The aim of this study was to evaluate the efficacy of noninvasive ventilation (NIV) on improving exercise tolerance of patients with heart failure (HF). A systematic review was performed in PubMed/MEDLINE, LILACS, Cochrane, CINAHL, Scopus and Web of Science for randomized and quasi-randomized clinical trials, without language and year of publication restrictions. Descriptors were defined as ‘heart failure’, ‘noninvasive ventilation’, ‘positive-pressure respiration’, ‘interactive ventilatory support’, ‘exercise test’ in addition to the keywords ‘BIPAP’, ‘CPAP’, ‘IPAP’, ‘EPAP’, ‘NIV’ and their Portuguese equivalents. Studies comparing NIV with one or two pressure levels to groups without intervention, other physiotherapy modalities without positive pressure or a sham group were included. Four studies were selected, including HF patients of various etiologies, considering the staging classification of New York Heart Association. Any included work realized the allocation concealment, all studies participants underwent blinding, but only two trials performed assessors blinding. None of the studies have described an intention to treat analysis and did not use appropriate statistical methods. All selected trials assessed functional capacity and in only two, dyspnea was assessed. The intervention protocols of the included trials were heterogeneous, three studies underwent a single intervention with NIV, two immediately before the functional capacity test and another study performed NIV during the exercise evaluation. The last trial held 14 sessions of NIV, with the functional capacity evaluation being performed on days 0, 4, 9 and 14. There is insufficient evidence on the effectiveness of NIV in increasing exercise tolerance.

Keywords | Heart Failure; Noninvasive Ventilation; Exercise Test.

RESUMO | O objetivo do estudo foi avaliar a eficácia da ventilação não invasiva (VNI) na melhora da tolerância ao exercício em indivíduos com insuficiência cardíaca (IC). Realizou-se uma busca sistemática nas bases de dados PubMed/MEDLINE, LILACS, Cochrane, CINAHL, Scopus e Web of science por ensaios clínicos randomizados e quasi-randomizados. Os descritores foram: ‘heart failure’, ‘noninvasive ventilation’, ‘positive-pressure respiration’, ‘interactive ventilatory support’, ‘exercise test’, além das palavras-chave ‘BIPAP’, ‘CPAP’, ‘IPAP’, ‘EPAP’, ‘NIV’ e seus equivalents em português. Foram incluídos estudos que comparassem a VNI com um ou com dois níveis de pressão a grupos sem intervenção, a outras modalidades fisioterapêuticas sem pressão positiva ou a um grupo sham. Foram selecionados quatro estudos, incluindo pacientes com IC de diversas etiologias. Os quatro estudos foram randomizados e controlados e realizaram o mascaramento dos participantes. No entanto, apenas dos trabalhos realizaram o mascaramento dos avaliadores. Em nenhum dos artigos selecionados foi feita a análise por intenção de tratar, e apenas um não utilizou métodos estatísticos adequados. Todos os estudos avaliaram a capacidade funcional e dois avaliaram a dispneia. Os protocolos de intervenção were heterogeneous, three studies underwent a single intervention with NIV, two immediately before the functional capacity test and another study performed NIV during the exercise evaluation. The last trial held 14 sessions of NIV, with the functional capacity evaluation...
Heart failure (HF) is a great worldwide problem that affects more than 20 million people. In Brazil, this prevalence reaches around two million subjects and there is an incidence of 240,000 new diagnosed cases every year. This complex syndrome brings several complications both for the society, like the high socioeconomic cost for the health system and early retirement due to productivity loss, and for the subject with the disease, like physical and psychological problems that cause his/her social isolation.

Alterations from the HF are not restricted to the cardiac scope. The main symptoms are dyspnea and fatigue, which can limit exercise tolerance. However, this population presents a musculoskeletal associated involvement due to the low cardiac output. The progression of these symptoms creates a decrease in the level of physical activity, which contributes to the worsening of the symptoms and exercise intolerance. Thus, the functional capacity and quality of life of this population is progressively reduced, which results in a frequent, high-cost and usually disabling clinical condition. In addition, this cardiac illness is associated with frequent hospitalization and re-hospitalization, functional involvement that evolves with significant morbidity due to low physical capacity and high mortality.

The noninvasive ventilation (NIV) arises as a co-adjuvant therapy in the attempt of improving patients’ functional capacity. Use of the NIV has been an option to reduce respiratory work, improve blood oxygenation, and lung complacency in order to provide an improvement in the tolerance during the physical exercise due to its sensible actuation in the cardiorespiratory interaction, therefore giving a better cardiac and respiratory response during the exercise.

Thus, the present review has as its aim to evaluate the NIV effectiveness with continuous pressure or with two levels of pressure in the exercise tolerance in adult patients from both genders with HF.
The primary outcome considered exercise capacity (oxygen consumption; distance) that was assessed through the maximum or sub-maximum exercise test and dyspnea level — evaluated using subjective scales. The oxygen peripheral saturation (SpO₂), blood pressure (BP), and cardiac frequency (CF) were the secondary outcomes.

A systematic search was performed to find articles published in indexed journals in the databases PubMed/ MEDLINE, LILACS, Cochrane, CINAHL, Scopus and Web of Science. The research strategy was based on the standards from the Cochrane Handbook for Systematic Reviews of Interventions. The keywords applied for the search followed the description of MeSH/DeCS terms, namely: ‘heart failure’, ‘noninvasive ventilation’, ‘positive pressure respiration’, ‘interative ventilator support’, ‘exercise test’. Besides these, the following keywords were used: ‘BIPAP’, ‘CPAP’, ‘IPAP’, ‘EPAP’, ‘NIV’. It was also done a research with the same words in Portuguese. The words were combined using the Boolean operators ‘OR’, ‘AND’ and ‘NOT AND’. There were no linguistic and publication year restraints. The investigation was carried out between April and May 2013.

In the initial phase, two independent reviewers (LAC and CRL) identified and evaluated titles and abstracts on the computer screen, in order to choose those that would meet eligibility criteria. Potentially relevant studies that raised doubts were retained for a posterior analysis of the complete text. In cases of disagreement in the process of selection and analysis of the articles, a third evaluator (DCB) would take part in the evaluation.

Data collection from the chosen studies was performed by two independent evaluators. Extracted data were: eligibility criteria, study population, participants’ flow, intervention details, outcome measures, results, and bias risk.

Two independent evaluators analyzed studies that met the inclusion criteria as to the methodological quality using PEDro scale, based on the Delphi list. PEDro scale was developed to be used in clinical trials and, recently, it is one of the most used in Physical Therapy. It allows a total score of ten points. For each criterion presented in it, a score of one or zero point may be attributed.

The oxygen peripheral saturation (SpO₂), blood pressure (BP), and cardiac frequency (CF) were the secondary outcomes.

In the end of data collection, it was seen the possibility of preparing a meta-analysis. However, this was not possible due to heterogeneity of the included studies and lack of data.

**RESULTS**

The search strategy was ample and resulted in 1,359 titles. From these, 1,300 were excluded by titles for not presenting the inclusion criteria. Of the 59 remaining studies, 12 were discarded due to duplicity, 19 were deviated due to not complying with the theme or inappropriate study design. Twenty-eight studies were chosen, which were assessed through abstract analysis, and 18 were excluded for not presenting inclusion criteria. The ten retained articles had their texts completely read, of which six did not present eligibility criteria, as described in Figure 1.

Therefore, four studies were included for qualitative synthesis. Studies included 58 adults with HF diagnosis, 18 women and 40 men, whose average of age varied from 33 to 68 years old. These studies were originally from Brazil and Canada and the population was composed by subjects with HF of several etiologies. However, two of the analyzed studies included only the disease congestive form. Staging was different between the studies, three of them included patients of II and III functional class and only one investigation had subjects of II-IV functional class. Nevertheless, they were always using the classification recommended by the New York Heart Association (Table 1).

Three studies evaluated the NIV effect with continuous pressure in the exercise tolerance through the 6-minute walking test, which is an effort sub-maximum exercise test. Only one of the studies evaluated and compared the NIV effect with two levels of pressure and with continuous pressure in the exercise tolerance, using the sub-maximum test of constant load in cycloergometer to assess functional capacity. There was also a variation as to the Control Group: in two studies, subjects did not perform any kind of NIV. However, in two others, they did the NIV in placebo with low pressure levels.

As to the control of patients’ selection, all studies mentioned performing randomization. However, only one of the included papers specified the used method: a raffle in an opaque and sealed envelope. As to the allocation secrecy, a study reported that the envelope used to keep secrecy was the opaque one. Nonetheless, none of the analyzed studies specified if one independent person performed the randomization.

In one article, there was an outcome selective description, besides expression of its results only in gain percentage...
Figure 1. Research and selection of studies for the systematic review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Table 1. Characteristics of the included studies

<table>
<thead>
<tr>
<th>Author, Country</th>
<th>Population</th>
<th>Age (mean)</th>
<th>Diagnosis</th>
<th>Criteria used for staging</th>
<th>Intervention protocol (Type of therapy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chermont et al. Brazil</td>
<td>G: 4 women and 8 men G: 56±12</td>
<td>HF (ischemic 6 and idiopathic 6) LVEF&lt;45%</td>
<td>NYHA II/III: 4/8</td>
<td>Type of therapy: NIV – CPAP 4 to 6 cmH₂O (individual pressure) through nasal blinding in decubitus position at 45°, before the 6WT. Placebo – ventilator support performance of 0.1 cmH₂O in decubitus position at 45°, before the 6WT. Frequency: once in each experimental situation, in different days (does not report the interval period). Duration: NIV and Placebo – 30 minutes. Functional capacity test: 6WT. Instrument: CPAP (Traquility, Healthdyne Technologies, Marietta, GA).</td>
<td></td>
</tr>
<tr>
<td>O’Donnell et al. Canada</td>
<td>G: 1 woman and 11 men; G: 61±4</td>
<td>CHF (ischemic and idiopathic)</td>
<td>NYHA II/IV</td>
<td>Type of therapy: NIV – CPAP (4.8±0.2 cmH₂O) during the constant load test in cycloergometer NIV 2 – SP (4.8±0.2 cmH₂O) during the constant load test in cycloergometer Placebo (CPAP 1 cmH₂O) during the constant load test in cycloergometer (75% of the maximum work). Frequency: once in every experimental situation with an interval of 1 hour between them. Duration: CPAP/SP/ Placebo during the Exercise test in cycloergometer. Functional capacity test: the constant load test in in cycloergometer Instrument: Resplicone. Ventilator (Resplicone, Murrayville, PA).</td>
<td></td>
</tr>
<tr>
<td>Wittmer et al. Brazil</td>
<td>NIV: 6 women and 6 men; Control: 4 women and 6 men. NIV: 598±37 Control: 52.7±11.4</td>
<td>CHF (diabetic 13, alcoholic 1, after rheumatic fever 8)</td>
<td>NYHA II and III</td>
<td>Type of therapy: NIV – CPAP performance (8 cmH₂O); 100 meters walk every day, respiratory exercises Control – 100 meters walk every day, respiratory exercises Frequency: NIV and Control – 14 sessions. Duration: NIV – 30 minutes of CPAP Functional capacity test: 6WT Instrument: CPAP (without specifications).</td>
<td></td>
</tr>
</tbody>
</table>

HF: heart failure; LVEF: left ventricle ejection fraction; NYHA: New York Heart Association; 6WT: 6-minute walking test; NIV: noninvasive ventilation; CPAP: continuous positive pressure; G: group; CHF: congestive heart failure; SP: support pressure.
(value after treatment – value before treatment/value before treatment x 100). It also analyzed its results, comparing the same group before and after intervention. Only one article\(^8\) mentioned sampling loss and none performed analysis with treatment intention. All articles\(^8\)-\(^11\) performed participants' blinding. However, only two did evaluator's blinding\(^8\)-\(^11\). Study scoring, according to PEDro scale, is presented in Table 2.

With regard to outcome evaluation, with the exception of one paper\(^11\), in which results were not clearly exposed, the functional capacity was assessed in a reproducible manner in three studies\(^8\)-\(^10\). As to the dyspnea outcome, only two studies tried to analyze data\(^9\),\(^10\), and both used Dyspnea Borg's Scale. In one of the articles\(^8\), there was data collection through Dyspnea Borg's Scale every two minutes during the 6'WT, but this was not presented. As to the secondary outcomes, two studies\(^9\),\(^10\) evaluated the oxygen saturation outcomes and three\(^8\)-\(^10\) assessed BP and CF (Table 2).

After observing intervention protocols in every study, a great heterogeneity was seen. In three studies\(^8\)-\(^10\), only one NIV session was performed, and two studies\(^8\),\(^9\) did it right after the functional capacity test (FCT), and the remaining study\(^19\) did the NIV during the FCT. The other included article\(^11\) performed 14 NIV sessions, and the functional capacity evaluation happened in days 0, 4, 9, and 14.

Studies had different information not only with regard to the number of sessions, but also to the used pressure. While two studies\(^9\),\(^11\) established the continuous pressure (CPAP) in 10 and 8 cmH\(_2\)O, respectively, other two ones\(^8\),\(^10\) performed one pressure titling. A study\(^8\) was titled based on subject’s satisfaction from a previous evaluation, adopting a CPAP of 3 to 6 cmH\(_2\)O. In addition, the other included study\(^10\) titled the pressure based on the benefits also found from a previous evaluation, adopting the same value as in the continuous pressure phase and for the two-pressure levels phase, mean of the adopted pressures was of 4.8 cmH\(_2\)O. In three articles\(^8\),\(^10\),\(^11\), intervention groups were described with details including important information, such as the positioning during NIV administration and used interface. Intervention groups and their respective protocols, outcomes and methods applied by the authors for measurement and intervention results are exposed in Table 1. Two studies were designed with the crossover kind, one\(^8\) did not mention the washout period between the experimental and placebo phases. Another research\(^4\) opted in performing the three effort sub-maximum exercise tests – two experimental and one placebo – at the same day, with a resting period of one hour between them.

Sampling calculation was performed in only one study\(^8\), which used literature information to carry it out. None of the articles defined the clinically important difference in the functional capacity or dyspnea evaluation. Furthermore, none of the analyzed articles presented conflicts of interest.

### Table 2: Evaluation of outcomes

<table>
<thead>
<tr>
<th>Author, Country</th>
<th>Type of outcome evaluations (intergroup analysis)</th>
<th>Functional capacity (instrument)</th>
<th>Dyspnea (instrument)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chermont et al(^8) Brazil</td>
<td>After the intervention and placebo phases</td>
<td>Distance (6'WT) NIV: 507±33 meters Placebo: 446±36 meters p≤0.001</td>
<td>Not assessed</td>
</tr>
<tr>
<td>Lima et al(^9) Brazil</td>
<td>Before and after for the intervention group and for the control group (intra and intergroup analysis)</td>
<td>Distance (6'WT) NIV: 534489±91 meters Control: 420±73±8 meters p≤0.03</td>
<td>Borg NIV: 11±0.8 Control: 13±1±16 p≤0.009</td>
</tr>
<tr>
<td>O’Donnell et al(^10) Canada</td>
<td>After the intervention and placebo phases</td>
<td>VO(_2) (Constant load test in cycloergometer: Ergometrics 800S, SensorMedics) SP: 1.47±0.14 L/min NS CPAP: 1.46±0.14 L/min 6 N5 Control: 1.5±0.14 L/min</td>
<td>Borg SP: 5.5±1.05 NS CPAP: 5.1±1.05 NS Control: 5.2±0.5</td>
</tr>
<tr>
<td>Wittmer et al(^8) Brazil</td>
<td>At days 0, 4, 9, and 14 for the intervention and control groups (intragroup analysis)</td>
<td>Distance (6'WT) Pre NIV: 344±25 meters Post NIV: increase of 28% p≤0.005 Pre control: 34±16 meters Post control: not reported p≤NS</td>
<td>Not assessed</td>
</tr>
</tbody>
</table>

NIV: noninvasive ventilation; 6'WT: 6-minute walking test; SP: support pressure; CPAP: continuous positive pressure; Pre NIV: basal value of the noninvasive ventilation group; Post NIV: post-test value of the noninvasive ventilation group; Pre Control: basal value of the control group; Post Control: post-test value of the control group; VO\(_2\): oxygen consumption.
DISCUSSION

Due to the methodological heterogeneity of clinical trials\(^8-11\) with regard to the applied protocols, as well as to the clear lack of demonstration of results, the effectiveness of NIV was not evident in the increment of exercise tolerance in subjects with HF. In addition, all studies analyzed their outcomes based only on the statistical significance, and it was not possible to calculate the magnitude of the treatment effect. The methodological aspects of the included articles deserve some considerations.

Some methodological artifices are used in order to reduce bias risk. The best way to minimize the selection bias with certification of the treatment effect is the proper performance of a randomization process and allocation secrecy. It was observed that a treatment can be overestimated in up to 40% when such items are not properly carried out\(^15,16\). Besides, the random allocation of the participants of a study is done to balance the characteristics of the groups, avoiding confusion factors\(^16\). With regard to the selection control of participants, it was seen that all included studies\(^8-11\) mentioned performing the randomization process. However, only one of the articles\(^9\) described the adopted process, despite the fact that it is not an appropriate method. On the other hand, as to the allocation secrecy, a study\(^9\) reported that the envelope used to keep secrecy was opaque; however, none of the analyzed studies specified if an independent subject performed the allocation. By knowing the importance of these methodological aspects, once they avoid overestimation of the treatment effect, it is possible that different results may be evidenced in studies that perform randomization and allocation secrecy correctly, suggesting a lower effectiveness of the technique.

In addition, to try minimizing the selection bias of the studies, very well established inclusion criteria were determined in order to homogenize the studied sample. Nevertheless, one of the studies\(^11\) evidenced a difference in the basal values of the ejection fraction (EF) for the studied groups. Therefore, the highest EF was for the intervention group compared to the Control Group. EF is an important variable able to interfere in the blood supply and consequently in the functional capacity of the assessed subjects\(^17\).

Participants’ blinding, although not possible of being performed in all situations, is as necessary

---

**Table 2. Evaluation of outcomes (continuation)**

<table>
<thead>
<tr>
<th>Author, Country</th>
<th>Type of outcome evaluations</th>
<th>Heart frequency (instrument)</th>
<th>Systolic blood pressure (instrument)</th>
<th>Diastolic blood pressure (instrument)</th>
<th>Oxygen Peripheral Saturation (instrument)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chermont et al.(^8) Brazil</td>
<td>After the intervention and placebo phases (intergroup analysis).</td>
<td>NIV: 99±4 bpm Placebo: 91±4 bpm ps003 Non-specified instrument</td>
<td>NIV: 129±6 mmHg Placebo: 136±8 mmHg ps01220</td>
<td>NIV: 60±5 mmHg Placebo: 58±5 mmHg ps00312 Non-specified instrument</td>
<td>Not assessed</td>
</tr>
<tr>
<td>Lima et al.(^9) Brazil</td>
<td>Before and after for the intervention group and for the control group (intra and intergroup analysis).</td>
<td>NIV: 996±13 bpm Control: 1178±193 bpm ps0086 Non-specified instrument</td>
<td>Not assessed</td>
<td>NIV: 140±12.6 mmHg Control: 150±8.9 mmHg ps0145</td>
<td>Not assessed</td>
</tr>
<tr>
<td>O’Donnell et al.(^10) Canada</td>
<td>After the intervention and placebo phases (intergroup analysis).</td>
<td>SP: 119±7 bpm NS CPAP: 122±8 bpm NS Control: 127±7 bpm Electrocardiographic Monitor (Cardiovit CS-6/12Z; Schiller, Baar, Switzerland).</td>
<td>Not assessed</td>
<td>SP: 97±0.3% NS CPAP: 975±0.2% NS Control: 975±0.3% Pulse oximeter (503 pulse oximeter, Criticare Systems, Wakesha, WI)</td>
<td>Not assessed</td>
</tr>
<tr>
<td>Wittmwe et al.(^11) Brazil</td>
<td>At days 0, 4, 9, and 14 for the intervention and control groups (intragroup analysis).</td>
<td>Not assessed</td>
<td>Not assessed</td>
<td>Not assessed</td>
<td>Not assessed</td>
</tr>
</tbody>
</table>

NIV: noninvasive ventilation; SP: support pressure; CPAP: continuous positive pressure.
as the randomization or allocation secrecy, since it decreases the probability that expectations of the investigators interfere in the real benefit of the treatment. All included studies in this review performed participants’ blinding studies. Blinding of examiners and evaluators of the outcomes, which may be performed in all analyzed studies, was only done in two papers. This item is necessary since it prevents execution and mensuration bias.

Only one clinical trial performed sampling calculation, which was based on literature data. Sample calculation is important for inferences and extrapolations of the results found for the general population. Only one of the included articles was classified as pilot study in the title. Therefore, more controlled and randomized clinical trials are needed with performance of sampling calculation to ensure a sufficient power, as well as a higher external validity of the results that were found.

Another important item to minimize the overestimation of results is the analysis with treatment intention, which compares patients in the group where they were primarily allocated, regardless of the sample loss. Use of this analysis is only possible when there are complete data with regard to all randomized subjects. Only one study reports patient loss. However, the patient’s inclusion and exclusion flowchart was not exposed in any analyzed paper. Absence of clarity in the exposure of subjects’ inclusion and exclusion and in the results concerning each outcome leads to a bias risk. Absence of data is a relevant bias in the effect estimation, which established the outcome selective description. This happens because studies with positive results are more easily published. Therefore, it is needed clarity and objectivity in the demonstration of results on outcomes, on the conduction of the volunteer during the research, as well as proper data treatment, which must be performed between groups after the intervention, but not before and after in the same group.

After qualitatively assessing the results presented, we observed a fragility in the internal validity of the studies, since many applied strategies, which could be used to minimize bias risk, did not do it correctly. Decision-making is based on the current concept of practice based on evidence, which requires choice of better evidence levels together with the clinical experience of the professional and patient’s choice. Thus, studies with more methodological scope need to be performed in order to better fundament the use of NIV as a coadjuvant therapy with physical exercises in HF subjects.

CONCLUSION

The present systematic review was inconclusive as to the efficacy of NIV in the increment of exercise tolerance in subjects with HF. From the assessed studies, it was concluded that there is a low methodological quality due to the high bias risk in the available studies about the present theme, thus it was not possible to recommend NIV in this context. Therefore, we recommend the performance of randomized and controlled trials following a stricter methodological criterion, with appropriate sampling power, which controls allocation secrecy and evaluator’s blinding. It is also needed more clarity as to data exposition in order to make feasible the analysis of the size effect of treatment for critical and important outcomes.

REFERENCES


