The use of microcirculatory techniques in the assessment of pathophysiology, diagnosis and management of critical limb ischemia

degraaff, J.C.

Citation for published version (APA):
degraaff, J. C. (2003). The use of microcirculatory techniques in the assessment of pathophysiology, diagnosis and management of critical limb ischemia

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 5

Design issues of a diagnostic randomized clinical trial; the value of toe and transcutaneous oxygen pressure measurements in the management of critical leg ischemia

Jurgen C. de Graaff, Dirk Th. Ubbink
Jan G.P. Tijssen, Dink A. Legemate

Submitted

Abstract

Objective The value of a new diagnostic test is usually performed by analyzing its accuracy in relation to a reference standard. Such a comparison is not possible in case a reference standard is not available and is, in principle, incomplete. Here we describe a potentially better - but still not very common - model of diagnostic research, viz. a diagnostic randomized clinical trial (D-RCT), and discuss its pros and cons.

Methods Here we use a trial investigating the diagnostic and therapeutic management of critical limb ischemia as an example. Patients clinically suspected of critical limb ischaemia will be randomized either for the conventional management strategy (treatment on the basis of ankle blood pressure, duplex and angiography) or the new strategy (based on transcutaneous oxygen and toe pressures). The impact of the diagnostic work-up on the diagnostic and therapeutic process and clinical outcome will be evaluated.

Results A D-RCT is suited when a true reference standard is lacking and two different concepts are to be compared. Furthermore, such a randomized clinical trial is the best available research method to control for confounding and bias. Besides, a D-RCT not only evaluates the effect on diagnosis, but also incorporates the total effect on clinical outcome, like side effects of the tests, use of the results by the physician, and side effects of the therapy. However, the D-RCT has some disadvantages as to the power and size of the trial, and the influence of treatment on the outcome parameter.

Conclusions Taking these considerations into account, a D-RCT can provide valuable information as to the evaluation of diagnostic tests. However, more experience with this new concept is required.
Introduction

Traditionally, the introduction of new diagnostic techniques is mainly based on the evaluation of the test performance in relation to a reference (or gold) standard. However, such an evaluation is not always possible and, in principle, incomplete. In many cases correct diagnosing in relation to a reference standard is not the ultimate goal. This goal is to improve the clinical outcome by enhancing the appropriate choice of treatment. Ideally diagnostic tests are evaluated in terms of their effect on improving clinical outcome, by means of a diagnostic randomized clinical trial (D-RCT).

In therapeutic research, the randomized clinical trial is generally accepted as the ultimate type of evaluation. However, the use of this research design in the evaluation of diagnostic procedures is less well established. In this paper we describe the rationale, theoretical considerations, design issues, limitations, and solutions of a randomized clinical trial in the evaluation and implementation of diagnostic tests, with the diagnostic and therapeutic issues around critical limb ischemia as an example.

Description of the clinical problem

Critical limb ischemia is a controversial concept. The question which critical limb ischemia requires intervention when is still under debate. According to the TransAtlantic Inter-Society Consensus (TASC) on the Management of Peripheral Arterial Disease the term critical limb ischemia can be used for all patients with chronic ischemic rest pain, ulcers, or gangrene attributable to objectively proven arterial occlusive disease. However, these clinical characteristics are not specific for critical limb ischemia but could also be caused by other diseases. Particularly in patients with multifactorial disease (peripheral vascular disease in combination with e.g. diabetes and/or venous insufficiency) the decision and timing of vascular intervention is difficult. Therefore, objective criteria to functionally quantify the degree of ischemia are necessary for the definition of critical limb ischemia. Various consensus documents have formulated criteria based on ankle blood pressure. However, objective criteria have never achieved a consensus. At present, diagnostic angiography or duplex scanning is frequently performed to evaluate the extent of vascular obstructions. The limitation of these investigations is that they do not provide functional information about tissue perfusion but only anatomical information, which is not always related to the severity of clinical symptoms and the clinical outcome of the disease.

Therefore, the indication for vascular intervention is usually based on the skilled view of a vascular specialist. The timing of vascular intervention is difficult, since vascular intervention is accompanied with serious side effects (risk of operation, wound infections, early occlusion, which may ultimately result in amputation or death). On the other hand, postponing an intervention may also be detrimental to the patient. Moreover, critical limb ischemia has a considerable impact on health-related quality of life. This leads to substantial variation between diagnostic procedures and indications for vascular intervention among vascular surgeons, hospitals and countries.

Hence, we are in want of simple objective criteria to identify when a vascular intervention is required. Previous retrospective investigations have shown that a combination of toe blood pressure (TP) and transcutaneous oxygen pressure (TcpO₂) measurements might be a good indicator for vascular intervention. Both are simple and quick tests that provide functional information about the
peripheral tissue perfusion and it is the disturbance in peripheral (micro-) circulation which causes the clinical signs and symptoms.\textsuperscript{6,14-16} However, the attributable value of TP and TcPO\textsubscript{2} in the identification of patients for vascular intervention has never been evaluated.

**Phases of the diagnostic process**

Over the past years, new insights have emerged to improve the evaluation of new diagnostic tests.\textsuperscript{1} Yet, methodological standards for a research design of (new) diagnostic techniques are not well established. According to Guyatt and colleagues, diagnostic technologies should be evaluated for: [1] technological capability, [2] range of possible use, [3] diagnostic accuracy, [4] impact on diagnostic thinking, [5] therapeutic impact and [6] patient outcome.\textsuperscript{17} More recently an evaluation in terms of costs has been advocated.\textsuperscript{18} Although important, it is beyond the scope of this article.

The first aspect, *technological capability*, encloses the precision of measuring a phenomenon, which includes reliability and, thus, the reproducibility and technical performance of a test at various times, places, equipment, and operators.

The second aspect concerns the exploration of the *range of possible uses* of a new diagnostic technique to provide important diagnostic information in various clinical situations and diseases. This includes the fact-finding phase in which the test must be correlated to as much clinical information as possible.

The third aspect, *diagnostic accuracy*, comprises the independent comparison with a reference standard using ROC analysis and calculation of likelihood ratios. When a reference standard is lacking, the consensus of a panel of experts, the combination of diagnostic tests, or the long term follow-up could serve as a substitute reference standard.\textsuperscript{10}

The fourth aspect concerns the *impact* of the outcome of a test combination on the diagnostic process (also called *diagnostic thinking efficacy*). The aim of diagnosing is to reduce the uncertainty of the physician, which is influenced by the result of a test. The impact on the diagnostic process addresses how the outcome of a test influences the judgment of the physician about the presence or absence of a disease. This is different in that the order of the tests and contribution of each test to the other (including clinical judgment of the physician) is evaluated, as opposed to the evaluation of the diagnostic accuracy of a single test. This phase is difficult to measure individually, but can be evaluated by measuring pre and post test diagnostic probabilities of a physician (the probability of the presence of the diagnosis).\textsuperscript{18}

The *therapeutic* (fifth) aspect evaluates whether the test actually influences the treating physician’s choice of treatment. To change morbidity and/or mortality or improve quality of life, a diagnostic test must change therapy.\textsuperscript{17} Ideally, all performed diagnostic tests contribute substantially to the choice of accurate therapy. But usually this is not the case, and frequently a whole series of diagnostic tests is performed, part of which is used only to reassure the physician without actually affecting the choice of therapy. In this phase the choice of therapy ought to be the result of evaluation.

The most important (sixth) aspect of diagnostic tests is the *influence on patient outcome* (morbidity and mortality). Even though a test may be good, the test is useless or even contraindicated if the treatment it engenders does not alter clinical outcome. In this phase, the clinical outcome of a treatment strategy using the conventional diagnostic information is compared with that of a strategy taking into account of the new diagnostic information.
The value of TP and TcpO₂ in relation to the phases in the diagnostic process

The evaluation of most diagnostic procedures used in medicine meets the first three demands, but only few address all aspects.¹⁹ The ankle pressure, TP and TcpO₂ have been evaluated on the first two levels. The technological capability of the ankle pressure is well accepted and established.²⁰-²² The inter-observer reproducibility of TP and TcpO₂ in clinical practice have recently been evaluated and is acceptable and comparable with the ankle pressure.¹⁴ The evaluation of the range of possible uses has been studied extensively. The ankle pressure, TP, and TcpO₂ are well correlated to the severity of vascular disease, symptoms, risk of amputation, prediction of wound healing, and effect of limb salvage therapy.⁷,¹²,¹³,²³-²⁵

The evaluation with regard to the third aspect (diagnostic accuracy) is impossible because a true reference standard is lacking. When diagnostic information is employed in such a way that it directly influences treatment decisions, a reference standard by definition does not exist. Some alternatives have been proposed. Firstly, in natural history studies, the diagnostic tool is evaluated as a predictor of clinical outcome without intervention.²⁶ However, a bad outcome does not necessarily constitute an indication for therapy. Secondly, other studies have related TP and TcpO₂ to clinical outcome after therapy.²⁷-²⁹ However this may not correspond to the indication for vascular intervention. Thirdly, other studies have evaluated the diagnostic accuracy using the actual indication for vascular intervention as indicated by the vascular specialist blinded for TP and TcpO₂ as the reference standard. Against this method one could argue that the potential for improving conventional decision making (the reference standard) is ignored. From these studies one can only conclude that TP and TcpO₂ correspond best with the clinical intuition of the vascular surgeon.⁶,¹² Therefore, these methods can only be used to establish cut-off values but can never improve the clinical outcome.

Yet the continuing discussion of identifying critical limb ischemia demonstrates that there is no solution for the classical concept of determining diagnostic. However, it must be born in mind that new diagnostic methods are employed with a view to improve patient outcome. Therefore, the clinical outcome should be incorporated in the evaluation.

The fourth (impact on the diagnostic process) and fifth aspect (influence on treatment) are interesting for theoretical considerations but cannot be evaluated independent of the last and most important aspect: clinical outcome. The impact of the latter three aspects is paramount and can simultaneously be evaluated in a diagnostic randomized controlled trial (D-RCT), since the evaluation of the impact on diagnostic process and influence on treatment can only be evaluated in terms of the eventual best clinical outcome. Therefore, the present study was designed to evaluate the value of the combination of TP and TcpO₂ (new management strategy) in comparison with the conventional management strategy (ankle pressure, duplex scanning, angiography and expertise of a team of vascular specialists) in terms of their influence on the diagnostic process, on the choice of therapy and most importantly, on patient outcome.
Implementation of a D-RCT in the study design

Our study includes patients clinically suspected of critical limb ischemia by a vascular specialist and referred to the vascular laboratory. It excludes the so-called clear-cut cases (patients with obviously severe and mild disease) leaving only patients with an uncertain decision. Patients are randomly assigned to the conventional or the new management strategy (figure 1).

Under the conventional strategy, decisions for further diagnostic imaging of the arteries (primary duplex scanning and - if indicated - followed by angiography), are only based on clinical symptoms, physical examination, and ankle pressure. Additionally, the TP is only measured in patients with diabetes mellitus and incompressible arteries (ankle brachial pressure index > 115%) since withholding functional information about the severity of the disease is unethical and unacceptable to the vascular surgeons in our Department. Subsequently, the

Figure 1. Flow chart of the trial. Patient clinically suspected of critical leg ischemia are randomized for either the conventional or new management strategy. According to the conventional strategy the decision for further diagnostic imaging of the arteries (primary duplex scanning and - if indicated - followed by angiography), and thus the intention for a vascular intervention, are based on clinical symptoms, physical examination and ankle pressure (and TP only in patients with incompressible arteries) by the vascular surgeon involved. According to the new strategy the results of the combination of TP and TcpO₂ measurements determine the intention for vascular intervention. A vascular intervention and, subsequently, a duplex and/or angiography to define the type and place of intervention, are indicated only if one of the two measurements is below the cut-off level. I.e., in this strategy the therapeutic decision is based on objective parameters as opposed to the clinical eye of the surgeon in the conventional strategy.
proposed therapy (conservative treatment, bypass or percutaneous transluminal angioplasty) is discussed at a consensus meeting of the vascular surgeons, intervention radiologist, and vascular technologist.

Under the new strategy, the results of the combination of TP and TcpO₂ measurements will determine the intention for vascular intervention in addition to the clinical symptoms, physical examination, and ankle pressure. A vascular intervention and, subsequently, a duplex and/or angiography are indicated if either of two measurements is below the cutoff level (TP ≤30 mm Hg and/or TcpO₂ ≤35 mm Hg). Thus, the therapeutic decision in the new strategy is based only on objective parameters, as opposed to mainly subjective parameters in the conventional strategy group. The final decision for vascular intervention (indication and type) will be made at the vascular consensus meeting, after which every possible and clinical useful vascular intervention is performed until TP and TcpO₂ will exceed the cut-off value.

All patients are followed for eighteen months. During this period the diagnostic and therapeutic regime is based on the policy of the randomization result. Thus, the randomization settles the diagnostic procedure, which in its turn determines the therapeutic decisions. During the follow-up period reiteration of the assigned diagnostic procedure may imply a change in therapy.

Table 1. Power analysis based on the scores of the bodily pain subscale of SF-36

<table>
<thead>
<tr>
<th>conventional strategy</th>
<th>Test result</th>
<th>‘Truth’</th>
<th>BP score</th>
<th>Mean BP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intervention</td>
<td>conservative</td>
<td>30% x 70 + 20% x 30</td>
<td></td>
</tr>
<tr>
<td>intervention</td>
<td>30%</td>
<td>20%</td>
<td>50%</td>
<td>30% x 70 + 20% x 30</td>
</tr>
<tr>
<td>conservative</td>
<td>20%</td>
<td>30%</td>
<td>50%</td>
<td>20% x 30 + 30% x 70</td>
</tr>
<tr>
<td>Total</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
<td>54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>new strategy</th>
<th>Test result</th>
<th>‘Truth’</th>
<th>BP score</th>
<th>Mean BP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intervention</td>
<td>conservative</td>
<td>45% x 70 + 5% x 30</td>
<td></td>
</tr>
<tr>
<td>intervention</td>
<td>45%</td>
<td>5%</td>
<td>50%</td>
<td>45% x 70 + 5% x 30</td>
</tr>
<tr>
<td>conservative</td>
<td>5%</td>
<td>45%</td>
<td>50%</td>
<td>5% x 30 + 45% x 70</td>
</tr>
<tr>
<td>Total</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
<td>66</td>
</tr>
</tbody>
</table>

Patients include in this study have severe bodily pain (Bodily Pain sub score of the SF 36: BP = 30; 100 means no pain), as compared to the reference population (BP = 70), with an overall standard deviation of 20.9 Correct diagnosing (and consequently correct treatment) increases the bodily pain score to the same value as in the reference population (BP = 70), whereas incorrect diagnosing keeps the bodily pain at the same level as at inclusion (BP = 30). Furthermore, in the conventional strategy 60% (30%+30%) of the patients are correctly diagnosed, while this increases to 90% (45%+45%) in the new strategy.12 Thus, the mean bodily pain in the group of patients in the conventional strategy increases to 54, while the bodily pain in the new strategy group increases to 66. Subsequently, power analysis shows that 90 (2 x 45) patients must be included to refute the null-hypotheses that the diagnostic accuracy between the subgroups was the same (p-value < 0.05 is considered significant, two-sided; and 1-β = 0.80).
**Outcome parameters**
The primary end point is the change in pain measured by the bodily pain sub-score of the SF-36. Secondary endpoints are the change in the clinical situation as judged by (limb-) survival, amputation frequency, wound healing, and change in health-related quality of life, as investigated by the SF-36, and the number of diagnostic procedures performed. Amputations are classified in minor and major amputations as suggested by Rutherford. The outcome parameters are assessed at 1, 3, 6, 9, 12 and 18 month after inclusion of the trial. The statistical methods include ANOVA for Pain, quality of life and severity of wounds and Kaplan-Meier survival analyses for patient and limb survival.

**Power analysis**
Power analysis on the basis of the bodily pain sub-scale of the SF-36 revealed that 90 patients must be included to refute the null-hypotheses that the diagnostic accuracy between the subgroups is the same (details of the calculation and assumptions are shown in table 1).

**Discussion**
The ongoing problem of defining critical limb ischemia is a typical research question where traditional methods of evaluation of diagnostic accuracy cannot be applied. As a result, we chose to evaluate the diagnostic value of TP and TcpO2 in terms of their effect on clinical outcome using randomization.

**Pros and cons of randomization**
A D-RCT as described here aims to prove whether the use of two simple, objective tests can improve patient outcome in comparison to the clinical eye of an (experienced) vascular specialist. The lack of a true reference standard makes a D-RCT especially suitable. Furthermore, a D-RCT also addresses a possibly harmful effect of diagnosis and therapy. This is particularly important in potentially harmful and/or invasive diagnostic procedures (such as angiography) and therapies with a high complication and failure rate (such as bypass surgery).

A randomized clinical trial is the best strategy to control for bias and confounding, and an effective strategy for an objective and controlled comparison. In addition, in a D-RCT two different concepts of diagnostic tests (in this particular instance anatomical information derived from the larger vessels versus functional information from the local, peripheral vessels) and a whole management strategy, including the effect on clinical outcome, can be compared. This is important as the decision for intervention (diagnosis) is not a snapshot, but a continuous interactive process and in which the vascular specialist continuously judges the clinical situation (e.g. pain, severity and progress of wound healing), new diagnostic findings (e.g. peripheral blood pressures or the extension of disorders as found by duplex scanning and angiography), comorbidity, and the risk of complications and failures of invasive therapy.

A D-RCT also has its limitations. A problem with the use of a D-RCT is that it might be unethical withholding a diagnostic test, especially when this test is already being used. This may be applicable to our situation. Although the TP is measured only rarely in hospitals in our country, withholding TP in patients with incompressible arteries is hard to accept for the physicians in our hospital and may therefore reduce their willingness to participate in this trial. Therefore, we
allowed TP measurements in patients with diabetes mellitus and incompressible arteries.

The incorporation of treatment in the outcome measurement introduces a so-called black box in which many factors (the efficacy and use of the test including subjective judgment of physicians) are involved, which could enhance or mitigate the value of the test.\(^1\) One aspect of the black box in D-RCT is that the outcome is dependent on the treating physician's capability of interpreting the role of the measurement. If, in theory, decision making with the new measurement does not lead to a better clinical outcome, this can be caused by two factors. Firstly, the measurement does not contain valid information. Secondly, the information is valuable, but the participating physicians do not know how to utilize the information. Differentiation between these aspects is not possible in a D-RCT but can be evaluated in a retrospective diagnostic review.\(^{17}\) Therefore, D-RCT is not suited in the early phase of the evaluation of a diagnostic test, but is a valuable adjuvant in the end-stage of the evaluation of diagnostic tests.

Judgment of the effect of a test in a D-RCT also depends on the effects of the chosen therapy on the outcome parameter. In other words, the effect of the test on the patient outcome is established through the effect of therapy. This is important in situations when adverse effects of the therapy have a large impact, when therapy is unsuccessful in many patients, or when the disease is self-limiting in time and withholding therapy does not largely influence the course of the disease.\(^{19}\) Moreover, a test is without value if current therapy provides no benefit. This does not imply that the application of future therapy may benefit from the diagnostic knowledge of the test. In other words, the diagnostic value of a test is determined by the therapeutic option used in the study. One could argue that the best possible test identifies only those patients who benefit from a specific therapy, independently of the 'diagnosis'. The total effect of all these factors on the major outcome parameters is difficult to analyze in the studies performed thus far and can best be analyzed in a D-RCT.

**Sample size**

The total influence of factors in the black box on the outcome parameter could be considerable and to determine empirically the effect of tests per se can therefore require a very large sample size. The size and power of a trial are the limiting factor of a D-RCT. The expected difference in clinical outcome between the groups of randomization defines the size and power of the D-RCT. As stated before, the clinical outcome is influenced by the effect size of therapy. Secondly, the result of an evaluation depends on the prevalence of the disease in the population of the trial. In populations with a low indication for therapy, a considerable number of patients should be included to show significant differences between diagnostic strategies. Therefore our study includes patients with a clinical suspicion of critical limb ischemia, with exclusion of clear-cut cases. Inclusion of borderline patients makes it more likely to detect a difference between the randomization groups, since it is not likely that different management strategies result in a different clinical outcome in undisputable (clear-cut) cases.

Furthermore, it is important to evaluate the value of a test in patients representative of those in whom the test will be applied in clinical practice.\(^{31,32}\) The clinical usefulness of a test is underestimated if the study is carried out in a population of patients with a clear cut indication for vascular intervention, or in patients who do not need an operation anyway. Since the value of a (new) test is not likely to change therapy and to alter clinical outcome in these group of
Figure 2. Scheme of research models. In the first research model (2a) all patients are randomized. However, differences in outcome will only be observed if the two test results disagree. This model requires larger sample sizes (or reduces the power of the trial). In the second (discordant) research model (2b) only patients in whom the two test results differ are randomized. The possibility of a different outcome between the two strategies is thereby increased.

Figure 2a

Figure 2b
patients. The evaluation of the tests in a group of ‘borderline’ patients also makes our patient group more representative for the clinical problem and reduces the number of patients required to detect a significant difference.

**Alternative design**

Our study includes all patients, and randomizes them for two management strategies; with or without the new diagnostic test (TP and TcpO₂). When comparing two diagnostic modalities one could also choose to perform both tests in all patients and to randomize only those with discordant results (figure 2).  

Theoretically we could have followed this design principle by recording a treatment decision before and after TP and TcpO₂. Subsequently, we could have randomized only those patients with discordant results. The advantage is that fewer patients have to be included. However our problem implies a chain of two management strategies with prolonged decision making. Therefore our problem lacks a single decision moment, which is a prerequisite for randomization of patients with discordant results.

In conclusion, a D-RCT is the study design of choice for the evaluation of TP and TcpO₂ in the management of critical limb ischaemia. It settles the problem of the lacking reference standard, and moreover it provides information on choice of treatment and patient outcome.

**Reference List**


