspiratoire maximale : une analyse systématique

HISTORIQUE : La pression inspiratoire maximale (PIM) est la mesure la plus utilisée pour évaluer la force des muscles inspiratoires. Les valeurs normatives de PIM varient considérablement selon les études, ce qui reflète peut-être des différences dans la démographie des participants et dans la technique pour mesurer la PIM.

OBJECTIF : Effectuer une analyse systématique au moyen de méta-analyses pour synthétiser les valeurs de PIM qui représentent des adultes en santé.

MÉTHODOLOGIE : En vue de l’analyse systématique, les chercheurs ont fouillé dans les bases de données Medline, Embase, Cochrane, Cumulative Index to Nursing and Allied Health (CINAHL) et Sport Discus. Deux analystes ont déterminé et sélectionné les articles et synthétisé les données. Ils ont évalué la qualité méthodologique au moyen duoutil QUADAS d’études d’évaluation de la qualité et d’exactitude diagnostique. Ils ont utilisé un modèle à effets aléatoires pour calculer les moyennes globales et les IC à 95 %.

RÉSULTATS : Les données de PIM ont été synthétisées selon le groupe d’âge et le sexe dans six des 22 articles inclus (n=840) dans la méta-analyse. L’indice QUADAS moyen était de 3,5 sur 7. La forcechette d’âge se situait entre 18 et 83 ans (426 hommes, 414 femmes). La PIM maximale s’est mise à diminuer avec le vieillissement dans la fourchette des 40 à 60 ans et a continué à dégresser progressivement avec l’âge. Dans le même groupe d’âge, les hommes avaient tendance à présenter des PIM plus élevées que les femmes. Les analyses de sensibilité pour retirer des études de la méta-analyse n’ont pas permis de repérer une seule étude qui contribuait davantage à l’hétérogénéité que les cinq autres.

EXPOSE : La PIM était plus élevée chez les hommes et diminuait avec l’âge, une tendance qui a commencé à être apparente à l’âge mur. Plusieurs caractéristiques des participants et la technique de PIM influent sur les valeurs chez les personnes en santé.

CONCLUSIONS : La présente méta-analyse fournit des valeurs de PIM normatives qui reflètent un énorme échantillonnage (n=840) et représentent probablement la plus vaste représentation des caractéristiques des participants par rapport aux rapports antérieurs de données normatives.

I mpairment of the respiratory muscles compromises ventilation, gas exchange and oxygen delivery to tissues (1). Respiratory muscle weakness has been reported in several conditions including those that are cardiovascular (2,3), pulmonary (4) and neuromuscular (5,6) in origin. Inspiratory muscle strength is reflected by the pressure developed within the thorax; these pressure measures are informative for clinical evaluation as well as physiological studies (1). Measurement of maximal inspiratory pressure (MIP) is a straightforward test in which individuals are asked to perform a forceful inspiration against an occluded mouthpiece (1,7). The advantages of this test are that it is noninvasive and performed quickly. On the other hand, its dependence on voluntary effort and the wide range of normative values limit the clinical utility of MIP. Consequently, a low value can reflect inspiratory muscle impairment but also can be due to poor test performance (8-10). Even a well-performed MIP test is difficult to interpret due to uncertainty of the most representative normative values.

Numerous sets of reference values for MIP have been reported (5,7-30); however, the usefulness of normal MIP values is obscured by the large variability among studies. The variation among these reports likely indicates differences in participant demographics and technical aspects of test performance (1,7). Participant characteristics that have been considered to influence MIP include effort and understanding of test performance, age, sex, height, weight, fitness level and smoking status (31-33). Apparatus set-up and test performance issues that can affect MIP values include type of mouthpiece, presence of a small leak, pressure evaluated (ie, peak or plateau), number of trials and lung volume at the starting point of test performance (1,7,19,27,34,35).

Due to the large interstudy variation of normative values for MIP, the data from a single study may not be appropriate to establish the lower limit indicating respiratory muscle weakness (25). The American Thoracic Society (ATS) /European Respiratory Society (ERS) (1)
A search of the Medline, EMBASE, Cochrane, Cumulative Index to Nursing and Allied Health (CINAHL) and Sport Discus databases (from inception to May 2012) was conducted. Inclusion criteria were as follows: participants were healthy adults (>18 years of age); the purpose of the study was to determine reference values for MIP; and published in English or Portuguese. Articles were excluded if they were a review article, thesis or dissertation; and the measurement was assessed in a standing position rather than sitting. The search strategy of search strategies are available on request from the corresponding author. Additional studies were identified by examining the reference lists of articles (manual search). Of which 19 full-text articles were reviewed. Eleven additional articles were identified from the reference lists of articles (manual search). Of the 30 full-text articles reviewed, eight were excluded because the study purpose was not to determine reference values of MIP (31,33,35,42,43); it was a review (43); the MIP manoeuvre was performed at functional residual capacity rather than residual volume (9) or in a standing rather than sitting position (20,44).

RESULTS

Search and selection

A flow diagram depicting the search and selection of studies is shown in Figure 1. The search of databases identified 4367 titles and abstracts, of which 19 full-text articles were reviewed. Eleven additional articles were identified from the reference lists of articles (manual search). Of the 30 full-text articles reviewed, eight were excluded because the study purpose was not to determine reference values of MIP (31,33,35,42,43); it was a review (43); the MIP manoeuvre was performed at functional residual capacity rather than residual volume (9) or in a standing rather than sitting position (20,44).

Methodological quality of studies

Methodological quality, as assessed according to seven items of the QUADAS, ranged between 0 and 7, with a mean score of 3 (Table 1). The most common items that were reported included: similar data available during test results and in practice (15 of 22 studies); study participants were representative of participants who would be tested in practice (12 of 22 studies); and the selection criteria were clearly described (12 of 22 studies). The quality criteria that were least often reported were whether the reference standard was likely to correctly classify the target condition (six of 22 studies); and whether uninterpretable/intermediate tests results were reported (six of 22 studies). Other issues that affected methodological quality were patient characteristics and MIP technique not queried by QUADAS. These issues are described in more detail in the following two sections and in Tables 2 and 3.
Maximal inspiratory pressure: A systematic review

TABLE 1
Ratings of Quality Assessment of Diagnostic Accuracy Studies (QUADAS)

<table>
<thead>
<tr>
<th>Author (reference), year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>9</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>Total (of 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cook et al (11), 1964</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Unclear 0</td>
</tr>
<tr>
<td>Ringqvist (12), 1966</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes 2</td>
</tr>
<tr>
<td>Black and Hyatt (5), 1969</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Unclear 0</td>
</tr>
<tr>
<td>Leech et al (13), 1983</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 5</td>
</tr>
<tr>
<td>Wilson et al (14), 1984</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Unclear 0</td>
</tr>
<tr>
<td>Carmelo Jr et al (15), 1985</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Unclear 2</td>
<td></td>
</tr>
<tr>
<td>Vincken et al (16), 1987</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear 1</td>
<td></td>
</tr>
<tr>
<td>McElvaney et al (17), 1989</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 6</td>
</tr>
<tr>
<td>Bruschi et al (18), 1992</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 3</td>
</tr>
<tr>
<td>Enright et al (19), 1994</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes 3</td>
</tr>
<tr>
<td>Enright et al (21), 1996</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 7</td>
</tr>
<tr>
<td>Johan et al (22), 1997</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Unclear 3</td>
<td></td>
</tr>
<tr>
<td>Pande et al (23), 1998</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 7</td>
</tr>
<tr>
<td>Neder et al (24), 1999</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Unclear 3</td>
</tr>
<tr>
<td>Hautohann et al (25), 2000</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes 4</td>
<td></td>
</tr>
<tr>
<td>Wohlgemuth et al (26), 2003</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Unclear 5</td>
</tr>
<tr>
<td>Windsch et al (27), 2004</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Unclear 5</td>
<td></td>
</tr>
<tr>
<td>Sachs et al (28), 2009</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes 3</td>
</tr>
<tr>
<td>Simões et al (29), 2010</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes 6</td>
</tr>
<tr>
<td>Costa et al (30), 2010</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Unclear 3</td>
</tr>
<tr>
<td>Gopalkrishna et al (10), 2011</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Unclear 0</td>
</tr>
</tbody>
</table>

Subtotals for items

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>9</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>Total (of 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>12</td>
<td>6</td>
<td>10</td>
<td>15</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

QUADAS Items (36): 1. Was the spectrum of persons representative of the demographics of the patients who will receive the test in practice?; 2. Were selection criteria clearly described?; 3. Is the reference standard likely to correctly classify the target condition?; 9. Was the execution of the reference standard described in sufficient detail to permit its replication?; 12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?; 13. Were uninterpretable/intermediate test results reported?; 14. Were withdrawals from the study explained?

Characteristics of participants

The characteristics of the participants are summarized in Table 2. All studies used convenience samples except one (24), which used a random sample procedure. The total number of participants was 9723. Age ranged between 18 and 90 years. The number of men and women ranged from 23 to 1269, and nine to 1602, respectively. Most studies had more women except four (11,12,22,23). Nine studies were performed in North America (5,8,13,16,17,19,21,28), four in Brazil (15,24,29,30), three in India (10,22,23), two in Germany (25,27), and one in Italy (18), Sweden (12), England (14) and The Netherlands (26). Only seven studies reported the height and weight for each age group (10,12,13,24,26,29,30). Most studies included normal and overweight participants (body mass index [BMI] between 18.5 kg/m² and 25.0 kg/m² and between 25.0 kg/m² and 29.9 kg/m², respectively). Some participants in two reports were obese (21,27). Five studies reported other pulmonary function measures apart from forced expiratory volume in 1 s (FEV₁) (11,12,19,22,26). There were 10 studies that did not perform any subgroup analyses on smokers and nonsmokers (5,8,13-16,18,25-27). Of these, six studies did not report FEV₁ values to substantiate the absence or presence of respiratory disease (5,14-16,18,26). Seventeen of the 18 studies stated that individuals with comorbidities, such as cardiac, neuromuscular or endocrine diseases, renal failure and active cancer treatment, were excluded (5,8,10,12,14,16,17,21-30).

Characteristics of MIP technique

The methodology for measuring MIP varied considerably (Table 3). A flanged mouthpiece was used in five studies (14,16,22,24,27). A small leak was reported in 18 studies but the size was not specified in four (14,22,24,25). Most studies reported using nose clips (16 of 22) and provided instructions (18 of 22). The number of repetitive manoeuvres to obtain MIP was at least five in two studies (12,18) and at least seven in two others (25,27). In six studies, the authors considered the learning effect for MIP test; it was reported that the final manoeuvre was not the best manoeuvre in five studies (12,16,24,25,28), the test could be stopped when the subject considered him or herself unable to perform better (22). Several studies considered the highest value within one or two measures that were within 5% to 10% (Table 3). The lower limit of normal was reported in nine studies (8,19,21,24,25,27-30). Most used the fifth percentile of the negative residuals of MIP (8,19,21,25,27-30), which was calculated in each age group for both sexes in one report (27). In contrast, two studies provided an alternative definition (Table 3) and one did not define the lower limit of normal (28).
<table>
<thead>
<tr>
<th>Author (reference), year</th>
<th>Age, years (range)</th>
<th>Height, cm, mean ± SD</th>
<th>Weight, kg, mean ± SD</th>
<th>BMI, kg/m^2, mean ± SD or 95% CI</th>
<th>FEV₁, mean ± SD or 95% CI</th>
<th>Lung disease</th>
<th>Smoking status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cook et al (11), 1964, n=32 (23M/9F)</td>
<td>M: Gp 3 18–47, Gp 4 49–64; F: Gp 6 17.5–32</td>
<td>M: Gp 3 181; Gp 4 177; F: Gp 6 164</td>
<td>Reported according to age</td>
<td>Reported according to age</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Ringqvist (12), 1966, n=200 (106M/94F)</td>
<td>18–83</td>
<td>Reported according to age</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>No NR</td>
</tr>
<tr>
<td>Black and Hyatt (5), 1969; n=120, (60M/60F)</td>
<td>20–74</td>
<td>NR</td>
<td>NR</td>
<td>Reported according to age</td>
<td>NR</td>
<td>No NR</td>
<td></td>
</tr>
<tr>
<td>Leech et al (13), 1983; n=595 (252M/343F)</td>
<td>15–35</td>
<td>Reported according to age</td>
<td>M: 74.5±8.5</td>
<td>M: 4.5–4.6 L*</td>
<td>F: 3.2–3.4 L</td>
<td>NR</td>
<td>Former and current smokers</td>
</tr>
<tr>
<td>Wilson et al (14), 1984; n=135 (48M/87F)</td>
<td>18–49</td>
<td>M: 179±6</td>
<td>Reported according to age</td>
<td>M: 20, F: 23</td>
<td>NR</td>
<td>No Former and current smokers</td>
<td></td>
</tr>
<tr>
<td>Camelo Jr et al (15), 1985; n=60 (30M/30F)</td>
<td>20–49</td>
<td>M: 170±7.7</td>
<td>Reported according to age</td>
<td>M: 24, F: 22</td>
<td>NR</td>
<td>No Former and current smokers</td>
<td></td>
</tr>
<tr>
<td>Vincken et al (16), 1987; n=106 (46M/60F)</td>
<td>16–79</td>
<td>M: 172±7</td>
<td>Reported according to age</td>
<td>M: 25, F: 23</td>
<td>NR</td>
<td>No Nonsmokers and smokers</td>
<td></td>
</tr>
<tr>
<td>McElvaney et al (17), 1989; n=104 (40M/64F)</td>
<td>&gt;55</td>
<td>M: 174±7</td>
<td>Reported according to age</td>
<td>M: 120±13%</td>
<td>F: 117±17%</td>
<td>No NR</td>
<td></td>
</tr>
<tr>
<td>Bruschi et al (18), 1992; n=669 (290M/379F)</td>
<td>18–70</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>No Nonsmokers and smokers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enright et al (21), 1995; n=228 (112M/116F)</td>
<td>&gt;65</td>
<td>M: 171.25</td>
<td>Reported according to age</td>
<td>M: 2.9 (2.0–3.7) L</td>
<td>F: 1.9 (1.2–2.6) L</td>
<td>No Nonsmokers and former smokers</td>
<td></td>
</tr>
<tr>
<td>Johan et al (22), 1997; n=452 (277M/175F)</td>
<td>20–80</td>
<td>M: 164–157</td>
<td>Reported according to age</td>
<td>M: 26.4 (22.32)</td>
<td>F: 26 (20.34)</td>
<td>No Nonsmokers and former smokers</td>
<td></td>
</tr>
<tr>
<td>Pande et al (23), 1998; n=273 (153M/120F)</td>
<td>20–65</td>
<td>M: 165±6.1</td>
<td>Reported according to age</td>
<td>M: 22.8±3.8</td>
<td>F: 24.6±4.8</td>
<td>No Nonsmokers</td>
<td></td>
</tr>
<tr>
<td>Harik-Khan et al (8), 1998; n=267 (139M/128F)</td>
<td>&lt;40 – &gt;75</td>
<td>M: 164±7.3</td>
<td>Reported according to age</td>
<td>M: 91.7±12.0%</td>
<td>F: 98.0±10.3%</td>
<td>No Nonsmokers, former, occasional, current smokers</td>
<td></td>
</tr>
<tr>
<td>Neder et al (24), 1999; n=100 (50M/50F)</td>
<td>20–80</td>
<td>M: 168±6.2</td>
<td>Reported according to age</td>
<td>M: 25±3.3%</td>
<td>F: 24.7±4.0</td>
<td>No Nonsmokers, former, occasional, current smokers</td>
<td></td>
</tr>
<tr>
<td>Haußmann et al (25), 2000; n=504 (248M/256F)</td>
<td>18–82</td>
<td>M: 176±6.82</td>
<td>Reported according to age</td>
<td>M: 106±21</td>
<td>F: 3.01±0.71 L</td>
<td>No Nonsmokers</td>
<td></td>
</tr>
<tr>
<td>Wohlgemuth et al (26), 2003; n=252 (126M/126F)</td>
<td>18–80</td>
<td>Reported according to age</td>
<td>M: 177±6.6</td>
<td>Reported according to age</td>
<td>NR</td>
<td>No Former and current smokers,</td>
<td></td>
</tr>
<tr>
<td>Windisch et al (27), 2004; n=490 (204M/286F)</td>
<td>10–90</td>
<td>M: 179.5±7.7</td>
<td>Reported according to age</td>
<td>M: 24.2±3.1</td>
<td>F: 23.9±4.1</td>
<td>No Never, former and current smokers,</td>
<td></td>
</tr>
<tr>
<td>Sachs et al (28), 2009; n=1755 (872M/883F)</td>
<td>45–84</td>
<td>M: 172</td>
<td>Reported according to age</td>
<td>M: 27.19; F: 27.31</td>
<td>&gt;65%</td>
<td>No Nonsmokers</td>
<td></td>
</tr>
<tr>
<td>Simões et al (29), 2010; n=140 (70M/70F)</td>
<td>20–89</td>
<td>Reported according to age</td>
<td>M: 92±8%</td>
<td>F: 93±7%</td>
<td>No Nonsmokers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costa et al (30), 2010; n=120 (60M/60F)</td>
<td>20–80</td>
<td>Reported according to age</td>
<td>M: 165.70±7.56</td>
<td>Reported according to age</td>
<td>NR</td>
<td>No Nonsmokers</td>
<td></td>
</tr>
<tr>
<td>Gopalakrishna et al (10), 2011; n=250 (125M/125F)</td>
<td>20–70</td>
<td>M: 155.99±5.81</td>
<td>Reported according to age</td>
<td>M: 23.54±3.21</td>
<td>F: 23.17±3.89</td>
<td>No Nonsmokers</td>
<td></td>
</tr>
</tbody>
</table>

*Range of mean data shown for different subgroups. BMI Body mass index; F Female; FEV₁ Forced expiratory volume in 1 s; Gp Group; M Male; NR not reported
TABLE 3
Technical aspects that influence maximal inspiratory pressure (MIP)

<table>
<thead>
<tr>
<th>Author (reference), year</th>
<th>Mouth-piece</th>
<th>Small leak (size)</th>
<th>Pressure evaluated</th>
<th>Starting volume</th>
<th>Time of MIP</th>
<th>Trials, n</th>
<th>Criterion for stopping</th>
<th>LLN (percentile or cmH2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cook et al (11), 1963</td>
<td>Tube</td>
<td>NR</td>
<td>Peak</td>
<td>RV and FRC</td>
<td>Without control</td>
<td>Min of 2</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Ringqvist (12), 1966</td>
<td>Tube</td>
<td>Yes (2 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Max 1.5 s</td>
<td>Min of 5</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Black and Hyatt (5), 1969</td>
<td>Tube</td>
<td>Yes (2 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 2</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Leech et al (13), 1983</td>
<td>NR</td>
<td>Yes (0.90 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>NR</td>
<td>Max of 3</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Wilson et al (14), 1984</td>
<td>Flanged</td>
<td>Yes, size NR</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 3</td>
<td>2 identical readings</td>
<td>NR</td>
</tr>
<tr>
<td>Camelo Jr et al (15), 1985</td>
<td>Tube</td>
<td>Yes (2 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 4</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Vincken et al (16), 1987</td>
<td>Flanged</td>
<td>Yes (1.27 mm)</td>
<td>Plateau</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 4</td>
<td>Highest 2 values within 5% difference</td>
<td>NR</td>
</tr>
<tr>
<td>McElvaney et al (17), 1989</td>
<td>Tube</td>
<td>Yes (0.6 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 3</td>
<td>Highest 3 values within 5% difference</td>
<td>NR</td>
</tr>
<tr>
<td>Bruschi et al (18), 1992</td>
<td>Tube</td>
<td>Yes (1.06 mm)</td>
<td>Peak</td>
<td>RV and FRC</td>
<td>Min of 1 s</td>
<td>Min of 5</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Enright et al (19), 1994</td>
<td>Tube</td>
<td>Yes (1 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>2 s</td>
<td>3–5</td>
<td>Highest 2 values within 10% difference</td>
<td>Yes (5th %); M: −41; F: −32</td>
</tr>
<tr>
<td>Enright et al (21), 1995</td>
<td>Tube</td>
<td>Yes (1 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>2 s</td>
<td>5</td>
<td>Highest 2 values within 10% difference</td>
<td>Yes (5th %); M: −42; F: −38</td>
</tr>
<tr>
<td>Johan et al (22), 1997</td>
<td>Flanged</td>
<td>Yes, size NR</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>3–5</td>
<td>Highest value of 3 similar trials</td>
<td>NR</td>
</tr>
<tr>
<td>Pande et al (23), 1998</td>
<td>NR</td>
<td>Yes (1.27 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 2 s</td>
<td>NR</td>
<td>Highest value</td>
<td>NR</td>
</tr>
<tr>
<td>Harik-Khan et al (8), 1998</td>
<td>Tube</td>
<td>Yes (1 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>2 s</td>
<td>Max of 5</td>
<td>Highest 2 values within 10% difference</td>
<td>Yes (5th %); M: −37; F: −32</td>
</tr>
<tr>
<td>Neder et al (24), 1999</td>
<td>Flanged</td>
<td>Yes, size NR</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>3–5</td>
<td>Highest value, &lt;10% of 3 trials</td>
<td>Yes (1.98 × SEE); M: −34; F: −18</td>
</tr>
<tr>
<td>Hautmann et al (25), 2000</td>
<td>NR</td>
<td>Yes, size NR</td>
<td>Plateau</td>
<td>RV</td>
<td>Min of 2 s</td>
<td>Min of 7</td>
<td>Highest value</td>
<td>Yes (5th %); M: −0.60; F: +0.59</td>
</tr>
<tr>
<td>Wohlgemuth et al (26), 2003</td>
<td>Face mask</td>
<td>Yes (2 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 3</td>
<td>Highest value varying 5%</td>
<td>NR</td>
</tr>
<tr>
<td>Windisch et al (27), 2004</td>
<td>Flanged</td>
<td>Yes (2 mm)</td>
<td>Peak and plateau</td>
<td>RV and FRC</td>
<td>Min of 1 s</td>
<td>Min of 7</td>
<td>Highest 2 values within 10% difference</td>
<td>Yes (5th %); according to age in decades</td>
</tr>
<tr>
<td>Sachs et al (28), 2009</td>
<td>Tube</td>
<td>NR</td>
<td>Plateau</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>5</td>
<td>Highest 2 values within 10% difference</td>
<td>Yes (NR); M: −40; F: −36</td>
</tr>
<tr>
<td>Simões et al (29), 2010</td>
<td>Tube</td>
<td>Yes (2 mm)</td>
<td>Plateau</td>
<td>RV</td>
<td>About 1 s</td>
<td>Min of 3</td>
<td>Highest value &lt;10% of all trials</td>
<td>Yes (1.645×SEE); M: −25; F: −69</td>
</tr>
<tr>
<td>Costa et al (30), 2010</td>
<td>NR</td>
<td>Yes (2 mm)</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 3</td>
<td>Highest value &lt;10% of 2 trials</td>
<td>Yes (5th %); M: −23.38; F: −28.83</td>
</tr>
<tr>
<td>Gopalakrishna et al (10), 2011</td>
<td>NR</td>
<td>NR</td>
<td>Peak</td>
<td>RV</td>
<td>Min of 1 s</td>
<td>Min of 3</td>
<td>Highest 2 values within &lt;10% difference</td>
<td>NR</td>
</tr>
</tbody>
</table>

5th % Fifth percentile of the negative residuals of maximal inspiratory pressure; F Female; FRC Functional residual capacity; LLN Lower limit of the normal range; M Male; Max Maximum; Min Minimum; n/a not applicable; NR Not reported; Peak The highest value reached during a brief maximal effort; Plateau The highest pressure sustained for a defined minimum period; RV Residual volume; SEE Standard error of the estimate of the model

DISCUSSION

Our meta-analysis provides a synthesis of normative values for MIPs, the most commonly performed measure to indicate inspiratory muscle strength. We reported data based on 840 subjects, the largest and most representative reference group available for this measure. This synthesis demonstrates strong age-related trends and sex differences in MIP, similar to trends reported by the included studies.

Our meta-analysis considered data from 840 subjects recruited from North America, Brazil, Sweden and The Netherlands. Therefore, the reference values are reflective of women and men from different ethnic backgrounds, with associated differences in body stature that could be expressed as variations in MIPs. Mean MIP data from each of the six studies that contributed to the meta-analyses had mean values (in different age groups for each sex) that overlapped with at least one other study. MIP values from three reports (29,12,24) were generally higher for all age groups in both sexes, whereas the data from the two reports (26,30) were generally lower in both sexes for most age groups. Interesting, the two studies that recruited American subjects and reported the highest quality scores (8,21) provided regression equations that produced MIPs that were higher than our mean values in one case (21) and lower in the other (8). In other words, the reports with high methodological quality provided mean MIP values that bracket the mean data derived from our synthesis, which lends further credence to the validity of the data derived by the meta-analyses.

Application of reference values require consideration of participant and methodological differences. Participant characteristics varied considerably among studies, and many of these could not be considered by our subgroup analyses. In addition to age and sex, other factors are reported to influence MIP (eg, height, weight, fitness and smoking status) (31-33). Unfortunately, there is no consensus regarding which of these variables have a significant influence on MIP or the directional influence of some of these variables. For example, height has been shown to be positively predictive (14,22), negatively predictive (8,22) and not predictive (5,16,18,19,21,23-25,29,30) of MIP. Smoking exemplifies another characteristic that could influence normative values if the participant had undetected lung disease.
Most studies demonstrated that MIP was significantly affected by sex and age in healthy subjects. This is in accordance with the study by Black and Hyatt (5), who reported a linear regression of MIP on age in both males and females; however, the regression was not significant in subjects <55 years of age. Our data are in accordance with this early investigation (5) and indicates modest decreases in MIP before 45 to 55 years of age (Table 4). Unfortunately, there is no consensus regarding the threshold when age-related decline in MIP occurs. It has been reported to occur after 30 years of age in one study (16) or after 60 years of age in another (25). It is plausible that MIP has not undergone study with sufficiently large sample sizes to determine the many factors that may contribute to valid reference values. A recent article on reference values for spirometry examined 55,136 healthy subjects (45), whereas the number of participants in the studies for the present systematic review ranged from 60 to 252. Considering the relatively small sample sizes and the ethnic diversity in studies performed on different continents, it is not surprising that MIP values varied among studies. Future studies should recruit larger numbers and factor in confounders such as ethnicity, smoking history, physical activity and parameters of body stature (ie, BMI, height and weight).

A primary methodological issue that likely contributed to the large interstudy variation of normative values for MIP is participant learning. Given that MIP is volitional, it requires full understanding of the task to be performed, maximal participant effort and enthusiastic coaching from the technologist. To attain a valid MIP, the selection of truly maximum efforts depends not only on setting a limit on variation between successive measurements, but also by carefully considering whether optimal learning has occurred. Several studies did not consider the need to continue the test if the final manoeuvre was the highest value (5,8,10,11,13-15,17-19,21-23,27,28,30). Incomplete learning combined with too few repetitions to achieve a maximal value are likely primary factors that influence heterogeneity of MIP values among studies.

The large interstudy variation of normative values for MIP may also be explained by technical aspects of test performance including variation in the type of mouthpiece, size of leak and the point of MIP measurement (peak or plateau) (Table 3). Five of the 22 studies used flanged mouthpieces (14,16,22,24,27), which have been reported to provide lower values than those obtained with rubber tube mouthpieces (1). A larger leak can also result in lower values (ie, MIP measurements obtained with leak of 2 mm internal diameter were 17% and 22% lower than those with leak of 1 mm or no leak, respectively) (35). A plateau measure that is averaged over 1 s would be lower than a peak pressure defined as the highest value reached during a brief maximal effort (18,19). Notably, not all studies after 2002 used the 1 s average pressure, which was recommended by the ATS/ERS, even though this measure is more reproducible than the peak pressure (1,46).

Limitations

The present synthesis was limited by variability in study quality, participant characteristics and MIP technique. Of the 22 included studies, only six were included in the present meta-analysis because of the necessity for MIP data to be reported according to sex and stratified into one or more of six comparable age groups (Table 4). Nine other included studies stratified MIP data according to age groups; however, the categories included either a wider age range or straddled the above-mentioned groupings (5,8,11,13,14,17,21,23,25). The methodological quality of studies (ie, QUADAS assessment) examined in the meta-analysis ranged between 2 and 6. Unfortunately, it was not possible to include some of the studies that had a QUADAS score of 7 of 7 because the MIP data were not reported in similar age groups (8) or they were presented as regression equations (8,21).

Despite an ATS/ERS statement on respiratory muscle testing, many subsequent reports included in the present review did not describe methods consistent with these guidelines. Several studies did not use the recommended flanged mouthpiece or did not report the type of mouthpiece (26-30). Description of a sustained pressure for 1.5 s to evaluate the maximum sustained pressure for 1 s was rarely detailed. It was also seldom noted whether the tester was experienced, and how much urging and encouragement was provided to the individual being tested. Regarding some of these technical issues, it was not possible to conclude whether some of these methods were not performed or simply not reported.

Attributing a bias toward larger or smaller MIP due to participant characteristics in a particular study was also difficult because of small sample sizes, the large number of potential confounders that could influence strength measures, and equivocal findings regarding their effect. For example, greater height and fitness were not consistently associated with higher MIPs. Because of these uncertainties and omissions in reporting, it was difficult to conclude whether MIP data were higher or lower in a particular study due to differences in MIP technique or participant characteristics. Thus, to date, no predictable correction factors can be applied to MIP data to normalize technical or participant differences.

Practical implications

Standardized techniques for measuring MIP will ensure more reliable and valid data. The ATS/ERS statement on respiratory muscle testing recommends that MIPs are measured using a flanged mouthpiece with a 2 mm leak at residual volume (1). The latter two criteria were consistently met by all six studies included in the present meta-analysis (12,15,24,26,29,30). A sustained pressure for 1.5 s to evaluate the maximum sustained pressure for 1 s is also recommended to obtain reliable measures (1). Finally, the importance of an experienced tester who can optimize volitional effort of the participant by clear instruction and strong encouragement is essential (1). These latter two issues, although not always reported in the included studies, are highly recommended for attaining MIP.

Regardless of the reasons for the differences in MIP among studies, appropriate normative values are required by pulmonary function laboratories to provide a valid reference of respiratory function for an individual patient. Ideally, appropriate reference values should be obtained in the same laboratory from a sample that represents the demographic characteristics of the patients to be tested. From a practical perspective, such reference values are usually not established in many centres due to limited resources. In these situations, the data
REFERENCES


Maximal inspiratory pressure: A systematic review

AUTHOR CONTRIBUTIONS: Study concept and design: ISP, VFP, WDR, GF; Data acquisition, analysis or interpretation of data: ISP, VFP, WDR, AWS, FC; Drafting the manuscript, revising and approval of the submitted and final version: ISP, VFP, AWS, FC, GF, WDR.

DISCLOSURES: The authors have no financial disclosures or conflicts of interest to declare.