INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

CLINICAL PREDICTORS OF DEEP VEIN THROMBOSIS IN PATIENTS WITH LEG SYMPTOMS

Susan Rebecca Kahn, MD
Department of Epidemiology and Biostatistics
McGill University, Montreal

November 1996

A Thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements of the degree of
Master of Science

Abstract

Clinical Predictors of Deep Vein Thrombosis in Patients with Leg Symptoms

Background: Deep vein thrombosis (DVT) is a common condition with significant mortality and morbidity. Proximal DVT is more often associated with pulmonary embolism and the post-phlebitic syndrome than calf DVT. Identifying which clinical variables predict DVT and proximal DVT could be useful for the effective targeting of diagnostic tests for DVT.

Purpose: To determine, in patients presenting with leg symptoms, which clinical variables best predict 1) DVT and 2) proximal DVT. To estimate the probability of DVT in an individual presenting with a particular grouping of these variables.

Methods: *Design:* A diagnostic test study design was used to develop a clinical prediction index. *Setting:* University teaching hospital inpatient and outpatient departments. *Population:* A series of 271 patients undergoing diagnostic procedures for a first episode of clinically suspected DVT at the Montreal General Hospital in 1989-90. Cases (n=73) were patients with DVT diagnosed by contrast venography (CV) (n=68), or if contrast venography could not be performed, by impedance plethysmography (IPG) (n=5). Controls (n=198) were patients from the same population who were free of DVT by CV (n=86) or IPG (n=112). Among cases, 52 patients (71%) had proximal DVT. *Measurements*: At baseline, information was collected on the following variables: age, sex, medical and surgical history, trauma and immobilization, and symptoms and signs. *Statistical analysis:* Case-control analysis, where cases were patients with DVT (primary analysis) or proximal DVT (secondary analysis). The univariate analysis identified individual variables associated with case status. The multivariate analysis used logistic regression with Bayesian model selection strategies to estimate a model that best predicted case status. The model was used to develop a clinical prediction index for DVT.

Results: Male sex, orthopedic surgery, and warmth and superficial venous dilation on exam were independent predictors of DVT (adjusted odds ratios and 95% confidence

intervals 2.8 [1.5, 5.1], 5.4 [2.2, 13.6], 2.1[1.2, 3.9] and 2.9 [1.4, 5.7], respectively) and proximal DVT (adjusted odds ratios 2.4 [1.2, 4.8], 4.1 [1.4, 12.3], 2.3 [1.2, 4.7] and 3.4 [1.6, 7.0], respectively) . A clinical prediction index that categorized patients into different levels of DVT risk was created, and its ROC curve showed moderate predictive ability. No single cutoff point was ideal in terms of desired sensitivity and specificity, however the index was useful in a strategy aimed to limit the need for contrast venography in patients with suspected DVT. Using this strategy, 78% of study patients could have avoided contrast venography.

Conclusions: Male sex, orthopedic surgery, warmth and superficial venous dilation are independent predictors of DVT and proximal DVT. In this population, a clinical prediction index that included these predictors was useful in choosing the optimal diagnostic test for patients with suspected DVT. This index should be evaluated prospectively in a larger population.

Abrégé

Prédicteurs Cliniques De La Thromboplébite Chez Le Patient Avec Des Symptômes Aux Jambes

Introduction: Les thrombophlébites profondes sont fréquentes et elles sont une cause importante de mortalité et de morbidité. Les thrombophlébites proximales sont plus souvent la cause d'embolie pulmonaire et du syndrôme post-phlébitique que la thrombophlébite au mollet. L'identification de variables cliniques capables de prédire le dévelopement de ces différents types de thrombophlébite pourrait être utile pour déterminer quel test utiliser pour fins diagnostiques.

Objectifs: Identifier, chez les patients se présentant avec des symptômes aux jambes, les variables cliniques qui prédisent le mieux 1) le dévelopement des thrombophlébites en général et 2) le dévelopement des thrombophlébites proximales en particulier. Estimer la probabilité d'une thrombophlébite chez un individu qui se présente avec un groupe particulier de ces variables.

Methodes: Type d'étude: Étude de test diagnostic dans le but de développer un index predictif clinique. Lieu de l'étude: Départements ambulatoires et hospitaliers d'un hôpital d'enseignement universitaire. Population: Un groupe de 271 patients ayant eu des procédures diagnostiques pour une première manifestation d'une thrombophlébite possible qui se sont présentés à l'Hôpital général de Montréal. Les cas (n=73) sont les patients avec un diagnosis de thrombophlébite tel que déterminé par un venogramme de contraste (n=68), ou, si ce dernier ne pouvait être fait, par pléthysmographie d'impédance (n=5). Les témoins (n=198) sont les patients venant de la même population qui n'avaient pas de thrombophlébite tel que déterminé soit à l'aide d'un venogramme (n= 86) ou d'une pléthysmographie (n=112). Parmi les cas, 52 patients (71%) souffraient d'une thrombophlébite proximale. Mesures: Les variables suivantes ont été documentées: age, sexe, histoire médicale et chirurgicale, de trauma et d'immobilisation et symptômes et

signes. Analyse Statistique: Dans l'analyse première, les cas comprennent tous les patients avec une thrombophlébite et dans l'analyse secondaire seulement les cas avec une thrombophlébite proximale sont inclus. L'analyse univariée aide à identifier les variables associées au status de cas. Tandis que l'analyse multivariée consiste en une régression logistique utilisant une stratégie de sélection de Bayes pour identifier le modèle qui prédit le mieux le status de cas. Ce modèle est utilisé pour construire un index clinique pour prédire la présence d'une thrombophlébite.

Resultats: Le sexe mâle, avoir subi une chirurgie orthopédique et de la chaleur ainsi qu'une dilatation des veines superficielles à l'examen physique étaient indépendamment associés avec la présence d'une thrombophlébite (odds ratios ajustés et intervales de confiance de 95%: 2.8 (1.5-5.1), 5.4 (2.2-13.6), 2.1 (1.2-3.9) et 2.9 (1.4-5.7), respectivement et avec la présence d'une thrombophlébite proximale (2.4 (1.2-4.8), 4.1 (1.4-12.3), 2.3 (1.2-4.7) et 3.4 (1.6-7.0), respectivement. L'index clinique de prédiction catégorisant les patients dans différents niveaux de risque fut crée et la courbe receveur-opérateur a demontré une abilité à prédire le diagnostique. Aucun point n'était idéal en ce qui a trait à la sensitivité et à la specificite. Cependant, l'index s'est avéré utile comme stratégie ayant pour but de limiter l'utilisation d'un venogramme chez les patients avec une thrombophlébite possible. En utilisant cette stratégie, un venogramme peut être évité chez 78% de ces patients.

Conclusions: Le sexe mâle, avoir subi une chirurgie orthopédique et de la chaleur ainsi qu'une dilatation des veines superficielles à l'examen physique sont des prédicteurs indépendants de la présence de la thrombophlébite en général et de la présence d'une thrombophlébite proximale en particulier. Dans cette population, un index clinique de prédiction incluant ces prédicteurs était utile pour choisir le meilleur test diagnostique pour les patients chez qui l'on suspecte une thrombophlébite. Cet index devrait être évalué prospectivement dans une plus grande population.

Preface

Thromboembolic disease, an area of medicine which encompasses deep vein thrombosis and pulmonary embolism, has been the subject of an explosion of research in the last 10-15 years, with new developments in (a) *classification* (e.g. the recognition that deep vein thrombosis and pulmonary embolus are both located on the same disease process continuum), (b) *etiology* (e.g. the identification of risk factors for the disease, including the hypercoagulable states), (c) *diagnosis* (e.g. development and validation of non-invasive tests), (d) *prevention* (e.g. preventive anticoagulation which has been shown to be both safe and effective for high risk groups) and (e) *treatment* (e.g. the development and use of low-molecular weight heparin compounds). Despite these impressive advances, however, numerous questions continue to surround this disease, and more research is required to answer them.

In responding to this need, I believe that the present project represents a particularly good fit between "the researcher" and "the researched". As an internist, I have had a long-standing clinical interest in thromboembolic disease. More recently, as a candidate for a Masters degree in Epidemiology and Biostatistics, I have found myself drawn toward the study of this challenging disease, a disease which not only intersects with so many medical subspecialties (hematology, cardiology, and vascular medicine, to name a few), but is also highly significant in terms of its international prevalence and its impact on affected individuals.

Acknowledgments

I wish to acknowledge the excellent guidance of my thesis supervisors, Dr. Lucien Abenhaim and Dr. Lawrence Joseph. I thank Dr. Abenhaim for his mentorship over the past three years, and for helping me learn how to see the forest for the trees. I am grateful to Dr. Joseph for encouraging me to take ownership of my statistical analysis, thereby ensuring a basic level of competence in biostatistics that will serve me well in the future.

I thank Dr. Jacques Leclerc for providing me with the database that allowed me to undertake this work, and for his helpful comments in the early planning stages of this project. I also thank my colleague and husband Dr. Michael Libman for his helpful critique of parts of this work.

I am grateful to my mother, Joan Kahn, for teaching me by example that learning is a lifelong pursuit, and for always encouraging my academic explorations. Finally, this work would not have been possible without the support and good humor of the four men in my life: Michael, Max, Eli, and Benjamin.

Table of Contents

List of figures, tables and graphs

Chapter 1

Chapter 2

Chapter 3

Chapter 4

1. Introduction

In this thesis I have undertaken to study the various symptoms, signs and baseline characteristics by means of which a physician is able to distinguish between patients who have a deep vein thrombosis and those who do not. By way of introducing this study, I shall review some of the basic aspects of venous thromboembolic disease or VTE, which is an "umbrella" term that includes both deep vein thrombosis (DVT) and pulmonary embolism (PE).

DVT refers to the presence of blood clot in the lumen of the deep veins of the body. It is a common condition which typically occurs in the veins of the legs. Less typically, it may also involve the veins of the arm, the brain, and the abdomen. As a consequence of the complete or partial obstruction of the vein by clot, DVT may produce a variety of acute, localized symptoms and signs, such as pain, swelling and redness. DVT, however, can also be silent.

DVT of the calf veins affects the peroneal or posterior tibial veins of the calf. DVT of the proximal veins affects the popliteal, superficial femoral, common femoral, external iliac or common iliac veins (see Figure 1, end of chapter). Calf DVT may undergo spontaneous lysis, may scar down within the calf vessels, or may extend proximally. Untreated calf DVT extends to the proximal veins with a rate of 25-30% (1-3).

PE is a life-threatening complication of DVT. It occurs when a fragment of clot in the deep vein breaks off and travels (embolizes) into the pulmonary circulation. If the embolus is large, respiratory failure, cardiac failure and sudden death can occur. Smaller PE may cause breathlessness, hypoxia, or may also be silent. In rare cases, chronic multiple pulmonary emboli can occur, which ultimately lead to pulmonary hypertension and right heart failure. Proximal DVT coexists with or leads to PE in approximately 50% of the cases, with an associated 25% mortality, whereas calf DVT rarely leads to symptomatic PE, but may be associated with asymptomatic, clinically unimportant PE in 20% of cases (4,5).

Accurate and timely diagnosis and treatment of DVT are important. Early

treatment of DVT with anticoagulants has been demonstrated to (a) reduce the incidence of pulmonary embolism and its associated mortality (6,7); (b) relieve acute symptoms in the leg; and (c) prevent extension of DVT from calf veins to more proximal veins. This latter effect of treatment avoids the worsening of symptoms, reduces the incidence of pulmonary embolus, and limits the extent of vein wall damage which, over time, can lead to chronic venous insufficiency, also referred to as the post-phlebitic syndrome. Ensuring an adequate duration of treatment (e.g. 3-6 months) prevents early recurrence of DVT (8) and may decrease the incidence of the post-phlebitic syndrome. Failure to diagnose and treat DVT can lead to chronic pulmonary thromboembolic disease, pulmonary hypertension and right heart failure.

Equally important to diagnosing DVT in patients with the disease is correctly identifying those who do not have DVT. The implication of a diagnosis of DVT is generally a 7-10 day hospitalization for administration of intravenous anticoagulation, followed by oral anticoagulation for at least six weeks and sometimes for life. Intravenous heparin anticoagulation is associated with a 5% risk of bleeding (9). Oral anticoagulation is associated with a 5-20% risk of bleeding (9) and requires frequent blood test monitoring and clinic visits. Also, because *prior* DVT is an important risk factor for *future* DVT, falsely labeling a patient with this diagnosis will result in needless anxiety and unnecessary tests each time he experiences leg symptoms. Furthermore, a false positive diagnosis of DVT in the case of women of childbearing age has special implications, for it is currently considered prudent, as a preventive measure, to treat women who have had a previous DVT with injected anticoagulants during pregnancy, a time of relative high risk for VTE disease. Not only is this inconvenient and uncomfortable, but it is also associated with complications both in the short-term, such as bleeding, and in the long-term, such as osteoporosis (10). One can appreciate, therefore, that correct classification of patients presenting with symptoms of DVT is important.

A diagnosis of DVT is achieved via a variety of invasive and non-invasive tests. Contrast venography (CV), which requires injection of contrast dye into a vein of the affected limb, is considered to be the reference test or "gold standard" for the diagnosis of DVT. However, its limitations include the potential to cause adverse reactions, its

invasiveness, and its cost.

Owing to these constraints, there has been an increasing interest over the past 20 years in the development of non-invasive methods to diagnose DVT. These include impedance plethysmography (IPG) and duplex compression ultrasonography (CUS). Aside from the obvious advantage of being non-invasive, these tests, when compared to CV, are both lower in cost and easier to apply. Their disadvantages, however, are their inability to visualize the entire venous system of the limb, the dependence of test performance on the operator, the lack of test availability in certain centers, and for some, the need for serial testing in the case of a negative result.

Symptoms suggestive of DVT are extremely common in the general population, and have a multitude of possible causes. By estimating the pre-test probability of DVT in an individual patient with leg symptoms, therefore, the choice and interpretation of diagnostic tests, as well as the subsequent decision-making regarding treatment, is likely to be greatly assisted. Such pre-test estimations would also help ensure the most rational and cost-effective use of diagnostic tests for DVT.

The set of issues generated above have prompted the following research questions, each of which will be addressed in this thesis:

A. In patients presenting with leg symptoms suggestive of DVT, which clinical variables at baseline predict the presence of deep vein thrombosis?

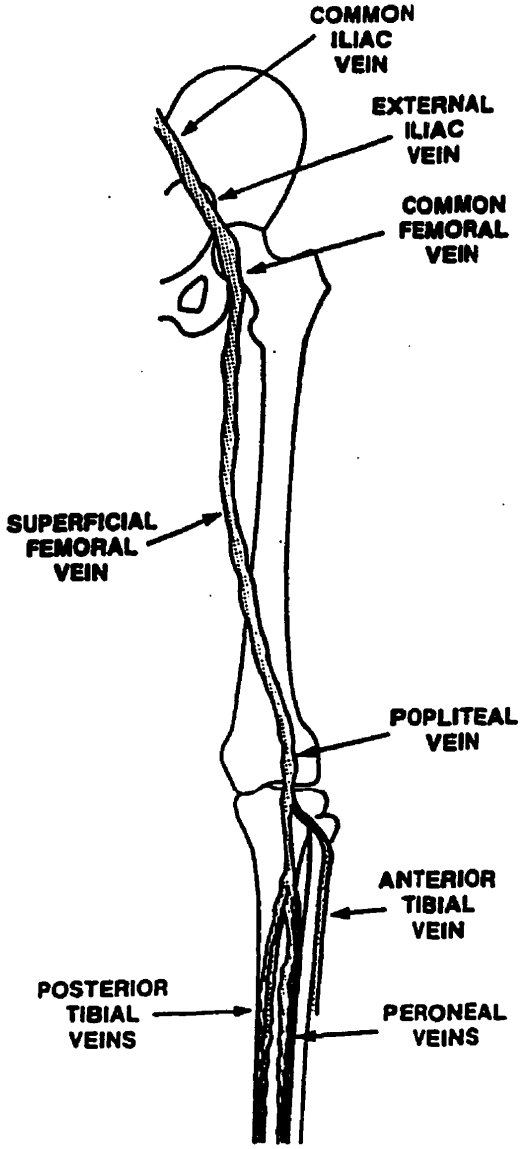
B. Among the clinical variables, which if any are particularly useful in predicting proximal vein DVT?

C. In an individual presenting with a particular grouping of these variables, can one accurately estimate the probability of deep vein thrombosis?

These questions were investigated within a pre-assembled series of patients, all of whom had participated in a randomized clinical trial in which the use of IPG and CV were compared in relation to their effectiveness in diagnosing DVT (IPG-CV study). At study

entry, information on a number of clinical variables was collected. Using a case-control analysis I studied the relationship of these clinical variables to the outcome. The primary outcome of interest was any DVT, and the secondary outcome of interest was proximal DVT. I then developed a predictive model to estimate the probability of DVT in an individual patient presenting with suspected DVT.

Figure 1. Anatomy of the deep venous system of the leg, posterior view



(borrowed from Leclerc (11), with permission)

This selective review of the literature of venous thromboembolic disease will focus on the incidence and known risk factors for DVT, the test characteristics of diagnostic tests for DVT, and the accuracy of both clinical symptoms and signs and published clinical prediction models for the diagnosis of DVT. This is the background required for the chapters to follow.

2.1. Incidence of deep vein thrombosis

2.1.1. Methodological difficulties

The true incidence of DVT in the population is difficult to assess for a number of reasons. Autopsy studies have not been useful in determining the frequency of DVT since DVTs themselves are not often fatal. Also, autopsy rates in general are low and overrepresent unusual cases. Most studies looking at DVT incidence have been performed in hospitalized patients, which overestimates the general incidence for two main reasons: hospitalized patients are at higher risk for DVT for a variety of reasons (see below), and in many of these studies, tests for DVT were performed in the absence of symptoms suggestive of DVT. It is uncertain how many of these subclinical DVTs would have become symptomatic and reached medical attention had they not been actively sought out.

Older community-based studies relied on clinical symptoms and signs to diagnose DVT, without the benefit of objective testing. This was problematic because less than half of patients suspected of having DVT had the diagnosis confirmed when objective tests were performed (12). Conversely, since symptoms can be vague or even absent, DVT may be underdiagnosed, especially among outpatients. Thus, clinical diagnosis may both overestimate and underestimate the true incidence of DVT. Ideally, for symptomatic DVT, one would need a totally captive population with 100% referral for suspected DVT using objective tests such as venography, IPG or CUS. For the diagnosis of asymptomatic DVT, however, these tests, although reasonably accurate, are too costly to use for mass surveillance of outpatients at low risk for DVT.

2.1.2. Incidence studies

In the Tecumseh Health study (1973), an 11-year longitudinal study of healthy individuals conducted in Tecumseh, Michigan using data provided by history and physical examination, estimates of the incidence of DVT were approximately 10.0 and 14.8 per 10,000 40-49 year old women and men respectively, which increased to 31.0 and 49.9 per 10,000 70-79 year-old women and men respectively (13). The main weakness of this study was the unavailability of objective testing, with complete reliance on clinical diagnosis.

In the United States, data from Vital Statistics and from the National Hospital Discharge survey, both based on hospital discharge diagnoses from 1970-1985, showed an age-adjusted rate for DVT (encompassing phlebitis and thrombophlebitis) of 79 per 100,000 and that for PE of 51 per 100,000 (14). Due to the method of data collection, it is not known how many of these were objectively verified with diagnostic testing.

A community-wide study conducted in 16 short-stay hospitals in Worcester, Massachusetts that has a catchment area of almost 400,000 predominantly white individuals retrospectively examined the incidence and case-fatality rates of DVT and PE in hospitalized patients over an 18 month period (15). Ascertainment of outcome was achieved by the use of International Classification of Disease codes that were likely to encompass most hospital-diagnosed venous thromboembolic conditions, with individual review of charts for objective diagnosis of DVT or PE, which was documented in 84% of cases with DVT and 61% of cases with PE.

The annual incidence of DVT was 48 per 100,000, while the incidence of PE with or without DVT was 23 per 100,000. Rates were higher in men than women, and increased with age. In-hospital case-fatality rates for DVT and PE were 5% and 23% respectively. The study only examined data from short-stay hospitals, hence cases arising from long-term facilities, rehabilitation centers and the general out-patient community were excluded, as were asymptomatic patients who did not seek medical attention. The authors estimate that there are approximately 170,000 patients with first-time VTE and 90,000 cases of recurrent VTE treated in short-stay hospitals in the U.S. each year, resulting in a minimum of 13,000 deaths each year. By extrapolation, and taking into account the almost certain underestimation of the true incidence of VTE, this likely

represents 600,000 cases in the general population overall.

A similar survey looked at the incidence of DVT in the region served by Malmo General Hospital in Sweden, which represented a population base of 281,000 people (16). This study also relied on hospital-based diagnoses, used a positive contrast venogram to define outcome and was thus oriented to symptomatic DVT, so cannot be called a true population survey. However, there was only one department at one hospital performing venography, and the patient population included both inpatients and referred outpatients, hence the figures for symptomatic DVT were likely to be quite accurate. The incidence of DVT was 160 cases per 100,000 per year, which included recurrent cases and cases associated with PE. Incidence rates increased with increasing age, but there was no difference in incidence rates between men and women.

Finally, a study of a random 5% sample of Medicare claims in the U.S. over a 3-year period identified all cases of DVT and PE in the elderly using International Classification of Disease (ICD) codes pertaining to the diagnosis and treatment of VTE (17). Annual incidence rates of DVT were 180 per 100,000 at age 65-69, which increased to 310 per 100,000 at age 85-89.

The studies described above considered mostly hospital-related cases. It is not possible from the data provided to accurately estimate the incidence of DVT in the general population. However, these studies convincingly show that venous thromboembolic disease is a significant health problem which affects all ages and which exacts considerable morbidity and mortality.

VTE is also a costly problem. The total cost per patient for a correct diagnosis by venography and subsequent treatment of DVT was estimated at approximately $5,000 in 1980 Canadian dollars, a figure which would now be at least $10,000 (18). This figure does not include costs incurred by the estimated 40-50% of DVT patients who develop long-term sequelae such as the post-phlebitic syndrome (19).

2.2. Risk factors for deep vein thrombosis

The etiology of DVT can still best be conceptualized by Virchow's triad, described in 1860, which delineates the pathophysiological factors which promote the development of venous thrombosis, namely *vein wall damage*, *stasis*, and *hypercoagulability*. For example, hip surgery which causes vein wall damage may result in DVT. Prolonged bedrest or car travel, via insufficient pumping action of the calf muscles, leads to stasis of blood in the deep veins which can promote DVT. Finally, hypercoagulability, as occurs with certain cancers, medications such as estrogen, and inherited abnormalities of the intrinsic blood anticoagulant system may lead to DVT. In some cases, DVT may be caused by combinations of these factors, for example in pregnancy, where stasis due to pressure of the enlarging uterus on the iliac veins and hypercoagulability due to the effects of high estrogen levels occur together. In other cases, no particular risk factor for the development of DVT can be identified.

Most studies on risk factors for DVT have been conducted in hospitalized patients, in whom the incidence of DVT and patient characteristics are more easily determined than in the community, since objective tests are more readily available and clinical and laboratory information is closer at hand. Much epidemiological information has been provided by the numerous published clinical trials on the primary prevention of DVT in high risk situations, which have prospectively assessed the risk of DVT in selected hospital populations using strict diagnostic criteria. However, there has been difficulty in identifying individual risk factors, since hospitalized patients are a disparate group with multiple underlying pathologies and numerous potential risk factors for DVT, some iatrogenic, which may interact with or confound one another. Nonetheless, among hospitalized patients, the following groups have identified themselves as being at increased risk for DVT:

2.2.1. Surgical Patients

<u>Orthopedic surgery</u>

Patients undergoing major orthopedic surgery of the lower extremity represent the

highest operative risk group for DVT and PE. Pooled data from prospective clinical trials of thromboprophylaxis that required mandatory post-operative venography have shown that among patients in the untreated or placebo arms, there was a 51% incidence of DVT after total hip replacement, a 71% incidence after total knee replacement, and a 48% incidence after hip fracture surgery. The rates of proximal DVT in the above groups were 23-36%, 9-20%, and 17-36% respectively, and of fatal PE 3.4-6%, 0.7%, and 3.6-12.9% respectively (20). The high rates in these patients reflect the presence of numerous underlying factors that promote the development of DVT, namely immobility, vessel injury, and activation of coagulation pathways.

Trauma

Interpretation of the literature on DVT incidence in trauma patients is difficult because of the high proportion of trauma patients with hip or lower extremity fractures, and the overall heterogeneity of this group of patients.

In a recent large prospective study of patients admitted to a regional trauma unit in Toronto, DVT was diagnosed by venography in 201 out of 349 patients (58%), only 3 of whom had clinical features suggestive of DVT. The rate of DVT was 69% in those with lower extremity fractures, but there was still a 50% incidence of DVT in trauma patients whose injury only involved the chest, face or abdomen. Independent risk factors for DVT among the study group were older age, blood transfusion, surgery, fracture of the femur or tibia, and spinal cord injury (21).

General abdominal and other surgeries

Approximately 25-30% of patients undergoing elective major abdominal surgery show post-operative evidence of DVT when surveyed by serial ^{125}I fibrinogen leg scanning (FLS), a technique which is sensitive to calf and low proximal DVT but insensitive to high proximal DVT. In pooled data from trials where DVT diagnosed by FLS was confirmed with venography, this incidence rate was closer to 20% (20). For unclear reasons, the DVT incidence in North American trials is about one half that of European trials. Overall, in North American studies, among general surgery patients, the incidence of any DVT was 16%, proximal DVT 7%, PE 1.6%, and fatal PE 0.9%. The

more serious endpoints are likely underestimated, since most patients received anticoagulant treatment as soon as the surveillance test became positive (20).

Patients undergoing surgery for malignant disease have higher DVT rates than those without malignant disease (20). Among patients undergoing urologic surgery, data collated from seven trials documented a 41% incidence of DVT (22). Vascular surgery conferred a 23-34% risk of DVT, as shown by two prospective screening studies using FLS (23,24). DVT as diagnosed by FLS also occurred in 17.5% of patients undergoing major gynecological surgery. Among these women rates were highest in those with malignancy, a past history of DVT or previous radiation therapy (25).

Anesthesia

For a given type of surgery, the type of anesthesia administered can influence the incidence of DVT. McKenzie et al. noted among patients with hip fracture undergoing orthopedic procedures that 75% who received general anesthesia developed venographically-proven DVT, compared to 40% who received subarachnoid blocks (26). Similarly, for urological procedures, a 12% rate of DVT was noted in retropubic prostatectomy patients randomly allocated to receive lumbar epidural analgesia, compared to 52% of those who received general anesthesia (27).

2.2.2. Medical Patients

Overall, the risk of DVT in various subcategories of medical patients has been less well studied than for surgical patients.

Malignancy

Trousseau, in 1865, first suggested the association between DVT and abdominal malignancies. Since then, numerous studies have been published which confirm the association between VTE and malignancy in general, however precise rate estimates are not available. DVT risk is also increased among cancer patients undergoing active treatment with chemotherapy. A randomized clinical trial comparing two adjuvant

chemotherapy regimens clearly showed that all thrombotic events occurred during months that the patients were receiving chemotherapy (28).

Intensive care patients

Like trauma patients, intensive care patients represent a heterogeneous group in terms of risk factors for DVT, rendering interpretation of incidence rates difficult. A prospective ultrasound case series of 100 medical intensive care patients uncovered 33 cases of DVT despite DVT prophylaxis in 58%. Interestingly, there were no differences in age, diagnosis of cancer, recent surgery, or duration of hospitalization between patients with and without DVT (29).

Myocardial infarction and ischemic stroke

Using FLS, the overall incidence of DVT was approximately 24% in MI patients, and 42% in the weak or paralyzed limb of stroke patients. These rates are derived from the pooled placebo arms of trials evaluating preventative antithrombotic therapy in these patient groups (20).

2.2.3. Other risks for DVT

Other important risk factors for DVT which affect both hospitalized and ambulatory patients have been recognized.

Age

It has long been known, based on clinical experience, that DVT incidence increases with age. DVT is extremely rare in children, and as demonstrated in trauma and surgery patients, increases in incidence sharply after age 40. Both the Worcester (15) and Malmo (16) community-based studies demonstrated an increased incidence of DVT with age. In the former study, age-specific rates of DVT were 16 per 100,000 population in those aged 20-29, compared to 265 per 100,000 in those aged 70-79. A similar trend was seen for PE. In the Malmo study, age-specific DVT rates were 6 per 100,000 population in those

aged 20-29, compared to 564 per 100,000 in those aged 70-79. These figures may have been influenced by suspicion bias, i.e. overdiagnosis in the elderly and underdiagnosis in the young based on different levels of clinical suspicion.

Gender

Both the Tecumseh and Worcester incidence studies found that DVT incidence was approximately 1.5 times higher in men than women in each age stratum studied (13,15). However, the Malmo incidence study found no sex differences in DVT incidence (16). At present, there is neither convincing evidence that male sex is a risk factor for DVT, nor a plausible explanation for why men might be at higher risk than women.

Immobilization

The association between immobility, its duration, and VTE has been confirmed in a number of autopsy and clinical studies. An autopsy study of 253 patients demonstrated DVT in 15% of patients immobilized for less than one week, compared to 80% in those with longer periods of immobilization (30). Kierkegaard found that from the second to the eighth day of immobilization, 13% of bedridden, non-surgical patients developed DVT as diagnosed by daily FLS. Over half of these developed by the fifth hospital day. Hence, even short periods of immobilization confer an increased risk for DVT (31).

Pregnancy and postpartum

Pregnancy and the postpartum state are considered to be high risk periods for VTE. Interpretation of existing data in this area is made difficult by the small number of patients studied, an overeliance on clinical diagnosis due to the adverse effects of radiation on the developing fetus, and varying definitions of the peripartum period. One large retrospective study using limited contrast venography in pregnant women found 11 documented cases of DVT among 14,869 women, 9 of which occurred postpartum, which is a pre- and post-partum rate of 10 and 61 per 100,000 respectively (32). In a prospective study during pregnancy using objective diagnostic criteria, the occurrence of 60 episodes of DVT were equally distributed during the three trimesters (33).

Previous venous thrombosis

An objectively confirmed previous venous thromboembolic episode is associated with an increased risk of DVT, especially in high-risk settings such as surgery, where studies using FLS have shown a two to three-fold increased risk of DVT (34). This risk likely results from permanent damage to the veins as well as the persistence of individual risk factors that promoted the development of the first episode of DVT. Of interest, most DVT prevention and treatment studies have excluded such patients, presumably because they represent a group at different risk than those without prior DVT, and because diagnostic tests do not perform as well due to altered venous anatomy and function.

Oral contraceptive use

Although burdened by various methodological flaws, mostly related to reliance on clinical signs to diagnose DVT, the weight of the evidence points to a 2 to 8 fold increased risk for DVT in women using oral contraceptives (35,36). A recent matched case-control study of 471 women aged 16-44 with VTE and 1772 controls found an odds ratio for VTE of 4.0 with use of oral contraceptives vs. non-use, and a four-fold probability of death due to VTE in users compared to non-users (37).

Blood abnormalities

Congenital deficiencies of protein C, protein S, and antithrombin III have been described frequently in association with recurrent DVT and DVT occurring at a young age or in unusual locations. However, the risk of DVT in individuals with these deficiencies has yet to be clarified. Overall, since these deficiencies are rare, DVT in the general population is rarely associated with these disorders (38). Activated protein C resistance, a recently described mutation that alters the binding site of factor V for activated protein C, has been reported to occur in 5% of the general population and in 20-40% of unselected patients with DVT, which would make it the most common inherited cause of DVT (39).

Other blood abnormalities that may confer an increased risk of DVT include the lupus anticoagulant, dysplasminogenemias and dysfibrinogenemias (38).

Other risks

There have been a number of other links made in the literature between certain clinical factors and the risk for DVT, for which sound data on causal association are not available. These include obesity, varicose veins, congestive heart failure, infection, inflammatory bowel disease, nephrotic syndrome, polycythemia, paroxysmal nocturnal hemoglobinemia and Behçet's disease.

2.2.4. Practical applications of risk data

Using data available on risk factors for perioperative VTE, a widely used risk classification system has been developed which takes into account both baseline individual risk factors and the type of surgery (20,40). As depicted in Table 2.1 below, this system outlines the rates of VTE associated with the various risk strata. The risk groups are defined based on the incidence of DVT detected by surveillance tests and the potential benefits of prophylaxis as demonstrated in clinical trials.

Table 2.1. Incidence of VTE by risk group in surgery patients

Risk Category*	CALF DVT %	PROXIMAL DVT %	CLINICAL PE %	FATAL PE %
Low risk	2	0.4	0.2	0.002
Moderate risk	10-20	2-4	1-2	0.1-0.4
High risk	20-40	4-8	2-4	0.4-1.0
Very high risk	40-80	10-20	4-10	1-5

* Low risk: Uncomplicated minor surgery in patients younger than 40 with no clinical risk factors (i.e. prolonged immobility, paralysis, prior DVT, varicose veins, congestive heart failure, myocardial infarction, stroke, malignancy, laboratory markers of increased DVT risk). Moderate risk: Major surgery in patients older than 40 with no other clinical risk factors. High risk: Major surgery in patients older than 40 who have additional risk factors. Very high risk: Major surgery in patients older than 40 with previous DVT or cancer or orthopedic surgery or hip fracture or stroke or spinal cord injury.

This risk index applies only to surgery patients. Other DVT risk indices have been developed for use in either high-risk asymptomatic patients or symptomatic patients with

suspected DVT. These will be critically reviewed later in this chapter.

2.3. Diagnosis of DVT

Until the 1960's, because of the unavailability of safe and reliable diagnostic tests, DVT was diagnosed clinically, with poor accuracy. With the advent of contrast venography, Haeger showed in 1965 that the venous system was completely normal in 46% of patients receiving treatment for DVT (12). Conversely, autopsy studies have demonstrated consistent underdiagnosis of DVT. In one series, among 195 patients who died of autopsy-proven pulmonary embolism, 162 (83%) had coexisting DVT. However, in only one fifth of these was DVT suspected antemortem, and only 3% had an objective test to confirm the diagnosis (41). Hence, clinical over- and underdiagnosis of DVT are both recognized problems, leading to a general consensus in the medical literature that the clinical diagnosis of DVT is inaccurate and cost-ineffective.

2.3.1. Tests used to diagnose DVT

Despite the publication in recent years of more than 50 studies on the reliability and validity of various diagnostic tests for DVT, the choice of the best test remains a controversial area. The nature of this controversy stems from two main problems, one general, and one particular to the study of DVTs. Firstly, the tests perform differently in different populations, hence the results of individual studies may be neither reproducible nor generalizable. Secondly, DVT that is diagnosed in surveillance studies of asymptomatic high risk patients, although providing us with interesting epidemiological information on incidence and risk, might not have the same natural history as DVT diagnosed in symptomatic patients.

As discussed in the Introduction, correctly identifying those who do not have DVT is as important as diagnosing DVT in those with the disease. In considering the strengths and weaknesses of individual diagnostic tests, the following points are important:

- The primary objective of diagnostic testing and subsequent treatment of DVT is avoidance of symptomatic PE, which can be fatal. Symptomatic PE results primarily from proximal (rather than calf) DVT (42).

- In most studies, the prevalence of DVT in patients with leg symptoms suspicious for DVT is 33-50%.
- 60-80% of patients with leg symptoms in whom the diagnosis of DVT is confirmed have proximal DVT (43).
- 90% of proximal thrombi originate in the calf (43).
- 20-30% of untreated calf thrombi propagate proximally (42).
- The likely explanation for the presence or absence of symptoms in an individual with confirmed DVT is that in symptomatic patients, occlusive clot obstructing the vein lumen is more likely (84% occluding; 16% non-occluding), whereas asymptomatic patients are more likely to have non-occlusive clot which adheres to the vein wall without disrupting venous flow (77% non-occluding; 23% occluding). Both can lead to PE (44).

There are three general patient categories that are important to distinguish in evaluating the literature on diagnostic testing for DVT: symptomatic patients presenting with a first episode of possible DVT, symptomatic patients with possible recurrent thrombosis, and asymptomatic patients at high risk for DVT. Thromboprophylaxis studies which use various surveillance tests for DVT typically enroll asymptomatic, high-risk patients (e.g. orthopedic surgery patients), whereas studies focusing on performance characteristics of diagnostic tests tend to enroll symptomatic patients who are referred for suspicion of DVT. Because of the difficulty in diagnosing recurrent DVT, patients with suspected recurrent thrombosis are often excluded, as in this study, and will not be addressed further here.

Even within these general categories, the performance of the test may differ in different patient groups, e.g. symptomatic inpatients vs. outpatients (45,46). Hence, both the choice of an optimal test and the interpretation of its results depend heavily on the target population of interest.

Contrast venography (CV)

CV is considered by most to be the 'gold standard' or reference test for the diagnosis of DVT. This technique involves injection of iodinated contrast material into a

small peripheral foot vein of the affected limb, followed by radiologic visualization of the venous system. Detailed visualization of the entire venous system is possible, and as such it is the only test that can identify all thrombi, whatever the size or location. Acute DVT can be distinguished from chronic venous disease.

The most reliable criterion for a positive test is the presence of a constant intraluminal filling defect visualized in at least two different projections (47). It is safe to withhold anticoagulation in those with negative venograms, as shown in a study by Hull in 1981. Among 160 consecutive patients with negative venograms, only 1.3% were subsequently proven to have thrombi which were likely induced by the procedure itself (48).

However, CV has several limitations: it is invasive, can cause pain and allergic reactions to the contrast dye, and carries a 2-4% risk of inducing DVT (49). Up to 20% of patients cannot have venography done for technical or other reasons (45). Also, because CV yields findings that are technically unsatisfactory for diagnosis in 10-15% of studies (50), and observer disagreement occurs in approximately 10% of cases (47), even venography cannot be assumed to have a sensitivity or specificity of 100%. It is not practical as a screening test in patients at high risk for DVT since it is not readily repeatable. For all of these reasons, it is not an ideal procedure, particularly when the pre-test probability of DVT is low. Recognition of these limitations has led to the development of several non-invasive tests over the last two decades. In studies examining the validity of non-invasive tests for DVT diagnosis, CV is generally used as the reference test, since although it is an imperfect "gold standard", it remains the best test available for this purpose.

Impedance plethysmography (IPG)

IPG is a low cost, non-invasive technique which measures the electrical impedance across the leg while venous flow is obstructed and then released using a cuff around the thigh. In effect, this measures the flow variability between patent and obstructed veins. IPG readily detects major venous occlusions of the proximal veins. IPG cannot be performed in the presence of a cast or traction (51).

False positive IPG can result from pregnancy, congestive heart failure, or from problems positioning the leg. Its major limitations are poor sensitivity (22%) for the detection of calf thrombi, probably since calf thrombi have little impact on the collective rate of venous outflow, and poor sensitivity in the setting of asymptomatic, high risk patients, such as those undergoing hip surgery (52), probably because in these patients, thrombi, although originating in proximal veins, are non-obstructing and thus do not interfere with venous hemodynamics.

In symptomatic patients, for thrombi in or proximal to the popliteal vein, the mean estimated specificity and sensitivity of IPG are 95% and 92% respectively, with positive predictive values of 83-95% and negative predictive values of 90-96% (51). However, some recent reports reevaluating IPG against venography have demonstrated a sensitivity of only 66% (53,54). This could reflect changes in referral practices, such that patients with smaller, less occlusive thrombi are being referred for testing, or could be due to improved venographic identification of smaller thrombi that were previously overlooked (55). Until further study, this issue remains unresolved.

Because of the poor sensitivity of IPG for calf DVT, serial testing on days 2 and 7 is required to detect missed calf DVT that might have propagated proximally. Data has shown that 14% of cases of proximal DVT are detected by serial testing performed after the initial IPG (46). Patients with repeatedly negative serial IPGs have a 2.5% chance of developing an episode of DVT during the subsequent 6 months (46).

Compression ultrasonography (CUS)

CUS is a non-invasive technique which simultaneously evaluates blood flow and images the veins. The presence of thrombus in the vein prevents venous collapse when pressure is applied over the vein with a Doppler probe. This lack of venous compressibility is the most reliable indicator of acute DVT using CUS (56). Visualization of thrombus in the vessel is an additional diagnostic criterion.

Sensitivity and specificity of CUS for the diagnosis of proximal DVT are both estimated to be 97% (56). CUS is insensitive to calf thrombi because of technical difficulties in visualizing the calf veins. Because of this, as with IPG, serial testing is

recommended in patients with an initial negative test. In a study by Heijboer of 491 symptomatic patients who had CUS, 6% of cases were detected only during serial testing. Patients who are obese or who have marked edema may have false negative results. The incidence of DVT during a six month follow-up in those with negative tests was 1.5% (46).

Because of its poor sensitivity for calf DVT and non-occluding proximal DVT, CUS, like IPG, is not adequately sensitive for use as a surveillance test in asymptomatic high risk patients.

^{125}I Fibrinogen leg scanning (FLS)

FLS requires the injection of radioactive fibrinogen through a peripheral vein, which becomes incorporated into thrombi that are actively forming. The radioactivity is subsequently detected from the surface of the limb using a scanning device. DVT is diagnosed if the scan is persistently abnormal after 24 hours. Its main use has been for DVT surveillance of asymptomatic, high risk patients. False negative results occur with old, inactive thrombus, small thrombus, and thrombus in the common femoral or iliac vein. False positive results occur in the presence of hematoma, inflammation, and urinary incontinence (51).

In a prospective analysis of 120 asymptomatic patients undergoing FLS for surveillance of DVT, the test had an overall sensitivity of 72%. Sensitivity for proximal DVT was 63% compared to 81% for calf DVT (44). FLS is 95% sensitive to symptomatic calf DVT of recent onset (57), however its sensitivity for symptomatic proximal thrombosis is only 60-80%, which is inadequate for use as a diagnostic test.

A study which examined the accuracy of the combination of IPG and FLS as an alternative to venography in symptomatic patients found that the combined sensitivity for all (proximal and calf) DVT was 94%, the specificity was 91%, the positive predictive value was 89% and the negative predictive value was 95% (57). If IPG was positive, FLS did not add any useful information to the management or outcome. Both tests performed better in patients whose symptoms were present for less than one week.

Pitfalls of FLS include the potential hazards of exposure to radiation, and the 24 hour delay in availability of test results. Because of the limitations discussed, it has fallen out of favor as a test for DVT.

Other tests

Other tests that have been used to diagnose DVT include measurement of hematological markers of ongoing thrombosis (e.g. d-dimer, fibrin degradation products) and the technetium RBC scan. They are not recommended because of poor test performance compared to the methods discussed above.

Table 2a (end of chapter) summarizes the properties and indications for use of the diagnostic tests discussed.

2.3.2. Clinical symptoms and signs of DVT

Despite the availability of various tests for DVT, strategies for referral of patients for diagnostic testing must of necessity be based on clinical suspicion, namely the use of clinical findings to estimate the likelihood of DVT.

The typical symptoms of DVT reported by patients are pain, warmth, redness and swelling of the lower extremity. The mechanism for pain is thought to be vein wall inflammation and venous distention. Warmth and redness are due to vein wall inflammation and shunting of blood from the obstructed deep vein to the superficial veins. Swelling is due to venous outflow obstruction. These symptoms may be present in various combinations and typically evolve over a few days, but more rapid (over hours) and more chronic (over weeks) evolution can both occur. Symptoms may be absent, as shown by autopsy studies and surveillance studies in high-risk surgical patients.

Signs of DVT as noted by physical examination include tenderness, warmth, erythema, cyanosis, edema (usually pitting), palpable cord (a palpable thrombosed vein), superficial venous dilation (SVD), and a number of elicitable signs named for the physicians who first described them: Homan's sign, the best-known of these, is present if sudden dorsiflexion of the ankle joint with the knee flexed to 30° produces discomfort in

the upper calf. Louvel's sign denotes worsening of pain along the course of a thrombosed vein by coughing or sneezing. The Lowenberg sign is present if, after inflation of a sphygmomanometer cuff around each calf, pain is experienced in the affected calf at a lower pressure than the unaffected one.

The differential diagnosis of a swollen, painful lower extremity is extensive, and includes cellulitis, arthritis, neuropathy, arterial occlusion, lymphedema, varicose veins, superficial thrombophlebitis, and chronic venous insufficiency. Some of these entities can be easily diagnosed at the time of initial presentation, avoiding further testing, whereas others can be inferred only after DVT has been objectively excluded. In a follow-up study of 87 consecutive patients who were clinically suspected to have DVT but who had negative venograms, the final etiologies of the symptoms were: muscle-related in 40%, cellulitis in 3%, leg swelling in a paralyzed leg in 9%, venous reflux in 8%, lymphatic in 8%, Baker's cyst in 5%, and unknown in 26% (48).

A number of studies have examined the accuracy of symptoms and signs in diagnosing DVT. Various methodological flaws affect the interpretation of these studies, such as inadequate description of selection criteria (58), highly selected populations with poor generalizability (58-60), lack of information regarding blinding of the clinician to the patient's ultimate diagnosis (12,58-63), use of retrospective clinical data gathered from charts after DVT was diagnosed (61), failure to provide criteria for a positive diagnostic test (58-63), and inadequate description of guidelines used to determine the presence or absence of clinical signs, which in themselves are often poorly defined and have untested reliability (58,61,62,64).

Nevertheless, it is possible from the methodologically more robust studies to estimate the sensitivities and specificities of various clinical symptoms and signs for the diagnosis of DVT. The specificities of many of these are low by design, since the clinical suspicion of DVT is what allows the patient entry into the study.

Haeger (12) prospectively studied 72 consecutive patients presenting to his vascular clinic with suspected DVT. All patients were examined by one or two experienced vascular surgeons, who documented the presence or absence of selected symptoms and signs before the patient underwent CV. Patients with four or more positive

signs were classified as "highly suspected". Overall, 46% of the sample had DVT proven by CV. Among the individual signs, calf pain and tenderness had the highest sensitivities (0.90 and 0.84 respectively), whereas superficial venous dilation (SVD) and Lowenberg's sign had the highest specificities (0.82 and 0.85 respectively). No individual sign had both high sensitivity and specificity, and even in the "highly suspected" group, positive predictive value (PPV) and negative predictive value (NPV) for DVT were .55 and .66 respectively. Haeger concluded that "clinical signs cannot be trusted" to diagnose DVT.

A later, well designed study by Richards et al. (65) prospectively compared the diagnostic accuracy of four non-invasive techniques, including physical examination, for the diagnosis of DVT. The study population consisted of 85 patients referred over a 12 month period for clinically suspected DVT or PE. There were no apparent exclusion criteria. Clear-cut criteria for a positive CV, the reference test used, were provided. All physicians performing non-invasive examinations and those performing and interpreting CV were blinded to each other's findings. The positive end-points of the 11 different physical exam maneuvers were explicit. In most patients bilateral CV was performed.

After exclusion of technically inadequate venograms, 60 DVT were found among 150 extremities, a prevalence of 40%. Of these, 37 (62%) were proximal, and 23 (38%) were calf DVTs. Excluding 3 of 11 physical exam maneuvers that are not well known and rarely used (Moses', Ramirez' and Peabody's signs), leg tenderness had the highest sensitivity (0.62). Difference in calf circumference, Homan's sign, Lowenberg's sign, warmth, SVD and palpable cord all had specificities of > 0.80, with palpable cord being both the most highly specific (0.98) and poorly sensitive (0.10) sign. However, due to the 40% DVT prevalence, the PPV of these signs was poor. Overall, all signs, whether taken individually or in combination, had poor predictive accuracy.

Sandler et al., in a similar study of the diagnostic accuracy of various non-invasive techniques for DVT, studied 50 patients with suspected DVT referred by various physicians, mostly general practitioners (66). A standardized clinical examination was carried out in each patient, focusing on the presence or absence of calf pain, pitting edema 3 cm. above the medial malleoli, and a palpable difference in temperature between the two legs. The percentage differences in circumference at the calf and thigh between the

affected and unaffected leg were noted. The reference test was CV, with well-described criteria for a positive study and independent review by two radiologists. It was not stated whether those performing the clinical exam were blinded to the diagnosis. In addition to sensitivity and specificity, the kappa index for each was reported in order to correct for chance agreement.

DVT was diagnosed by CV in 29 (58%) of patients; 24 (83%) had proximal DVT, and 5 (17%) had calf DVT. This prevalence is higher than that noted in most studies of patients referred for suspected DVT, and could indicate that the population examined was at high risk for DVT or had clinical features which were particularly typical for DVT. None of pain, edema, or temperature difference had good specificity, but they appeared to be reasonably sensitive (0.86, 0.97 and 0.72 respectively). However, after taking into account chance agreement, only the sensitivity for edema remained robust (0.78). There was a large overlap of differences in leg circumference between those with and those without thrombosis. If clinical signs alone had been used to make the diagnosis of DVT, 42% of the patients would have received anticoagulation unnecessarily, since the predictive accuracy of the clinical features, apart from ankle edema, was little greater than might have occurred by chance.

The above studies are summarized in Table 2b (end of chapter). In brief, symptoms and signs in themselves do not appear to be useful in discriminating between patients with and without DVT. The overall poor specificities and PPVs of the various symptoms and signs are not surprising, given that patients are referred for testing because of these features and that the prevalence of diagnosed DVT in symptomatic patients is typically only 40%. The poor sensitivities of individual factors could indicate that combinations of these factors would be more helpful in predicting DVT. In any case, focusing solely on symptoms and signs is artificial, since clinicians typically have other data at hand which aids clinical judgment when assessing the individual patient, such as demographic factors, concurrent disease status, past medical history and medication use. I will now focus on studies that have combined clinical factors and other relevant data for the development of prediction indices in an attempt to more accurately predict DVT.

2.3.3. Clinical prediction indices

Methodological considerations

Clinicians traditionally make diagnostic predictions informally and nonquantitatively, using some combination of clinical experience and published evidence (67). Several cognitive principles, or heuristics, for using personal experience to estimate probability have been described (68). A clinician is using the representativeness heuristic when he judges the probability that a patient has a disease by the extent to which the clinical findings resemble the essential features of the disease. This can be misleading if the disease is rare, if the findings are poor predictors of the disease or if the clinician's experience of the disease is based on a small, atypical sample. The availability heuristic is being used when the probability of a diagnosis is judged by the ease with which similar diagnoses are remembered. This is misleading when the clinician has recently diagnosed a rare condition. Finally, a clinician is using the adjustment heuristic when initial probability estimates are adjusted to take into account unusual patient features. Studies have shown that clinicians do not adjust their initial estimates enough and in general overestimate the probability of disease (68,69).

Clinical prediction rules or indices are statistical models based on information procured from numerous patients which quantitatively estimate the probability of a diagnostic outcome. Methodological standards have been described for the development and validation of clinical prediction rules (67,70). The definition of the event to be predicted should be clear and free of ascertainment bias. The predictive findings should be precisely defined, easily available to the clinician, and ideally have proven reliability. Assessment of outcome and predictive findings should be blinded, and both should be clinically relevant. The patient selection process should be described. The population should include a wide spectrum of patients and should be representative of the clinical practice in which the prediction rule is to be used. The margin of error in the point estimate of probability and the misclassification rate should be provided as a measure of the accuracy of the prediction rule. Cross-validation techniques, or ideally, testing of the prediction rule prospectively in a new clinical setting should be done. The mathematical

techniques for developing the prediction rule should be identified. Finally, the ultimate measure of a clinical prediction rule is its effect on patient care, such that even when the above methodological standards have been met, the prediction rule may have little clinical utility.

Published clinical prediction indices for DVT

These include indices developed for use in asymptomatic high risk surgical patients and indices for symptomatic patients presenting because of suspected DVT.

-Gynecological surgery patients

Crandon, in 1980, developed a predictive index for patients who were undergoing gynecological surgery (71). The index was derived in 124 patients and validated in a further 62 patients. Data was obtained on height, weight, age, hospital length of stay before surgery, smoking habits, pre-operative hemoglobin level, varicose veins, history of DVT, nature of surgery, presence of malignancy, and a number of laboratory measures reflecting coagulation status. All data were collected prospectively, and stepwise logistic discriminant analysis was performed. Pre-operative and serial post-operative FLS was performed to assure the absence of DVT pre-operatively and to capture post-operative DVTs.

The best single predictor of post-operative DVT was euglobulin lysis time (ELT), a laboratory measure of fibrinolysis, followed by age, varicose veins, fibrin-related antigen (FRA), and percentage overweight. The equation for the index was I = -11.3 + 0.009(ELT) + 0.22(AGE) + 0.085(VARICOSE VEINS) + 0.043(FRA) + 2.19(%OVERWEIGHT). The index was applied prospectively to 62 patients undergoing similar gynecological surgery, 16% of whom developed DVT, and was demonstrated to have a sensitivity of .90, specificity of .87, PPV of .56 and NPV of .98 when I = -2 was taken as the cut-off point for the equation.

Problems with this index include the need for blood tests not widely available, applicability to a restricted surgical population only, and validation on a small number of patients with post-hoc determination of the best cut-off point for the index's equation.

Patients with negative FLS did not undergo CV. This exposes the sample to misclassification bias given the known poor sensitivity of FLS, especially for proximal DVT. With regard to clinical applicability, this index is not applicable to symptomatic patients, and would be difficult to use even in its intended population.

Using logistic regression analysis, Clarke-Pearson attempted to develop a more clinically useful index to predict DVT post-operatively in women undergoing major abdominal or pelvic surgery for gynecological disease (25). The variables in the final model were anesthesia duration greater than 300 minutes, age, prior DVT, race, edema, and severity of varicose veins. The authors state that the degree of concordance (0.82) demonstrates the effectiveness of the model, but it is not clear what this means or how it was derived. Only patients with FLS that were positive in popliteal or more proximal regions underwent CV, hence serious misclassification in the direction of underdiagnosis was likely. The authors do not state how patients were recruited for this study, or whether those interpreting the leg scans and CV were blinded to the patients' clinical status. Also, the model has not been prospectively validated on other populations of gynecological surgery patients. To the authors' credit, this study did prospectively identify risk factors for the development of DVT in this patient population, which is of epidemiological importance.

-Abdominal surgery patients

Lowe developed an index for use in patients undergoing elective abdominal surgery (72). The index was derived in 63 such patients, via prospective collection of clinical and lab data followed by FLS performed pre-operatively and then serially post-operatively. One-third developed DVT.

Using linear discriminant analysis, the best model was AGE + 1.3(% MEAN POPULATION WEIGHT FOR HEIGHT, AGE AND SEX). In the derivation group, a cut-off of 175 had good sensitivity (.90) but poor specificity (.52). The validation group consisted of 41 similar patients, half of whom developed DVT. Using a cut-off point of 170, the sensitivity and specificity of the index was unchanged. The model includes a variable for obesity, which is not a proven independent risk factor for DVT. In Lowe's

study population, obesity could have been spuriously associated with DVT or with another, unmeasured risk factor for DVT. This could explain the poor specificity of the index.

A small effectiveness study of Lowe's index was also performed, in which only those patients with a score greater than 170 were given anticoagulant prophylaxis pre-operatively. The incidence of DVT in this patient subset was reduced from 51% (derived from the derivation and validation studies) to 8%. However, DVT also developed in 12.5% of patients with a score less than 170 ("low-risk" patients), which could be argued is unacceptably high.

The problems with this index are the small sample size available for the derivation and validation studies, and poor specificity, such that the PPV achieved with use of the index is barely an improvement over the pre-test probability, as seen by the likelihood ratio-positive (LR+) of 1.8. The index was subsequently applied by another author to 47 similar patients and was completely unable to predict DVT (73).

Sue-ling developed a predictive index for DVT in similar patients using similar methods (74). The incidence of DVT in both the derivation (n=85) and validation (n=43) groups was one-third. Numerous clinical and laboratory measures were gathered at baseline. DVT was diagnosed prospectively by FLS, with confirmation by CV if FLS was positive. Using stepwise logistic discriminant analysis, the most powerful predictors of DVT in this population were age (like Lowe and Crandon) and ELT (like Crandon). The model, $I = -11.5 + 0.133(AGE) + 0.006(ELT)$, had a sensitivity of .93 and a specificity of .83 in the validation group when a cut-off of -1.5 was chosen. With a DVT prevalence of 33% in this group, the PPV was .72 and NPV was .96. Hence, this index was more specific than that of Lowe, and had better predictive accuracy.

The advantage of Sue-ling's index is its simplicity. However, ELT is not a readily available test in most hospitals, and the index has yet to be validated in other populations of pre-operative patients. Also, the index seems too simple, probably because due to small patient numbers, there was inadequate power to detect significant differences in other important variables between groups. For example, the prevalence of varicose veins, percent lower (vs. upper) abdominal surgery and the duration of pre-op stay were all

higher in the DVT group, yet none of these differences achieved statistical significance.

-Symptomatic patients with suspected DVT

Vine in 1981 (75) retrospectively studied 150 consecutive patients who had contrast venography, one third of whom had DVT. Various elements of the clinical history, exam, and laboratory results were collected via retrospective chart review, and likelihood ratios for each variable were calculated. Receiver Operating Characteristic (ROC) curves were constructed using disjunctive analysis to assess the additive contribution to the risk for DVT of the baseline variables with the highest likelihood ratios (malignancy, recent blood transfusion, recent surgery, congestive heart failure, immobilization, infection, erythema of legs, anemia, and leg swelling).

Weaknesses of this study include lack of information on patient selection, retrospective data collection with no information on amount of missing data, applicability of the ROC curves to this patient population only (ie. not validated in other populations), difficulty in deciding which cut-off point was most meaningful, and inability of the average clinician to use these curves in daily practice.

Landefeld used retrospective methods similar to Vine's to identify clinical findings useful in estimating the probability of acute proximal deep vein thrombosis (76). The population studied was 355 consecutive patients who had contrast venography over a two year period and for whom medical records were available. The authors state that most DVT patients in their hospital were captured, yet they also note that among 20% of patients with a discharge diagnosis of DVT, venography was never performed, which raises questions about the hospital's selection process for venography. Data on 76 clinical items, including symptoms, signs, comorbid conditions and laboratory tests were gathered retrospectively from chart review. In order to avoid ascertainment bias, i.e. the chance that knowledge of venographic results would affect the observation or recording of data, an attempt was made to only gather data recorded before venography was performed.

Venograms were normal in 185 patients (52%), showed proximal DVT in 96 patients (27%), and were equivocal in seventy-four patients (21%), i.e. were either non-diagnostic due to inadequate filling of the deep veins (45 patients) or showed only calf

DVT (29 patients).

A derivation group and a validation group were randomly selected from within the study population. Linear discriminant analysis was used to identify independent predictors of the venographic diagnosis. The clinical findings associated with proximal DVT in the univariate analysis were male sex, age 65 years or older, active cancer, fever, recent immobilization, shorter duration of leg symptoms, swelling above the knee, and swelling below the knee. Only swelling above and below the knee, recent immobility, cancer and fever were independent predictors of proximal DVT. Patients in the validation group were classified according to the number of independent predictors present. The risk for proximal DVT was 5% among patients with none of the five predictors (LR+ 0.15), 15% among patients with one predictor (LR+ 0.47), 50% in those with 2 predictors and 30% in patients with 3 or more predictors (LR+ 2.0 for 2 or more predictors). In patients with at least one of the five predictors, the sensitivity of the index was .97 and the specificity was .26. In patients with more than 3 predictors, sensitivity fell to .20 but specificity rose to .85.

Although this index would be very easy to apply in clinical practice, there were methodological flaws in its development and validation. Symptoms and signs were recorded as "present" or "not known to be present", so that an absent finding was treated in the same manner as one not recorded. This could have underestimated the diagnostic value of certain findings. Data was collected retrospectively, and despite the authors' assurances, ascertainment bias very likely occurred. Variables were considered for multivariate analysis based primarily on a p-value cut-off point, without consideration of clinical relevance for those variables not achieving this cut-off point. DVT was more likely in those with 2 predictors than in those with 3 or more predictors, which is not logical, and could be due to peculiarities in the data resulting from small patient numbers or to a general lack of validity of this index. Also, in patients with 2 or more predictors, the probability of DVT (PPV) was 42%, indicating that this index adds little to predictions based simply on the prior probability of DVT in symptomatic patients, i.e. prevalence of 30-50%. Finally, the index has yet to be validated in other patient populations who may be at different risk for DVT.

-Wells' prediction index

The final, and most clinically useful index for DVT prediction was that developed by Wells and colleagues in 1995 (77). Prior to the study the authors developed a clinical model, based on literature review and clinical expertise, which stratified patients into three pre-test probability categories for DVT: high, moderate or low. The study was conducted at three centers in Canada and Italy. Those eligible for participation were outpatients with clinically suspected DVT who had symptoms for less than 60 days. Patients were excluded if they had prior VTE, could not tolerate contrast dye, had suspected PE, were pregnant, or were on anticoagulants. Also excluded were patients with an obvious alternative cause for their symptoms who did not go on to have diagnostic testing. DVT was diagnosed by CV, which was interpreted by readers blinded to the patient's clinical history. Of 887 consecutive patients, 358 were ineligible, mostly because of prior DVT, alternative diagnosis for their symptoms, or inability to perform or evaluate CV. Table 2.2 shows the model used:

Table 2.2. Wells' *a priori* prediction index

Major Points	Minor Points
Active cancer	Recent trauma
Paralysis, paresis, or recent cast	Pitting edema in symptomatic leg
Recent immobilization or surgery	Dilated superficial veins in symptomatic leg
Tenderness along deep vein distribution	Hospitalization in last 6 months
Swollen thigh and calf (measured)	Erythema
Family history of DVT	

Wells' index: clinical probability rating

High: ≥ 3 major points, or ≥ 2 major and ≥ 2 minor points, and no alternative diagnosis.

Low: 1 major point and at least 2 minor points and an alternative diagnosis, or 1 major point and at least one minor point and no alternative diagnosis, or at least 3 minor but no major points and an alternative diagnosis, or at least 2 minor points but no major points and no alternative diagnosis.

Moderate: all other combinations.

Among the 529 study patients, 135 (25.5%) had DVT. The patients were assigned to a probability group before undergoing CV. The values for sensitivity, specificity, PPV

and NPV of the three pre-test probability groups are shown in Table 2.3:

Table 2.3. Accuracy of Wells' index for the prediction of DVT

Pretest probability	# Patients	DVT n(%)	Sensitivity	Specificity	PPV	NPV
High	85	72 (85%)	.53*	.97*	.85	.86*
Moderate	143	47 (33%)	.75#	.75#	.33	.95#
Low	301	16 (5%)			.05	

* compared to low and moderate risk groups combined

compared to low risk group

The accuracy of the clinical model was similar in all three hospital centers despite differences in DVT prevalence in the centers. The model was found to have excellent interobserver reliability (kappa=.85). It is relatively easy to apply, uses readily available data, and could be combined with non-invasive testing in order to improve the efficiency of the diagnostic process in patients with DVT, especially in cases where pretest probabilities and non-invasive test results are discordant. Limitations of this index are that other clinical risk factors not considered for inclusion in the model could be as or more useful in predicting DVT, and, since multivariate analysis was not performed, it is not known whether or not some of the clinical predictors are collinear, interact with one another, or are acting as confounders in the association between unmeasured predictors and DVT. Its ultimate utility can only be assessed after prospective validation of the model on other populations.

Table 2c (end of chapter) summarizes the test characteristics of the prediction indices discussed above. Table 2d (end of chapter) shows the extent to which they adhered to suggested methodological standards for clinical prediction rules (67,70). No one index met all methodological standards.

The next two chapters describe the methods and results of the study I performed to develop a prediction index for DVT in symptomatic patients with suspected DVT.

Table 2a. Diagnostic tests for DVT: summary of properties and indications for use

TEST	Advantages	Disadvantages	Indications for use
CV	• "gold standard" for use in patients with or without symptoms • anatomic visualisation of venous anatomy of whole leg	• invasive • side effects • potential for serious allergy • can induce DVT • inconvenient for repeat testing in screening of high risk patients	Diagnosis of DVT (calf and proximal) in symptomatic patients Confirmation of an abnormal non-invasive test result Screening studies of high risk patients
IPG	• non-invasive • simple, easily used • detects virtually all occlusive thrombi in thigh	• insensitive to calf thrombi • may miss non-occlusive thigh thrombi • serial testing required if initial test negative • not for use in asymptomatic patients	Diagnosis of proximal DVT in symptomatic patients
CUS	• non-invasive • may perform slightly better than IPG	• similar to IPG	Diagnosis of proximal DVT in symptomatic patients
FLS	• less invasive than venography • good sensitivity for symptomatic calf DVT	• exposure to radiation • poor sensitivity to proximal DVT, or DVT > 7 days old • poor specificity overall • test result available after 24 hours • may take 72 hours for test to be positive	Diagnosis of DVT in combination with other tests Screening selected high risk patients when confirmed by CV

CV=contrast venography IPG= impedance plethysmography CUS= compression ultrasonography FLS=^{125}I fibrinogen leg scanning

Table 2b. Studies on the diagnostic accuracy of clinical symptoms and signs of DVT

Study	Patient population	Reference test for DVT	Diagnostic criteria for + test stated?	DVT Prevalence and site n (%)	Clinical symptoms /signs studied	Sens	Spec	PPV	NPV
Haeger 1969 (12)	clinically suspected DVT n=72; 50% male; age: range 16-68	CV	yes	33 (46%) site NR	calf pain tenderness cool skin ankle edema calf edema SVD Homan Lowenberg highly suspected (≥4 signs)	.90 .84 .42 .76 .42 .33 .33 .20 .67	.03 .26 .62 .24 .68 .82 .79 .85 .54	.44 .49 .50 .46 .52 .60 .58 .60 .55	.25 .67 .56 .56 .56 .61 .57 .52 .66
Johnson 1974 (58)	n=30; selection process and patient characteristics not defined	CV	no	16(53%) 10(63%) proximal 6(37%) calf	whole leg edema calf edema tenderness Homan	.06 .56 .56 .13	.88 .69 .44 .81	.33 .64 .50 .40	.48 .61 .50 .48
Cranley 1976 (59)	clinically suspected DVT; highly selected n=133 limbs (124 patients)	phleborhe-ography and CV	no	72 (54%) site NR	muscle pain tenderness swelling Homan	.83 .82 .90 .48	.15 .28 .08 .59	.54 .57 .54 .62	.43 .57 .42 .46

CV=contrast venography NR=not reported SVD= superficial venous dilation Homan=Homan's sign Lowenberg=Lowenberg's sign diff.= difference inpts= inpatients outpts= outpatients adeq= adequate Sens= sensitivity Spec= specificity PPV= positive predictive value NPV= negative predictive value

Table 2b. (con'd). Studies on the diagnostic accuracy of clinical symptoms and signs of DVT

Study	Patient population	Reference test for DVT	Diagnostic criteria for + test stated?	DVT Prevalence and site n (%)	Clinical symptoms /signs studied	Sens	Spec	PPV	NPV
Richards 1976 (65)	150 limbs in 85 patients with suspected DVT	CV	yes	60 (40%); 37 (62%) proximal 23 (38%) calf	ankle edema	.40	.52	.36	.57
					diff. in calf size	.35	.89	.68	.67
					tenderness	.62	.71	.59	.74
					Homan	.42	.84	.64	.68
					Lowenberg	.37	.81	.56	.66
					warmth	.33	.87	.63	.66
					SVD	.27	.91	.67	.65
					palpable cord	.10	.98	.75	.62
Lindqvist 1977 (63)	47 patients with suspected DVT	CV	no	24 (51%) 15 (63%) proximal 9 (37%) calf	swelling	.79	.30	.54	.58
					pain	.75	.13	.47	.33
					stiffness	.75	.52	.62	.67
					edema	.63	.35	.50	.47
					tenderness	.58	.26	.45	.38
					SVD	.33	.70	.53	.50
					Homan	.29	.61	.44	.55
					Lowenberg	.25	.70	.46	.47
					cyanosis	.25	.57	.38	.72
Cooperman 1979 (62)	98 limbs in 67 patients with suspected DVT; 40% male; age: mean 54; range 18-75	CV	no	23 (23%)	tenderness	.74	.56	.34	.88
					swelling	.74	.57	.35	.88
					Homan	.43	.81	.42	.82
					heat	.35	.92	.57	.82
					redness	.26	.91	.46	.80

CV=contrast venography NR=not reported SVD= superficial venous dilation Homan=Homan's sign Lowenberg= Lowenberg's sign
diff.= difference inpts= inpatients outpts= outpatients adeq= adequate Sens= sensitivity Spec= specificity PPV= positive predictive value
NPV= negative predictive value

Table 2b. (con'd). Studies on the diagnostic accuracy of clinical symptoms and signs of DVT

Study	Patient population	Reference test for DVT	Diagnostic criteria for + test stated?	DVT Prevalence and site n (%)	Clinical symptoms /signs studied	Sens	Spec	PPV	NPV
Simpson 1980 (64)	non-surgical patients with calf pain and suspected DVT n=43; 40% male; age: mean 50.7, range 17-86	CV	no	14 (33%) (5 had DVT and Baker's cyst) site NR	pitting edema, calf swelling>2 cm, cyanosis, SVD, calf pain, warmth: any 5-6 signs	.22	.92	.50	.77
Singer 1980 (60)	non-surgical patients n=92: 46 with + CV, 46 with -CV; 58% male; age: 64± 14	CV	no	45 (50%) by design; 31 (67%) proximal 15 (33%) calf	pain, temperature, color change, induration, tenderness, diff. in calf size: ≥ 4 signs	.72	.65	.67	.69
Sandler 1984 (66)	50 patients with suspected DVT 54% male; age: mean 55 range 18-85	CV	yes	29 (58%); 24 (83%) proximal 5 (17%) calf	pain edema warmth	.86 .97 .72	.19 .33 .48	.60 .67 .65	.50 .88 .53
Vaccaro 1987 (61)	patients who had IPG and CV and retrospective chart review of physical exam n=150	CV	no	68 (45%) site NR	tenderness swelling heat redness Homan	.60 .81 .29 .16 .10	.40 .45 .77 .86 .88	.45 .55 .51 .50 .41	.54 .74 .57 .55 .54

CV=contrast venography NR=not reported SVD= superficial venous dilation Homan=Homan's sign Lowenberg=Lowenberg's sign
diff.= difference inpts= inpatients outpts= outpatients adeq= adequate Sens= sensitivity Spec= specificity PPV= positive predictive value
NPV= negative predictive value

Table 2c. Clinical prediction indices for the diagnosis of DVT

Study	Patient population	DVT reference test, prevalence and site	Methodology	Variables in final model	Predictive accuracy of model
Crandon 1980 (71)	**derivation group**: 124 patients having major gyne. surgery (16% developed DVT) **validation group**: 62 patients having major gyne. surgery	FLS with CV confirmation 10 (16%) developed DVT post-op	-pre-op collection of clinical + lab data -stepwise logistic discriminant analysis	-ELT -age -varicose veins -fibrin antigen -% overweight	Sens .90 Spec .87 PPV .56 NPV .98
Vine 1981 (75)	150 consecutive patients who had CV; 40% male; 57% inpatient	CV 50 (33%) positive	-retrospective chart review of all variables considered relevant -ROC curves with disjunctive analysis for combinations of variables	-malignancy -recent transfusion -recent surgery -congestive heart failure -immobilization -infection -erythema in legs -anemia -swelling	malignancy only: Sens .24 Spec .94 ranging to... all variables Sens 1.0 Spec .16
Lowe 1982 (72)	**derivation group**: 63 patients >age 40 having elective abdominal surgery (33% developed DVT) **validation group**: 41 patients	FLS 18 (44%) positive	-pre-op collection of clinical + lab data -linear discriminant analysis	-age -percent mean weight for age, sex and height	Sens .89 Spec .52 PPV .59 NPV .86
Sue-ling 1986 (74)	128 patients >age 40 undergoing elective abdominal surgery **derivation group**: 85 patients (27% developed DVT) **validation group**: 43 patients	FLS with CV confirmation 14 (33%) developed DVT post-op	-pre-op collection of data -stepwise logistic discriminant analysis	-age -ELT	Sens .93 Spec .83 PPV .72 NPV .96

CV=contrast venography FLS= ^{125}I fibrinogen leg scanning Gyne=gynecological ELT=euglobulin lysis time
ROC=receiver-operating characteristic curve LR=likelihood ratio PPV=positive predictive value NPV=negative predictive value
Sens=sensitivity Spec=specificity

Table 2c. (con'd). Clinical prediction indices for the diagnosis of DVT

Study	Patient population	DVT reference test, prevalence and site	Methodology	Variables in final model	Predictive accuracy of model
Clarke-Pearson 1987 (25)	411 gyne. patients undergoing major abdominal and pelvic surgery	FLS with CV confirmation for positive scans 72 (17.5%) developed DVT post-op	-pre-op and intra-op collection of data -stepwise logistic regression	-type of surgery -anesthesia duration -age -leg edema -race -severity of varicose veins -prior radiation -prior DVT	concordance of 0.82
Landefeld 1990 (76)	355 patients who had CV for suspected proximal DVT **derivation group**: 236 randomly chosen patients **validation group**: remaining 119 patients	CV 27% had proximal DVT	-retrospective chart review to collect data on 76 clinical items -multiple linear discriminant analysis	-swelling below knee -swelling above knee -recent immobility -cancer -fever	0 variables: LR .15 PPV .05 1 variable: LR .47 PPV .15 $\geq$ 2 variables: LR 2.0 PPV 0.42
Wells 1995 (77)	529 outpatients with first episode suspected DVT	CV 25% had DVT: 21% proximal 4% calf	prior development of reliable clinical stratification model based on literature review and clinical experience; testing of model on study population	based on clinical factors, pre-test probability of DVT was estimated to be either: -high -moderate -low	 PPV .85 PPV .33 PPV .05

CV=contrast venography FLS= ^{125}I fibrinogen leg scanning Gyne=gynecological ELT=euglobulin lysis time
ROC=receiver-operating characteristic curve LR=likelihood ratio PPV=positive predictive value NPV=negative predictive value
Sens=sensitivity Spec=specificity

Table 2d. Adherence of published clinical prediction indices to methodological standards for clinical prediction rules

Methodological standard (67,70)	Crandon 1980 (71)	Vine 1981 (75)	Lowe 1982 (72)	Sue-ling 1986 (74)	Clarke-Pearson 1987 (25)	Landefeld 1990 (76)	Wells 1995 (77)
Clear definition of outcome	yes	yes	yes	yes	yes	yes	yes
Blind assessment of outcome	no	N/A	no	no	no	N/A	yes
Precise definition of predictive finding	yes	no	yes	yes	no	yes	yes
For retrospective studies: blind assessment of predictive finding	N/A	no	N/A	N/A	N/A	no	N/A
Relevance of predictive findings	no	yes	no	no	yes	yes	yes
Patient age and sex stated	yes	yes	yes	yes	yes	yes	no
Mathematical technique described	yes	yes	yes	yes	yes	yes	yes
Test of misclassification rate	yes	no	yes	yes	no	yes	no
Prospectively validated in new clinical setting	no	no	no	no	no	no	no
Effects of clinical use prospectively measured	no	no	yes	no	no	no	no

N/A=not applicable

3. Methods

This study was a secondary analysis of a subset of patients who participated in a NHRDP-funded study during 1989-90 at the Montreal General Hospital. The latter study, which compared two methods of diagnosis of deep vein thrombosis, was titled *A Randomized Trial of Impedance Plethysmography (IPG) Versus Contrast Venography (CV) in Patients with a First Episode of Clinically Suspected Deep Vein Thrombosis*. It will be referred to throughout the text as the "IPG-CV study". I refer to my study as the CPOD study (*Clinical Predictors of DVT Study*).

3.1. IPG-CV Study

This section will outline the subjects and methods used in the IPG-CV study. These details are required to understand the design, source population, potential for bias, and generalizability of the CPOD study.

The specific objectives of the IPG-CV Study were to determine the effectiveness of IPG compared to CV for the diagnosis of DVT in patients with a first episode of clinically suspected DVT, and to perform a cost-effectiveness analysis of the diagnostic techniques.

3.1.1. IPG-CV Study population

Patients screened

Beginning January 19, 1989, all consecutive Montreal General Hospital patients over age 18 with a first episode of clinically suspected DVT were screened for inclusion into the IPG-CV study. Patients could be inpatients or outpatients from any hospital ward, department or clinic. A total of 1034 patients were screened.

Exclusion criteria

Patients meeting one or more of the following exclusion criteria were excluded from the IPG-CV study:

1. Clinically suspected recurrent DVT, as defined by prior positive contrast venography and/or prior administration of anticoagulants for DVT or "phlebitis". These patients were excluded because of the difficulties in reliably diagnosing DVT in a limb with abnormal venous architecture caused by prior DVT. Also, patients with prior DVT represent a population at increased risk for DVT, due both to alterations of the venous anatomy as well as possible persistence of underlying risk factors which promoted the original DVT. They therefore are likely to have different characteristics than patients presenting for the first time with leg symptoms.
2. Allergy to contrast material. Allergic reactions to contrast dye can be life-threatening and thus precluded being randomized.
3. Patients on long-term anticoagulants. These patients were excluded for several reasons. The likelihood of developing DVT while taking anticoagulants is very low. Many of the conditions for which anticoagulants are recommended are themselves associated with leg symptoms (e.g. congestive heart failure, peripheral arterial insufficiency). Therefore, including patients taking anticoagulants would have reduced the yield of patients with DVT. Also, those patients who develop DVT while taking anticoagulants undoubtedly are at particularly high risk for DVT, hence including these patients would have limited the generalizability of the study.
4. Pregnancy. Contrast venography is contraindicated during pregnancy due to the hazards of radiation exposure to the fetus. Therefore, pregnancy would preclude randomization.
5. Patient refused or was incapable of giving informed consent. The study protocol stipulated that patients randomized to the IPG arm who had a negative baseline IPG required repeat testing on day 1, day 3-5, and day 10-14. All study patients required follow-up interviews at 3 and 6 months. Therefore, those who refused consent were probably less able to return to the hospital for multiple testing, or would not be available for the 3 and 6 months interviews. These patients were likely a mix of the frail elderly, people who were ill, and full-time workers who found it impractical to return for frequent visits.
6. Patient is geographically unsuited for repeated IPG testing. Patients living far from the

hospital were excluded since they would be unable to complete the study protocol.

7. Patient with acute or chronic renal failure (defined by serum creatinine >200 mmol/L. Contrast venography can worsen renal function and precipitate renal failure in patients with abnormal renal function.
8. Patients with a leg in traction or wearing a plaster cast. IPG, which requires placement of an inflatable cuff on the leg, cannot be performed over a rigid plaster cast. Contrast venography cannot be performed if the foot veins to be injected are covered by plaster, and rotating the patient (to ensure adequate dispersion of contrast dye throughout the venous system) is contraindicated if the leg is in traction.

Table 3a (end of chapter) shows the number of patients excluded for each exclusion criterion. Table 3b (end of chapter) shows the age and sex distribution of patients included in the IPG-CV study versus those excluded from the study. Overall, 60% of patients considered for participation in the study were female, and 66% of excluded patients were female. In each of the eight exclusion criteria, more females than males were excluded.

3.1.2. Study protocol

Information on a number of variables was collected at baseline including demographic data, baseline health status, and symptoms and signs. All data was collected by the study nurses before the diagnostic procedure (IPG or CV) was performed.

Study recruits were randomized to IPG or CV. Patients randomized to IPG had a baseline study, which, if negative, was repeated on day 1, day 3-5, and day 10-14. This is usual practice, because as discussed earlier, IPG is most sensitive for proximal DVT, hence if symptoms are due to a calf DVT that has not yet extended proximally, IPG, initially negative, may later become positive as the thrombus propagates proximally. Patients randomized to CV had a single study only.

Patients with positive CV were treated for DVT. Patients with positive IPG had immediate confirmatory CV and, if positive, they were treated for DVT. Patients with negative serial IPGs or negative CV were not treated.

All test results (IPG and CV) were interpreted blindly by panels of three experts, without knowledge of any clinical information. Disagreements between two panel members were adjudicated by the third member. The IPG panel was composed of two hematologists and one senior IPG technician. The CV panel was composed of three vascular radiologists.

CV was read as positive if all deep veins were adequately visualized and a constant intraluminal filling defect was seen in two or more projections or an abrupt cut-off was seen in a deep vein above the knee. Positive CV was further characterized as calf DVT or proximal DVT, depending on which vein(s) contained clot. CV was read as negative if all deep veins were adequately visualized and were free of intraluminal filling defects. IPGs were read as normal or abnormal, as previously defined in the literature (78).

All randomized patients were followed for six months. During this time, they were instructed to return to the clinic if they experienced recurrent leg pain or swelling. They were contacted by telephone for interviews at 3 and 6 months to obtain the following information: alive or dead, general health status, persistence or resolution of the original symptoms, development of any new symptoms, and the interim requirement for new investigations or treatments for DVT or PE. The primary aim of this follow-up was to ascertain whether or not the diagnosis of DVT had been missed during the initial phase of the study (false negative test result).

The IPG-CV study flow chart is depicted in Figure 3.1 (end of chapter).

3.1.3. Final study population of IPG-CV Study, and adherence to study protocol

324 patients were recruited, all of whom gave informed consent to participate in the study. 165 patients were randomized to IPG, and 159 were randomized to CV. There were 22 cross-overs: 8 from IPG to CV (6 for geographic reasons, 2 had tremor thus IPG could not be performed), and 14 from CV to IPG (7 no venous access, 4 vein rupture, 3 miscellaneous reasons). One patient dropped out after randomization when a diagnosis of cellulitis was made and neither test was performed. 270 (83.3%) completed the 3 month

interview, 286 (88.3%) the 6 month interview, and 242 (74.7%) completed both interviews. Ten patients (3.1%) did not complete any interview. Twenty-seven patients, on follow-up interview, had symptoms suspicious for DVT or PE for which they underwent diagnostic testing. Of the 27, five had confirmed DVT or PE. There were 23 deaths by the 6 month follow-up point.

3.2. CPOD study

3.2.1. Study population

The source population for the CPOD study was the 324 participants in the IPG-CV study. The aim of the CPOD study was to develop a clinical prediction index for the diagnosis of DVT in patients presenting with leg symptoms. Therefore, it was critical to exclude patients in whom the presence or absence of DVT could not be ascertained with certainty, and to carefully classify CPOD study patients into those with confirmed DVT ("cases") and those without confirmed DVT ("controls"). This was achieved using the criteria described below.

Exclusion criteria

Patients were excluded if they dropped out of the study before or immediately after the initial diagnostic test, if they did not complete at least 2 out of 3 serial IPGs, or if their CV was considered to be technically inadequate by the blinded reviewer panel as defined by the protocol of the IPG-CV study, i.e. inadequate visualization of any of the following veins: common iliac, external iliac, common femoral, superficial femoral, popliteal, peroneal, and posterior tibial veins, or failure to obtain two or more views of any intraluminal filling defect(s). Also excluded were any patients with initial negative CV or IPG who were diagnosed with DVT or PE during the 6 month follow-up period.

Definition of confirmed diagnoses ("cases")

Confirmed diagnoses (patients with DVT) were defined as patients from the source population with a positive CV at one or more anatomical sites or patients with a positive

IPG where no CV confirmation was possible and treatment for DVT was administered. I will refer to these patients as "cases". Cases had proximal DVT if, by CV, a clot was found in a proximal vein, whether or not clot was also present in a calf vein, or if IPG was positive and the patient was treated for DVT even if no CV confirmation was possible. Cases had calf DVT if, by CV, a clot was found in a calf vein only.

Definition of non-confirmed diagnoses ("controls")

Non-confirmed diagnoses (patients without DVT) were defined as patients from the source population with a negative CV, even if IPG was positive, or the combination of a negative baseline IPG and at least 2 out of 3 serial IPGs completed and negative and no DVT or PE diagnosed by 3 or 6 month follow-up. I will refer to these patients as "controls".

It is important to note that these patients are not controls in the strict sense of the term as used in case-control studies. They are not representative of controls in the target population (ie. the world of patients with leg symptoms who are free of DVT and who are in the catchment area of the Montreal General Hospital) since they were selected into the study in a biased way: they had leg symptoms that 1) on initial assessment by their physicians, were not attributed to a specific diagnosis, which might have avoided the need for further testing and 2) furthermore, aroused in their physicians the suspicion of DVT, thus leading to referral to the IPG-CV study. Thus, the process by which "controls" were selected into the study was associated with the probability of their having DVT. Nevertheless, this terminology is used because of the analytic approach employed in this study.

From the source population of 324 patients, there were 53 exclusions, 73 cases and 198 controls. There were 9 deaths among cases (12.3%) and 11 deaths among controls (5.5%) by the end of the six month follow-up period. Figure 3.2 (end of chapter) depicts the assembly of the CPOD study population. Tables 3c, 3d, and 3e (end of chapter) summarize the criteria for selection of cases and controls, and criteria for exclusion of patients from the study.

Potential for misclassification

Since positive IPGs were followed by confirmatory CV, and the criteria for classifying CV as positive were strict, a false positive diagnosis of DVT was unlikely. Although in 5 cases with positive IPG, CV confirmation of DVT was not possible, given the predictive accuracy of IPG for proximal DVT of 90-95% in similar populations, it is unlikely that any of these five would be false positive results.

Of greater concern was the possibility of false negative results (i.e. missed DVTs) since, by study protocol, patients with negative IPGs were not required to undergo CV. However, patients with negative IPGs had repeated serial testing, and any patients with IPGs that turned positive on serial testing were referred for confirmatory CV. Patients also had 3 and 6 month telephone interviews to assess for clinically suspected and/or objectively confirmed DVT or PE that could have represented DVTs that were missed initially. Therefore, thorough attempts were made to capture all cases.

Nonetheless, since IPG only reliably detects proximal DVT (specificity 95%, sensitivity 92%), calf DVT that did not propagate proximally by the last serial IPG or that resolved spontaneously without proximal extension could have been missed in patients randomized to undergo IPG. By extrapolation, in this sample, based on finding 21 DVTs restricted to the calf out of 154 CV studies (13.6% prevalence), it is conceivable that out of 86 negative IPGs, 12 (13.6%) calf DVTs could have been missed, 3 (25%) of which would have extended proximally and been ultimately detected by serial IPG, and 9 (75%) of which would likely have remained localized to the calf and spontaneously resolved with time (42). Thus, based on a theoretical worse-case scenario where IPG had 0% sensitivity for calf DVT, the maximum number of missed calf DVTs that would have been misclassified as controls in this sample would be 9.

Adequacy of follow-up of study population

Of the 271 patients, 194 (71.6%) completed both the 3 and 6 month follow-up interviews, 25 (9.2%) completed only the 3 month and 45 (16.6%) completed only the 6 month interviews, for a total of 97.4% of patients with follow-up of at least 3 months. 7 (2.6%) had neither the 3 month nor 6 month follow-up interview, all of whom were in the

control group. (4 had negative IPG, and 3 had negative CV).

3.2.2. Outcome variables

The outcome variable for the primary analysis was DVT (any site), thus cases had DVT (any site) and controls had no DVT. For the secondary analysis, the outcome variable was proximal DVT, thus cases had proximal DVT and controls had no DVT. For the secondary analysis, patients with isolated calf DVT were excluded.

3.2.3. Predictor variables

At study entry, prior to diagnostic testing, information on a number of variables was collected on a standard reporting form by the study nurse, who was therefore blinded to the outcome of the diagnostic tests. Three categories of data were collected:

Demographic information

-Age

-Sex

-Patient location (inpatient vs. outpatient)

Baseline health status

-Medical illness, including type and whether active (congestive heart failure, chronic lung disease, diabetes, liver disease, peripheral vascular disease, collagen vascular disease, other)

-Leg trauma in the last 30 days and number of days immobilized

-Cancer, including type and whether active (breast, lung, colon, prostate, hematological, pancreas, kidney, other)

-Surgery in the last six months, and type (abdominal, gynecological, orthopedic, hip, knee, prostate, thoracic, cardiac, neurologic, other)

Clinical symptoms

-Pain, swelling, redness, any symptom

-Duration of above symptoms

Clinical signs

-Tenderness, pitting edema, non-pitting edema, erythema, superficial venous dilation, warmth, palpable cord, Homan's sign, any sign

Tables 3f-3j (end of chapter) display the outcome and predictor variable names, their description, and how they were coded.

3.2.4. Ethical considerations

The IPG-CV study was approved by the Montreal General Hospital Research and Ethics Committee, and all participating patients gave informed consent. The CPOD study was a secondary analysis of IPG-CV data, using aggregate and not case-specific data, hence confidentiality of study participants was preserved.

3.2.5. Sample size calculation

In regression analyses in general, the larger the number of independent variables included in the model, the better the fit will be to the data set from which the model is built. In the extreme case, if there are as many independent variables as there are subjects in the data set, a model can be found that fits the data exactly, even if none of the variables have any predictive ability at all. In these cases and others less extreme, there can be much doubt about whether the model will be generalizable to other data sets. Therefore, in building regression models, one must balance model parsimony against overfitting to a particular data set, which will occur if too many independent variables are included. Finding the right balance increases the likelihood that the results will be generalizable to other data sets.

However, there are no specific guidelines as to the number of observations recommended per predictor in logistic regression modeling. In linear regression, an often quoted rule is to have at least 10 observations per predictor in the maximum model (79). The CPOD study population of 271 would thus allow consideration of at most 27 predictor variables if a linear model were being used. However, since each subject in a logistic regression model provides only a "yes" or "no" outcome, each subject contributes

less information on average compared to that provided by each subject in a linear regression on a continuous outcome variable. Therefore, higher ratios of subjects to number of predictor variables might be advisable.

Bayesian methods of model selection were used in this thesis. These are specifically designed to balance parsimony versus fit in logistic regression modeling, and can be used with samples of any size. Rather than using backwards or forwards model selection techniques that typically greatly increase type I errors that can lead to overfitting, Bayesian techniques calculate the support for any given model by the data, and compare this to the support for other models. Furthermore, the procedures specifically account for the sample size. Therefore, while larger sample sizes can result in better models with more accurately established parameters, smaller sample sizes are not likely to result in overfitting, as is possible with other techniques (80).

I had no control over the sample size in this study since the data were already collected when I started my project for a completely different purpose than my own. The original sample size calculation was based on a two-sided α of .05 and a β of .20 (power of 80%) to detect a 50% reduction in the anticipated failure rates of IPG (10%) and CV (20%). It should be noted that the final sample size of 324 patients in the IPG-CV study fell short of the 418 needed based on the sample size calculation.

For the final model, I calculated a confidence interval for each parameter (odds ratio), the width of which indicated to what degree of accuracy the value of the coefficient was known. A sufficiently narrow confidence interval thus implied that the sample size was sufficient to accurately estimate the effect of that parameter.

3.2.6. Statistical analysis

The primary analysis was carried out for cases (DVT any site) vs. controls (no DVT), and the secondary analysis was carried out for cases (proximal DVT) vs. controls (no DVT).

Univariate (simple) analysis

Cases and controls were compared on all baseline variables. For continuous

variables, results are presented as mean ± standard deviation, or for variables with heavily skewed distributions, median with interquartile range. Dichotomous variables are presented as proportion (percent) affected in each outcome group.

Means were compared using Student's t-test. Difference in means and the associated 95% confidence interval for the difference[1] are presented. Proportions were compared using chi-square test, or Fisher exact test for cells with expected cell frequency less than 5. Differences in proportions and the associated 95% confidence interval for the difference[2] are presented. Medians were compared using the Mann-Whitney test, a non-parametric two sample median test.

The sensitivity, specificity, PPV, NPV, LR+ and LR- were calculated for all variables that appeared to be good predictors of case status, as judged by the confidence interval for the difference in the variable between cases and controls.

Bivariate (stratified) analysis

Spearman's rank correlation was run on all variables to look for highly correlated variables which might be collinear or confounders in the association between a given variable and case status.

Based on substantive evidence, in an attempt to explain some of the relations found in the univariate and correlation analyses, cases and controls were further stratified on the following variables: presence/ absence of active medical illness, over/ under 65, inpatient/ outpatient, gender, recent trauma, recent surgery, and recent orthopedic surgery. The relationships between predictor variables and case status were examined, and when the strength of the association between a given predictor variable and case status was different in the two strata, effect modification by the stratifed variable on the predictor variable was tested, if substantively plausible. SAS software (release 6.11, SAS systems, Inc.) was used for the univariate and bivariate analyses.

---

1 $X_1 - X_2 \pm 1.96 * \sqrt{(var_1^2/n_1 + var_2^2/n_2)}$

2 $p_1 - p_2 \pm 1.96 * \sqrt{(p_1q_1/n_1 + p_2q_2/n_2)}$

Multivariate analysis

Logistic regression analysis was performed using the bic.logit procedure (S-plus software), which employs Bayesian methods of model selection (80).

Standard frequentist model selection methods (e.g. forward, stepwise, backward) rely on P-value based significance tests to include or exclude a predictor variable from the model. One difficulty with these methods is that P values for individual variables change depending on the number of variables considered in the model and the selection method used, thus use of a P value cut-off as a basis for variable inclusion or exclusion can be dramatically misleading. A second difficulty is that several different models may all seem reasonable yet lead to different conclusions about the question being studied. Thus, selecting a single model and basing inference on it becomes somewhat arbitrary and ignores issues of model uncertainty and the resulting uncertainty of inferences.

The Bayesian approach to hypothesis testing, model selection and accounting for model uncertainty overcomes these difficulties to a large extent. Uncertainty about the unknown parameters of the model is expressed in terms of the probability of the parameter given the data, using Bayes' theorem:

$$p(\theta|D) \propto p(D|\theta)p(\theta)$$

where D= data observed, and θ=the vector of unknown parameters. Thus, the posterior distribution is proportional to the likelihood times the prior. The bic.logit procedure provides estimates of the beta coefficients for each predictor variable using the mean of individual model estimates, weighted by the posterior probabilities of each parameter as defined above (posterior mean). For model selection, the Bayesian Information Criterion (BIC) (also called the Schwarz Criterion (SC)) provides an accurate approximation to Bayes' factor, which is the ratio of the integrated likelihoods of the models being compared, say M_2 vs. M_1. When this ratio is >1, the data favor M_2 over M_1 and the magnitude of Bayes' factor can be used to assess the strength of this evidence. BIC is a function of the likelihood ratio statistic or model deviance:

$$BIC_k = L_k^2 - df_k \log n$$

where L_k^2 is the deviance for model M_k and df_k is the corresponding number of degrees of freedom. Models can be compared by taking the difference of their BIC values, with the model having the smaller (ie. more negative) BIC having better model "fit". As with Bayes' factors, the magnitude of the difference in BICs can be used to assess the strength of evidence for one model against another. Finally, for each model, model uncertainty is expressed as the posterior probability that each model is true, given that one of them must be true. This probability is derived from the BIC (80). Use of these methods and their interpretation for my data will be discussed in the next chapter.

Variables judged to be important predictors of DVT by clinical grounds or by the univariate and bivariate analyses, and interaction terms that were judged to be plausible based on results of the stratified analyses were entered into the bic.logit logistic regression program (80). Final model selection was based on the BIC value.

Standard regression diagnostics were not performed, since it was considered that the BIC criterion of model selection gave the best possible model given the data, and no model could be expected to precisely predict case status for each individual, especially since all predictor variables except age were binary. Hence, some deviation from predicted outcome is to be expected. The width of the 95% confidence interval surrounding the parameter (calculated using $e^{\beta \pm 1.96(S.E.)}$) was used as an estimate of the degree of accuracy with which the parameter (here, the odds ratio) was known.

Model validation is important especially when the model is to be used to predict outcome for future subjects, since the fitted model always performs better on the data set it was derived from (81). Cross-validation techniques using split-samples and other methods (79) were considered for use but were rejected because of the small sample size of the study, which would render parameter estimation in the derivation and validation subsets unstable. Ideally, the model should be validated externally and prospectively on a population of patients presenting with leg symptoms suspicious for DVT. To help assess the model's predictive accuracy within the sample (goodness of fit), I examined, for each covariate pattern, the probability of case status based on the model vs. the probability

observed in the data.

The variables in the final model were used to develop the clinical prediction index. A sensitivity analysis was performed for the different cutoff points of the index, and the corresponding Receiver Operating Characteristics (ROC) curve was plotted.

Chapter 3: Figures and Tables

Figure 3.1. IPG-CV STUDY: study flow chart

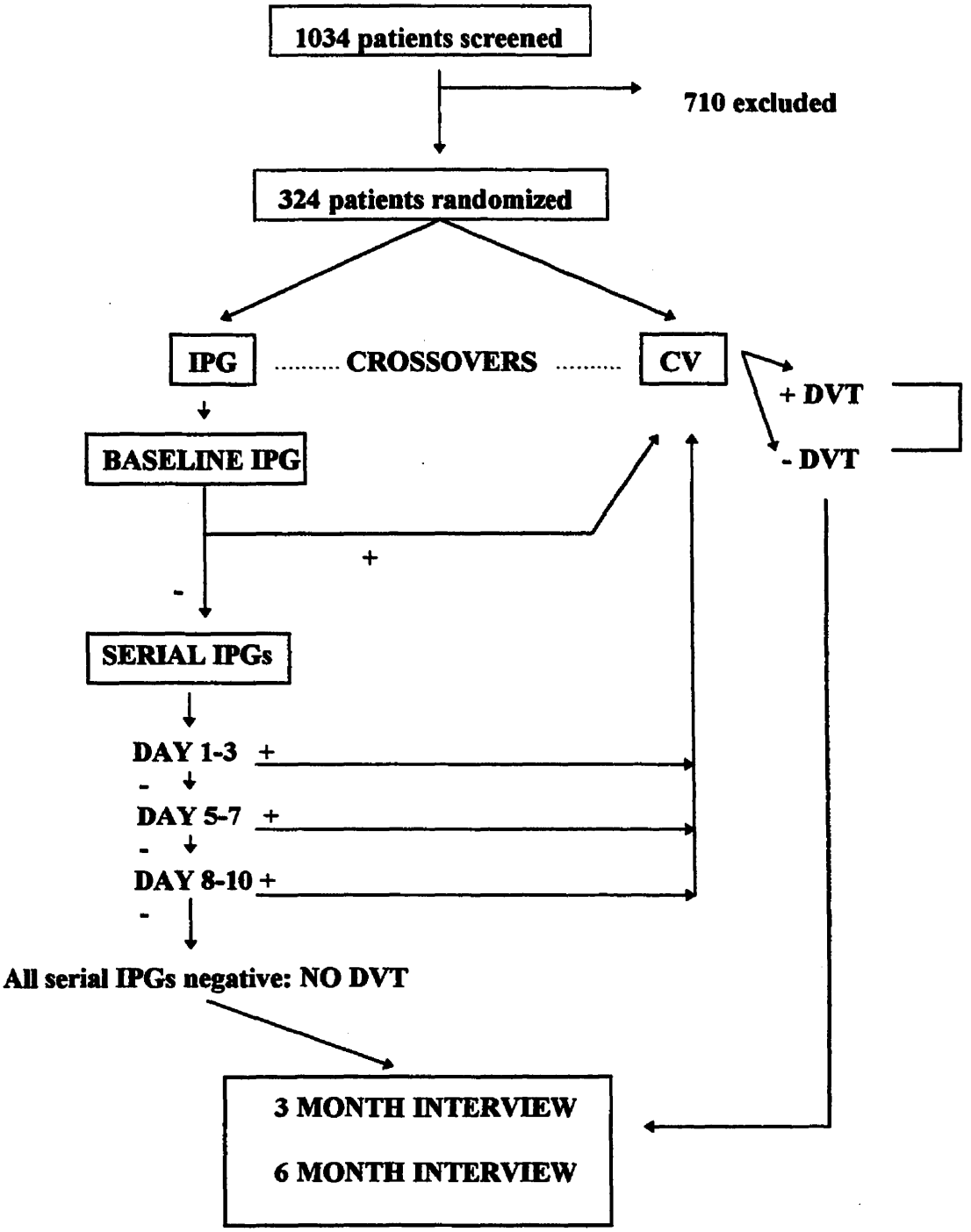


Table 3a. Exclusion criteria for IPG-CV study (total number of patients excluded=710)

EXCLUSION CRITERION	NUMBER OF PATIENTS EXCLUDED*
Suspected recurrent DVT	101
Allergy to contrast dye	34
Long-term anticoagulants	12
Pregnancy	24
Refused/incapable of consent	510
Renal failure	34
Leg in traction/plaster cast	10

Note: * Sum is greater than total (n=710) due to fulfilment of >1 exclusion criterion in some patients

Table 3b. Mean age and sex distribution of patients included and excluded for IPG-CV study (total screened=1034)

	PATIENTS INCLUDED (N=324)	PATIENTS EXCLUDED (N=710)	DIFFERENCE [95% CI]
Age(years)*, mean±SD	56.7±17.2	60.3±18.5	-3.6 [-5.9,2.3]
SEX, % male	53%	34%	.19 [.13,.25]

Note: * ages unavailable for 15 of 710 excluded patients

Figure 3.2. Flowchart illustrating assembly of Clinical Predictors of DVT (CPOD) study population

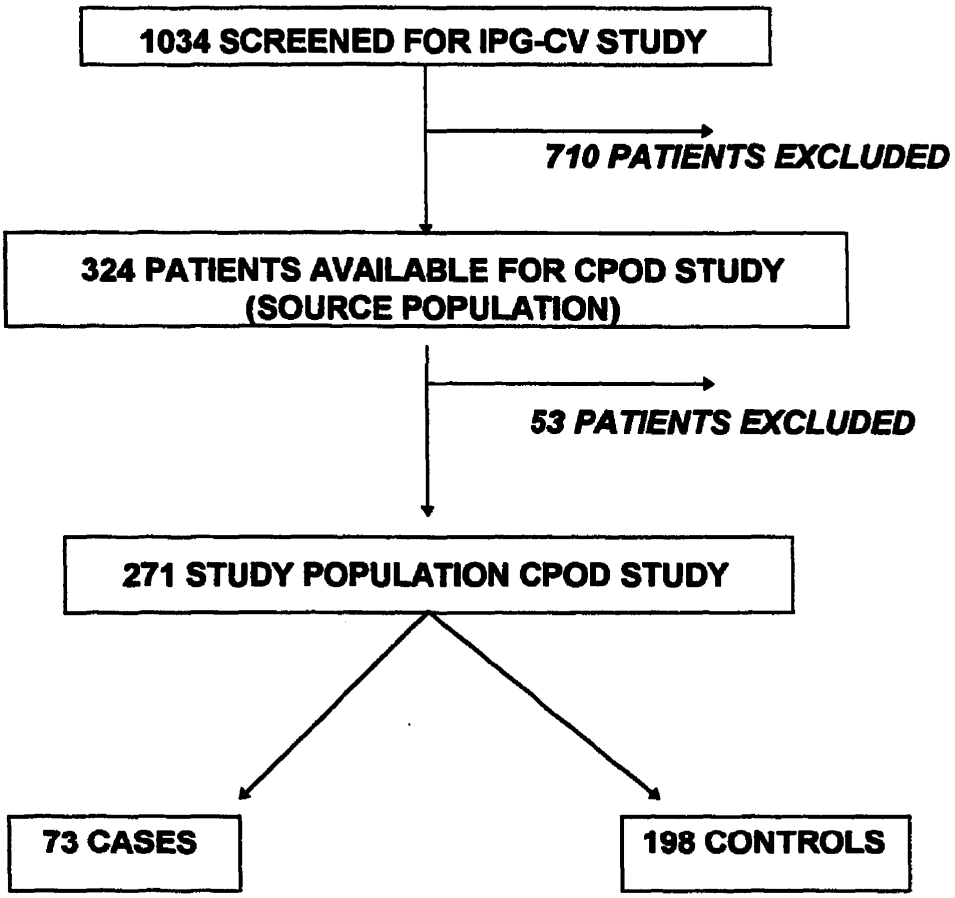


Table 3c. CPOD STUDY POPULATION : definition of CASES, n=73

CRITERION	N
Positive CV	68
Positive IPG, no CV confirmation possible *	5

Notes:
* 3 patients had no venous access; 2 patients had vein rupture

Table 3d. CPOD STUDY POPULATION: definition of CONTROLS, n=198

CRITERION	N
Negative CV	86
Negative baseline IPG and at least 2/3 serial IPGs completed and negative, and no DVT or PE diagnosed by 3 or 6 month follow-up interview	112

Notes for Tables 3c and 3d:

CV= contrast venography IPG= impedance plethysmography
DVT= deep vein thrombosis PE= pulmonary embolus

Table 3e. Reasons for exclusion of patients from CPOD study

REASON FOR EXCLUSION	NUMBER OF PATIENTS
Inadequate CV [1]	37
Incomplete follow-up of negative IPG[2]	9
Drop-outs [3]	2
Pull-outs [4]	5
TOTAL EXCLUDED	53

Notes:

[1] Among the 37 patients with inadequate CV:

29 never had IPG

8 had IPG: of these,

5 had positive IPG

3 had negative IPG: of these,

2 had inadequate follow-up

1 completed all follow-ups with no evidence of DVT

[2] 5 patients had baseline IPG only (ie. none of the serial IPGs were done)
4 patients had baseline IPG and only 1 of 3 serial IPGs

[3] 1 patient had neither IPG nor CV and was treated for cellulitis
1 patient had negative baseline IPG and then dropped out of study

[4] 1 patient had DVT diagnosed at 3 month follow-up, which could have represented a missed DVT
4 patients: blinded interpretation of CV not available

Table 3f. CPOD study: outcome variables

VARIABLE DESCRIPTION	VARIABLE CODING	VARIABLE TYPE
Case status for any DVT	case=1 control=0	dichotomous
Case status for proximal DVT	case=1 control=0	dichotomous

Table 3g. CPOD study: predictor variables- demographic

VARIABLE DESCRIPTION	VARIABLE CODING	VARIABLE TYPE
Patient age	age, in years	continuous
Patient age	years > 65=1 years ≤ 65=0	dichotomous
Gender	male=1 female=0	dichotomous
Patient location	inpatient=1 outpatient=2	dichotomous

Table 3h. CPOD study: predictor variables- baseline health status

VARIABLE DESCRIPTION	VARIABLE CODING	VARIABLE TYPE
Active medical illness (except cancer)	yes=1 no=0	dichotomous
Leg trauma in last 30 days	yes=1 no=0	dichotomous
Number of days immobilized in last 30 days	number of days	continuous
Immobilized $\geq$ 1 day in last 30 days	yes=1 no=0	dichotomous
Prior history of cancer	yes=1 no=0	dichotomous
Active cancer	yes=1 no=0	dichotomous
Surgery in last 6 months	yes=1 no=0	dichotomous
Orthopedic surgery	yes=1 no=0	dichotomous
Hip surgery	yes=1 no=0	dichotomous
Knee surgery	yes=1 no=0	dichotomous

Table 3i. CPOD study: predictor variables- clinical symptoms

VARIABLE DESCRIPTION	VARIABLE CODING	VARIABLE TYPE
Presence of any symptom	yes=1 no=0	dichotomous
Symptom duration	number of days	continuous
Log symptom duration	log number of days	continuous
Presence of pain	yes=1 no=0	dichotomous
Pain duration	number of days	continuous
Swelling noted by patient	yes=1 no=0	dichotomous
Swelling duration	number of days	continuous
Redness noted by patient	yes=1 no=0	dichotomous
Redness duration	number of days	continuous

Table 3j. CPOD study: predictor variables- clinical signs

VARIABLE DESCRIPTION	VARIABLE CODING	VARIABLE TYPE
Any sign noted by examiner	present=1 absent=0	dichotomous
Tenderness on palpation	present=1 absent=0	dichotomous
Pitting edema	present=1 absent=0	dichotomous
Non-pitting edema	present=1 absent=0	dichotomous
Erythema	present=1 absent=0	dichotomous
Superficial venous dilation	present=1 absent=0	dichotomous
Warmth	present=1 absent=0	dichotomous
Palpable cord	present=1 absent=0	dichotomous
Homan's sign	present=1 absent=0	dichotomous

4. Results

In the CPOD study population (n=271) there were 73 patients with DVT (27%). The primary analysis compared cases with DVT (all sites) vs. controls (patients without DVT), and the secondary analysis compared cases with proximal DVT vs. controls (patients without DVT). Before presenting these results, I will briefly present the baseline characteristics of patients included in the CPOD study compared to those excluded. This has relevance as to the generalizability of the results and the potential for selection bias.

As seen in Table 4.1, CPOD study subjects and those excluded from the study were similar on all variables, except that a higher proportion of excluded patients were male. Since many patients were excluded from CPOD because of technically inadequate contrast venography (CV), patients with adequate and inadequate CV were also compared on a number of baseline variables. There were no important differences between these groups (Table 4.2).

Table 4.1. Comparison of CPOD study subjects vs. patients excluded from study

VARIABLE	SUBJECTS (N=271)	EXCLUDED PATIENTS (N=53)	DIFFERENCE [95% CI]
Age(years), mean± SD	57.1±17.0	54.3±18.4	2.8 [-2.6, 5.4]
Sex (% male)	49.4%	69.8%	- 20.4 % [-34, -7]
Location (% in-patient)	26.2%	35.8%	- 9.6 % [-24, 4]
Medical disease at baseline (%)	37.2%	41.5%	- 4.3% [-19, 10]
History of cancer (%)	16.2%	13.2%	3% [-7, 13]
Active cancer (%)	8.8%	3.8%	5% [-1, 11]
Surgery in past 6 months (%)	32.4%	35.8%	- 3.4% [-17, 11]
Orthopedic surgery in past 6 months (%)	8.9%	13.2%	-4% [-14, 5]
Symptom duration(days), median (IQR)	5 (2-14)	4 (1-8)	—

Table 4.2. IPG-CV subjects with technically adequate vs. inadequate contrast venography (CV)

VARIABLE	ADEQUATE CV N=154	INADEQUATE CV N=37	DIFFERENCE [95% CI]
Age(years) mean± SD	57.6±15.7	53.1±18.3	4.5 [-1.9,6.4]
Sex (% male)	57.9%	67.6%	- 9.7% [-27, 7]
Patient location (% inpatient)	25.8%	35.1%	- 9.3% [-26, 8]
Symptom duration(days), median (IQR)	5 (2-14)	4 (2-8)	---
Recent leg trauma (%)	23.9%	24.3%	- 0.4 % [-16, 15]
Surgery in past 6 months (%)	34.0%	37.8%	- 3.8% [-21, 14]
Orthopedic surgery past 6 months (%)	9.4%	13.5%	- 4.1% [-16, 8]
Hip surgery past 6 months (%)	2.5%	5.4%	- 2.9% [-11, 5]
Knee surgery past 6 months (%)	6.9%	8.1%	- 1.2% [-11, 8]
Active intercurrent disease (%)	31.5%	37.8%	- 6.3% [-24, 11]
History of cancer (%)	17.0%	18.9%	- 1.9% [-16, 12]
Active cancer (%)	10.7%	5.4%	5.3% [-3, 14]

4.1. Primary Analysis

4.1.1. Univariate analysis

Demographic variables (Table 4.3)

The mean ages of cases and controls were similar (58.9 ± 16.8 vs. 56.4 ± 17.0 respectively). Among cases, 65.8% were male compared to 43.4% of controls, and DVT occurred in 36% of men but only 18% of women. Thus, in this population DVT was more common in men than women. Among cases, 38.4% were inpatients compared to 21.7% of controls.

Table 4.3. Demographic data in cases and controls

VARIABLE	CASES n=73	CONTROLS n=198	95% CI FOR DIFFERENCE	P VALUE
Age (years), mean±SD	58.9±16.8	56.4±17.0	2.5 [-2.0,4.5]	.29
Sex, % male	65.8	43.4	22% [10, 35]	.001
Patient location, % inpatient	38.4	21.7	17% [4, 29]	.006

Baseline health status

i. Active medical disease (Table 4.4)

Active medical disease (excluding cancer) was significantly more common in controls (41.4%) than cases (26%). When broken down by disease type, diabetes and "other illness" (which was not further characterized in the data entered into the database but likely included illnesses such as hypertension, angina and arthritis) were more common in controls than cases (8.1% vs. 1.4%, and 32.3% vs. 17.8% respectively).

Table 4.4. Active medical disease in cases and controls

VARIABLE	CASES n=73	CONTROLS n=198	DIFFERENCE [95% CI]	P VALUE
Active disease %	26	41.4	- 15.4% [-27,-3]	.02
Type of disease %				
Congestive heart failure	0	1.0	- 1% [-2, .3]	1.0
Chronic lung disease	0	0.5	- 0.5% [-1, .5]	1.0
Diabetes	1.4	8.1	- 6.7% [-11, -2]	.048
Liver disease	0	2.5	- 2.5% [-5, -.3]	.33
Peripheral vascular disease	1.4	0.5	0.9% [-2, 4]	.47
Collagen vascular disease	0	0.5	- 0.5% [-1, .5]	1.0
Other	17.8	32.3	-14.5% [-25, -4]	.02

ii. Immobilization and leg trauma (Table 4.5)

Immobilization for more than one day in the last 30 days was more common in cases than controls (24.7% vs. 9.6% respectively). Similar results were seen when the cut-off was moved to more than 0, 2, 3, 4, or 5 days of immobilization. At more than 6 days of immobilization, the difference between groups was lost. This may have been due to the small number of patients that had immobilization for 6 days or longer. Among patients immobilized for any length of time in the last 30 days, the median number of days of immobilization was similar in cases and controls (5 days in both groups). Hence, immobilization for a period of time up to 5 days was a predictor of DVT in the univariate analysis. I did not find evidence that the actual duration of immobilization was a predictor,

perhaps due to the small number of patients with immobilization. For the remainder of the analysis immobilization was dichotomized into more than one day vs. one day or less, since this cut-off is easily quantifiable clinically.

Leg trauma in the last month occurred in 30.1% of cases and 20.7% of controls, but the 95% CI for this difference included zero ([-3%, 21%]).

Table 4.5. Recent leg trauma and immobilization in cases and controls

VARIABLE	CASES N=73	CONTROLS N=198	DIFFERENCE [95% CI]	P VALUE
Immobilized >1 day in last 30 days (%)	24.7	9.6	15.1% [4, 26]	.002
Among patients immobilized in last 30 days, number of days immobilized, median (IQR)	5 (2-8) (n=24)	5 (1-10) (n=27)	----	.95
Recent leg trauma (%)	30.1	20.7	9.4% [-3, 21]	.103

iii. Cancer (Table 4.6)

Active cancer was present in 15.1% of cases and 6.6% of controls. Because of the small patient numbers in each category, I could not investigate whether any individual cancer type was particularly thrombogenic. A past history of cancer that was currently inactive was found in similar proportions of cases and controls.

Table 4.6. Cancer in cases and controls

VARIABLE (%)	CASES N=73	CONTROLS N=198	DIFFERENCE	[95% CI]	P VALUE
Active cancer %	15.1	6.6	8.5%	[-.4, 17]	.03
History of cancer %	21.9	14.1	7.8%	[-3, 18]	.12
Type of cancer: [1] %				[-6, 5]	
Breast [2]	4	4.5	- 0.5%		1.0
Lung	4.1	1.0	3.1%	[-2, 8]	.12
Colon	0	1.0	- 1%	[-2, .3]	1.0
Prostate [3]	4.2	1.2	3 %	[-2, 8]	.29
Hematological	1.4	1.0	0.4%	[-3, 3]	1.0
Pancreas	0	0	0	---	---
Kidney	0	0.5	- 0.5%	[-1, .4]	1.0
Other	13.7	7.6	6.1%	[-3, 15]	0.12

Notes: [1] Some patients had >1 cancer type [2] denominator = women [3] denominator = men

iv. Surgery and orthopedic surgery (Table 4.7)

More cases than controls had surgery in the last six months (43.8% vs. 28.3%, 95% CI for difference [3, 28]). When analyzed by surgery type, almost all of this difference was attributable to orthopedic surgery. Among cases, 19.2% had orthopedic surgery compared to 5.1% of controls (95% CI for difference [5, 24]). With regard to the site of orthopedic surgery, 11% of cases had knee surgery, compared to 2% of controls (95% CI for difference [2, 16]). For hip surgery, the difference in proportions among cases (8.2%) and controls (3.5%) was not statistically significant.

Table 4.7. Recent surgery, and type, in cases and controls

VARIABLE (%)	CASES N=73	CONTROLS N=198	DIFFERENCE [95% CI]		P VALUE
Surgery in last 6 months[1] %	43.8	28.3	15.5 %	[3, 28]	.015
Type of surgery %					
Abdominal	9.6	6.6	3%	[-5, 11]	.40
Gynecological [2]	4	0	4 %	[-.5, 8]	.18
Orthopedic	19.2	5.1	14.1%	[5, 24]	.001
Hip	8.2	3.5	4.7%	[-2, 11]	.12
Knee	11.0	2.0	9 %	[2, 16]	.004
Prostate [3]	2.1	0	2.1%	[-1, 5]	.36
Thoracic	1.4	1.5	- 0.1%	[-3, 3]	1.0
Cardiac	0	2.0	- 2%	[-4, -.5]	.58
Neurologic	2.7	2.5	0.2%	[-4, 5]	1.0
Other	11.0	12.6	- 1.6%	[-10, 7]	.71

Notes: [1] some patients had > 1 surgery type, hence total is less than sum
[2] calculated in women [3] calculated in men

Clinical symptoms (Table 4.8)

Patients were recruited for participation in the IPG-CV study because of suspected DVT. Thus, by design, symptoms were not expected to be good discriminators between case and control status. This was confirmed by the presence of "any symptom" in 100% of cases and 99.5% of controls. Similarly, with regard to individual symptoms, pain, swelling and redness were found equally in cases and controls. Although there was a trend to longer median duration of symptoms in controls (6 days, interquartile range (IQR) [3-20])

than cases (4 days, IQR [2-11]), this difference did not achieve statistical significance. One might expect that the more chronic the symptoms, the less likely they are due to DVT, since DVT is an acute illness that progresses if not treated.

Table 4.8. Type and duration of symptoms in cases and controls

VARIABLE	CASES N=73	CONTROLS N=198	DIFFERENCE [95% CI]	P VALUE
Any symptom (%)	100	99.5	0.5% [-.5, 1]	1.0
Global symptom duration (days), median (IQR)	4 (2-11)	6 (3-20)	---	.25
Log mean symptom duration(days) ± SD	1.75±1.3	2.11±1.4	-.37 [-.74,.37]	.06
Pain (%)	78.8	78.1	0.7% [-10, 12]	.90
Number of days of pain (days), median (IQR)	3 (1-7)	4 (1-9)	---	.07
Swelling (%)	83.6	77.2	6.7% [-4, 17]	.26
Number of days of swelling (days), median(IQR)	2 (1-7)	4 (1-11)	---	.20
Redness (%)	68.2	61.6	6.6% [-6, 19]	.31
Number of days of redness (days), median (IQR)	1 (0-6)	3 (0-7)	---	.08

Clinical signs (Table 4.9)

All cases and 97.5% of controls had at least one of eight clinical signs. However, for individual signs, superficial venous dilation (SVD) and warmth of the lower extremity occurred in significantly higher proportions of cases than controls (28.8% vs. 14.7% respectively for SVD; 71.2% vs. 49.5% respectively for warmth). Of the two, SVD is likely to be more specific for DVT, since it is caused by diversion of blood from obstructed deep veins to superficial veins. Warmth, on the other hand, can be caused by a variety of inflammatory conditions including DVT, and on substantive grounds would not be expected to be a good predictor of case status. Homan's sign was also more prevalent in cases than controls (52.1% vs. 38.4% respectively).

Table 4.9. Clinical signs in cases and controls

VARIABLE %	CASES n=73	CONTROLS n=198	DIFFERENCE [95% CI]	P VALUE
Any sign	100	97.5	2.5% [.3, 5]	.33
Tenderness	82.2	78.3	3.9% [-7, 14]	.48
Pitting edema	41.1	39.9	1.2% [-12, 14]	.86
Non-pitting edema	78.1	69.7	8.4% [-3, 20]	.17
Erythema	48.0	35.9	12.1% [-1, 25]	.07
Superficial venous dilation	28.8	14.7	14.1% [3, 26]	.008
Warmth	71.2	49.5	21.7% [9, 34]	.001
Palpable cord	15.1	14.7	0.4% [-9, 10]	.93
Homan's sign	52.1	38.4	13.7% [.4, 27]	.04

The presence of a palpable venous cord, shown by Richards to be highly specific for DVT (65), occurred with equal prevalence in cases (15.1%) and controls (14.7%), leading one to wonder what anatomical structure was being palpated in controls. As expected by their high prevalence in the general population and their multifactorial etiology, the clinical signs edema, erythema and tenderness were poor predictors of case status.

Summary of test characteristics of potential predictor variables (Table 4.10)

Table 4.10 below displays the sensitivity, specificity, predictive values and likelihood ratios for each variable that was shown to be associated with case status in the univariate analysis.

No predictor had both high sensitivity and specificity. Sensitivity was poor for all predictors, with warmth having the highest estimated sensitivity (.71). High estimated specificities were found for SVD (.85), immobilization (.90), active cancer (.93), and orthopedic and knee surgery (.95 and .98 respectively). No predictor had high PPV (the highest PPV was .67 for knee surgery). Warmth and male sex had the highest NPVs (.83 and .82 respectively).

The likelihood ratio-positive (LR+) is the ratio of sensitivity to 1-specificity. It represents the odds of a positive "test" (here, the presence of a given predictor) in cases

compared to controls. The likelihood ratio-negative (LR-) is the ratio of 1-sensitivity to specificity. It represents the odds of a negative "test" (here, the absence of a given predictor) in cases compared to controls (82). A test that is highly discriminatory would have a LR+ >>1, and a LR- close to 0. The variables with the highest LR+ were knee surgery (5.5), orthopedic surgery (3.8), and immobilization (2.5). The variables with the lowest LR- were warmth (.58) and male sex (.60).

Table 4.10. Accuracy of individual predictor variables for the diagnosis of DVT

Predictor	Sens	Spec	PPV	NPV	LR+	LR-
Male sex	.66	.57	.36	.82	1.5	.60
In-patient location	.38	.78	.39	.78	1.7	.79
Active illness	.26	.59	.19	.68	.63	1.3
Immobilization	.25	.90	.49	.76	2.5	.83
Active cancer	.15	.93	.46	.75	2.1	.91
Recent surgery	.44	.72	.36	.78	1.6	.78
Orthopedic surgery	.19	.95	.58	.76	3.8	.85
Knee surgery	.11	.98	.67	.75	5.5	.91
SVD	.29	.85	.42	.76	1.9	.84
Warmth	.71	.50	.35	.83	1.4	.58
Homan's sign	.53	.62	.33	.78	1.4	.76

Notes: Sens= sensitivity Spec= specificity PPV= positive predictive value NPV= negative predictive value LR+ = likelihood ratio for positive test LR- = likelihood ratio for negative test

4.1.2. Bivariate analysis

Spearman's rank correlation (Table 4.11)

Spearman's nonparametric rank correlation was computed for all variables. An $|r|$ of $\geq .20$ was considered to be a conservative indicator of a potentially important correlation between two variables. Table 4.11 shows the r values for these variable pairs. Not shown in the table are variable pairs for which high correlation would be expected, such as presence of a symptom and its duration, surgery and surgery type, swelling and edema, etc.

While none of the correlations were very strong, inpatient status was correlated with surgery, orthopedic surgery, knee surgery, and immobilization, whereas being an

outpatient was correlated with symptom duration. Trauma was correlated with surgery, orthopedic surgery, knee surgery, hip surgery, and immobilization. Superficial venous dilation was correlated with age. Immobilization was correlated with surgery, orthopedic surgery, hip surgery, and knee surgery. Finally, immobilization was negatively correlated with symptom duration. Of note, sex, active medical illness and active cancer were not correlated with any other predictor variables (not shown).

Table 4.11. Spearman's rank correlation for variables with $|r| \geq .2$

	Case status	Outpatient location	Age	Surgery	Trauma	Immobilization
Surgery		-.48			.29	.39
Orthopedic surgery	.22	-.29			.41	.36
Knee surgery	.20	-.24			.31	.23
Hip surgery					.29	.31
Symptom duration		.27				-.22
Immobilization	.20	-.40		.39	.39	
SVD			.23			

Stratified analyses

Stratified analyses were performed on selected variables in order to explore potential confounding effects of a given variable on the association between a predictor variable and case status. The choice of stratifying variables was both hypothesis-driven, based on previously known confounding effects between variables, and data-driven, to try to explain associations found in the univariate analysis that were not known in the literature to predict case status.

The stratified analyses were also used to explore potential effect modification, which would be suspected if there was a difference in the strength of the association between case status and one predictor variable for the two levels of the stratification variable. For effect modification, only *a priori*, substantively plausible effects were

considered for entry into the multivariate model.

For confounding effects, the data presented are only for variables that showed a significant difference in proportions, as determined by the 95% CI for this difference, between the two levels of a given stratification variable. For effect modification, the data presented are for variables that showed a significant association with case status in one level of the stratum and a non-significant association with case status in the other level of the stratum, as determined by the 95% CI for the difference in proportions.

i. Stratified by age (Table 4.12)

Knee surgery was more common in those under 65 than those over 65, whereas hip surgery and SVD were more common in those over 65 than those under 65.

Table 4.12. Selected variables in those under 65 compared to those over 65

VARIABLE	UNDER 65 % n=170	OVER 65 % n=101	95% CI FOR DIFFERENCE
Knee surgery	6.5	1.0	[-10, -1]
Hip surgery	2.4	9	[.6, 13]
SVD	12.4	28.7	[6, 26]

ii. Stratified by sex (Table 4.13)

Homan's sign and warmth were more common in men than women. There were no sex differences noted for the following variables (not shown): age, active medical illness, active cancer, orthopedic, knee or hip surgery, and immobilization.

Table 4.13. Selected variables in males compared to females

VARIABLE	MALES % n=134	FEMALES % n=137	95% CI FOR DIFFERENCE
Homan's sign	53	31.4	[-33, -10]
Warmth	61.9	48.9	[-25, -1]

iii. Stratified by patient location (Table 4.14)

Table 4.14. Selected variables in inpatients compared to outpatients

VARIABLE	INPATIENTS % n=71	OUTPATIENTS % n=200	95% CI FOR DIFFERENCE
Over 65	50.7	32.5	[5, 32]
Surgery	70.4	19	[39, 63]
Ortho surgery	22.5	4	[8, 29]
Knee surgery	12.7	1.5	[3, 19]
Hip surgery	11.3	2.5	[1, 16]
Immobilization	36.6	5.5	[19, 43]

There were significantly higher proportions of inpatients who were over 65, had surgery, orthopedic surgery, knee surgery and hip surgery, and who were immobilized compared to outpatients. Median symptom duration was 3 (IQR 1-6) in inpatients compared to 7 (IQR 3-21) in outpatients, P value for difference 0.0001 using a Mann-Whitney test.

iv. Stratified by active medical illness (except cancer)

Significantly fewer patients with active illness had surgery, compared to those without active illness (24.7% vs. 37.1%, 95% CI for difference [1, 24]).

v. Stratified by surgery

A higher proportion of surgery patients than patients who did not have surgery had been immobilized (33% vs. 4.4%, 95% CI for difference [18, 39]).

vi. Stratified by orthopedic surgery (Table 4.15)

Immobilization and trauma were more common among patients who had orthopedic surgery. There was no active cancer among orthopedic surgery patients.

Table 4.15. Selected variables in patients with and without orthopedic surgery

VARIABLE	ORTHOPEDIC SURGERY % n=24	NO ORTHOPEDIC SURGERY % n=247	95% CI FOR DIFFERENCE
Immobilization	54.2	9.7	[24, 65]
Trauma	79.2	18	[44, 78]
Active cancer	0	9.7	[-13, -6]

vii. Stratified by trauma (Table 4.16)

Among trauma patients, there were higher proportions of male sex, surgery, orthopedic, knee and hip surgery and immobilization compared to patients without trauma. Trauma patients had less baseline disease compared to patients without trauma.

Table 4.16. Selected variables in patients with and without trauma

VARIABLE	TRAUMA % n=63	NO TRAUMA % n=208	95% CI FOR DIFFERENCE
Male sex	60.3	46.2	[.2, 28]
Surgery	57	25	[18, 46]
Ortho surgery	30.2	2.4	[16, 39]
Knee surgery	15.9	1	[6, 24]
Hip surgery	15.9	1.4	[5, 24]
Immobilization	38.1	6.3	[19, 44]
Baseline disease	25.4	40.9	[-28, -3]

Summary of potential confounding effects

Recognizing potential confounding effects is important for interpreting the univariate analysis, for designing and interpreting the multivariate analysis, and for selecting the most appropriate variables to include in the clinical prediction rule. From the correlation and stratified analyses, the following conclusions are drawn:

1. There is no *a priori* reason to believe that inpatient location, per se, is a risk for DVT. The apparent association between inpatient location and case status was likely confounded by one or more of the following variables, which were present in higher

proportions of inpatients compared to outpatients: age over 65, surgery, orthopedic surgery, knee surgery, hip surgery and immobilization.

2. Male sex is not a known risk for DVT. However, the stratified analysis failed to show any confounding effects that could explain the association between male sex and case status found in the univariate analysis.
3. Trauma is a known risk factor for DVT. However, in this sample, the association between trauma and case status, although weak, might have been confounded by the higher prevalence of male sex, surgery, orthopedic surgery, hip surgery, knee surgery and immobilization in trauma patients compared to patients without trauma.
4. Increasing age is a recognized risk for DVT. The lack of association in this sample between age and case status could have been confounded by the higher proportions of patients with knee surgery in those less than 65 years old.
5. The apparent association between SVD and case status could have been confounded by the higher proportion of patients over 65 who had SVD.
6. The inverse of the expected association between active illness and case status could have been confounded by the lower rates of surgery in those with active illness.
7. The apparent association between surgery and case status could have been confounded by the higher rates of immobilization in patients who had surgery. Alternately, since surgery is usually followed by a period of immobilization, there might be an important degree of collinearity between these variables.
8. The apparent association between orthopedic surgery and case status could have been confounded by the higher rates of immobilization and trauma in patients who had orthopedic surgery. Alternately, since trauma may lead to orthopedic surgery, and since both are followed by a period of immobilization, these three variables might be providing the same prediction information.

If this had been an etiological study of DVT, identification of and adjustment for confounders would be critical to the understanding of disease causality. However, for the development of a clinical prediction rule, identification of confounders is important

primarily to ensure inclusion of the most relevant predictors in the prediction rule. In some cases, it may be more relevant to retain the confounding variable in the prediction rule instead of the true "risk factor", provided that it helps in the prediction of DVT. This would be the case if, for example, the confounder was easier to measure than the true risk factor.

Effect modification

SVD was a predictor of DVT in males but not females (Table 4.17). Thus, SEX*SVD was identified as a potential effect modifier with the hypothesis that, since SVD is more common in females due to their higher prevalence of varicose veins (note the prevalence of SVD in 18.8% of female controls but only 9.3% of male controls), SVD might have better predictive accuracy for DVT in males than females.

Table 4.17. SVD in cases and controls, stratified by sex

Level of stratification variable	Cases n=73	Controls n=198	P value	95% CI for difference
Males n=134	29.2% n=48	9.3% n=86	.003	[6, 34]
Females n=137	28% n=25	18.8% n=112	.30	[-9, 28]

Although orthopedic surgery and knee surgery were also better predictors of DVT in males than females (data not shown), there were no theoretical reasons to believe that this represented true effect modification.

There were other variable pairs for which effect modification was plausible substantively, e.g. higher risk of DVT in older patients with surgery or immobilization compared to younger patients with surgery or immobilization; higher risk of DVT in orthopedic surgery patients with trauma compared to non-orthopedic surgery patients with trauma. However, since these theoretical effects were not substantiated by the data (data not shown), they were not pursued further.

4.1.3. Multivariate analysis

Logistic regression

A matrix of twenty-seven predictor variables was entered into the bic.logit program (80) of S-plus, a logistic regression program that uses the Bayesian Information Criterion (BIC) to select the best model out of all possible models (here, 2^{27} possible models). The variables included were those associated with case status in the univariate analysis, and potential confounding variables as described in the bivariate analysis. The dependent variable was case status, where cases had DVT (any site) and controls had no DVT.

For model selection, the option "strict = T" was used, which excludes models that are 20 or more times less likely than the most likely model (this corresponds to a BIC difference of 6 between models) and models that are likely to be nested within better models.

The most probable model, given the data, contained the variables SEX (male=1 female =0), DIS (active illness, present=1, absent=0), ORTHO (recent orthopedic surgery, yes=1, no=0), and SVD (superficial venous dilation on exam, present=1, absent=0). Table 4.18 displays the characteristics of this model and the next 3 most likely ones:

Table 4.18: Four most likely models generated by bic.logit, S-plus

Model order of likelihood	Variables in model	P (Model | data)	BIC	Δ BIC
1st	SEX, DIS, ORTHO, SVD	.28	-1177.779	--
2nd	SEX, ORTHO, SVD, WARMTH	.26	-1177.681	.1
3rd	SEX, ORTHO, SVD	.19	-1177.068	.7
4th	SEX, ORTHO, WARMTH	.05	-1174.460	3.3

The posterior probabilities, or p (model | data), are interpreted as "out of all the models currently being considered, these are the relative probabilities of each model assuming one of them must be true". The posterior probability that the 1st model was the

true model was .28. For the 2nd and 3rd models, the respective posterior probabilities were .26 and .19. The 4th most likely model had a probability of .05, ie. was about 6 times less likely to be "true" than the 1st model. The BIC for the 1st model was -1177.779, which was marginally different from the BICs of the 2nd and 3rd models. The Δ BIC between the 1st and 4th model was 3.3, which indicated positive, but not strong evidence that the 1st model was more likely than the 4th. Table 4.19, adapted from Raftery (80), shows general guidelines for preference of one model over another.

Table 4.19. Grades of evidence for one model against another

BIC difference	p (model | data)	Evidence
0-2	50-75	weak
2-6	75-95	positive
6-10	95-99	strong
>10	>99	very strong

It is evident that the 1st model generated by the data showed only weak-to-positive evidence of being more likely than the 4th most likely model. Therefore, given the data, the first three models were probably about equally likely, and these were only slightly more predictive than the 4th model. This indicated that there was no one subset of variables among those in the variable matrix that had a superior ability to explain the data over other subsets.

As discussed, the 1st and 2nd models had similar BICs and posterior probabilities, indicating no real preference of one over the other. I selected the 2nd model for further consideration, since, in the first model, the parameter for DIS (active illness) was negative, indicating a protective effect. This is contrary to the literature on DVT risk factors, and my hypothesis as to why it appeared to be protective in this data set will be discussed in the next chapter.

The selected 2nd model (which I call model A) was logit(p) = -2.4302 + 1.0133 SEX + .7634 WARMTH + 1.6910 ORTHO + 1.0536 SVD. The characteristics of model A are detailed in Table 4.20.

Table 4.20. Model A

VARIABLE	PARAMETER ESTIMATE	STANDARD ERROR	ODDS RATIO	95% CI FOR ODDS RATIO
SEX	1.01	0.33	2.8	1.5, 5.1
WARMTH	.76	0.31	2.1	1.2, 3.9
ORTHO	1.69	0.47	5.4	2.2, 13.6
SVD	1.05	.35	2.9	1.4, 5.7
Intercept	-2.43	0.33	---	---

Notes:

1. Variable coding: SEX: male=1 female=0 WARMTH limb warmth on exam 1=yes 0=no ORTHO: recent orthopedic surgery 1=yes 0=no SVD: superficial venous dilation 1=yes 0=no
2. Odds ratio $= e^{\text{parameter}}$
3. 95% CI$= e^{\text{parameter} \pm 1.96(\text{standard error})}$

In model A, the adjusted odds ratio for male sex as a predictor of DVT was 2.8, for warmth on exam was 2.1, for orthopedic surgery was 5.4, and for SVD was 2.9.

Since both SEX and SVD were included in this model, and the bivariate analysis suggested interaction between these two variables, the next step was to enter the interaction term SEX*SVD term into the variable matrix.

The model containing the interaction term (model B) was logit(p) = -2.27 + .72 SEX + .80 WARMTH + 1.72 ORTHO + .43 SVD + 1.2 SEX*SVD. The parameter estimates, standard errors and odds ratios for model B are shown in Table 4.21:

Table 4.21. Model B

VARIABLE[1]	PARAMETER ESTIMATE	STANDARD ERROR	ODDS RATIO[2]	95% CI FOR ODDS RATIO[3]
SEX	.72	0.35	2.1	1.1, 4.1
WARMTH	.80	0.32	2.2	1.2, 4.1
ORTHO	1.72	0.47	5.6	2.2, 14
SVD	.43	.53		
			FEMALES 1.5	.54, 4.3
			MALES 10.5	1.9, 13.5[4]
SEX*SVD	1.2	.73		
Intercept	-2.28	0.34	---	---

Notes:

[1] Variable coding: SEX: male=1 female=0 WARMTH: limb warmth present=1 absent=0 ORTHO: recent orthopedic surgery 1=yes 0=no SVD: superficial venous dilation 1=yes 0=no SEX*SVD: =1 if male and SVD; =0 if female; =0 if male and no SVD

[2] Odds ratio $= e^{\text{parameter}}$

[3] 95% CI$= e^{\text{parameter} \pm 1.96(\text{standard error})}$

[4] 95% CI for odds ratio of interaction term obtained using a Monte Carlo simulation (83)

In model B, the adjusted odds ratio for male sex as a predictor of DVT was 2.1, for warmth was 2.2, for orthopedic surgery was 5.6, for SVD in a female was 1.5, and for SVD in a male was 10.5 ($e^{.72+.43+1.2}$). Note that for SVD in women, the 95% CI for the odds ratio included zero.

The BIC difference between model A and model B was 2.4, positive but not strong evidence in favor of model A. The posterior probability of model A (.0014) was three times greater than that of model B (.004), and the odds ratios of the parameters of model B were less stable than those from model A. Nevertheless, because of the substantive plausibility of the interaction term in model B, both models were retained for further analysis.

Goodness of fit

To assess the internal validity of the models, a goodness of fit analysis was performed comparing, in each covariate category, the probability of DVT predicted by models A and B compared to that observed in the data (Table 4.22). This technique was also used as a regression diagnostic device, in that the fit of the two competing models could be directly compared. For reasons discussed in the methods section, standard regression diagnostics were not performed.

Overall, the predicted and observed probabilities were very close (within $\pm$ 20% of each other), indicating good fit between the models and the data. As expected, for covariate patterns with small or zero cell size, the fit cannot be determined because the percent observed is very unreliable. Also, since the parameter estimates represent an averaging across the data, patients with covariate patterns of small cell size would be more likely, by chance alone, to deviate from this "average".

In general, model B was slightly more accurate in predicting case status than model A, especially for covariate patterns where SVD=1 (e.g. covariate patterns 1001, 1101, 0001, 0101), suggesting that there might be effect modification between sex and SVD in the data set. This was supported by the observation that for women, the increase in observed probability of DVT when SVD was present was .02 (covariate patterns 0000 vs. 0001), but for men was .31 (covariate patterns 1000 vs. 1001). It is also possible that

model B predicted better simply because it contains an extra variable, and hence will automatically fit the observed data better.

Table 4.22. Goodness of fit analysis: Predicted probability of DVT vs. that observed, by covariate pattern

Covariate Pattern				Cell size	Predicted probability*		Observed probability
SEX	WARMTH	ORTHO	SVD	N=271	Model A	Model B	
1	0	0	0	42	.19	.17	.19
1	1	0	0	59	.34	.32	.29
1	1	1	0	10	.74	.72	.80
1	1	1	1	0	.89	.93	N/A
1	0	0	1	8	.41	.52	.50
1	0	1	1	0	.79	.86	N/A
1	1	0	1	14	.60	.70	.71
1	0	1	0	1	.57	.54	1.0
0	0	0	0	54	.08	.09	.09
0	1	0	0	45	.16	.19	.20
0	1	1	0	5	.51	.56	.40
0	1	1	1	1	.75	.66	1.0
0	0	0	1	9	.20	.14	.11
0	0	1	1	2	.58	.47	0
0	1	0	1	16	.35	.26	.31
0	0	1	0	5	.32	.36	.40

Notes: N/A = not assessable due to cell size of zero

* Predicted probability calculated from:

Model A: $p = \frac{\exp(-2.4302 + 1.0133\ SEX + .7634\ WARMTH + 1.6910\ ORTHO + 1.0536\ SVD)}{1 + \exp(-2.4302 + 1.0133\ SEX + .7634\ WARMTH + 1.6910\ ORTHO + 1.0536\ SVD)}$

Model B: $p = \frac{\exp(-2.27 + .72\ SEX + .80\ WARMTH + 1.72\ ORTHO + .43\ SVD + 1.2\ SEX*SVD)}{1 + \exp(-2.27 + .72\ SEX + .80\ WARMTH + 1.72\ ORTHO + .43\ SVD + 1.2\ SEX*SVD)}$

Final model chosen

As presented above, neither of the two models was indisputably the "best". Model A was chosen as the final model for the following reasons: it had more negative BIC and higher posterior probability than model B, and the odds ratios for its parameters had slightly narrower 95% confidence intervals, implying more accurate parameter estimation. In this analysis, I used non-informative priors (prior probabilities) for all the variables,

despite the availability of literature on risk factors for DVT. The reasons for this were to avoid making potentially inaccurate assumptions regarding the applicability of prior risk data to my study population, and the great difficulty of summarizing past literature into a single prior distribution. Had I inputted informative priors in favor of the interaction term, the posterior probabilities for it would have been higher.

The issue of which of the two models is superior could be more definitively settled in a second data set with a higher sample size, which was not available here.

Confounding

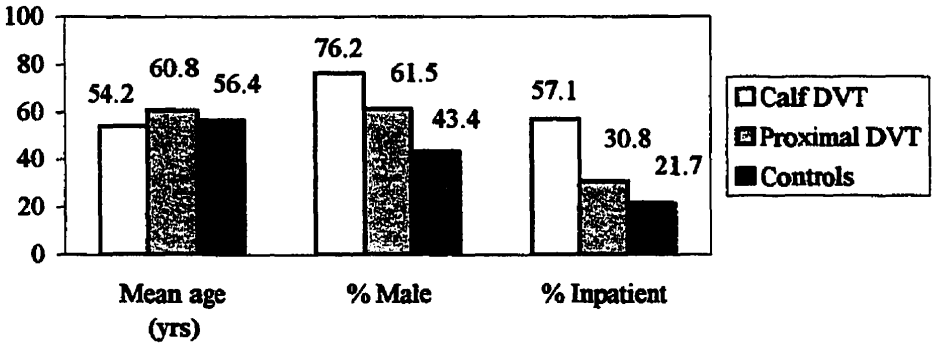
The suspected confounding influences of immobilization, patient location, and trauma were confirmed by the failure of these variables to appear in the final model, and by noting that the maximum likelihood estimates (MLE) for these parameters changed meaningfully when relevant predictor variables were added to the model (e.g. MLE for patient location was -.93 in a model that did not contain the variable for orthopedic surgery, and -.57 in a model that included orthopedic surgery). Interestingly, sex was confirmed as an independent predictor of case status.

4.2. Secondary Analysis

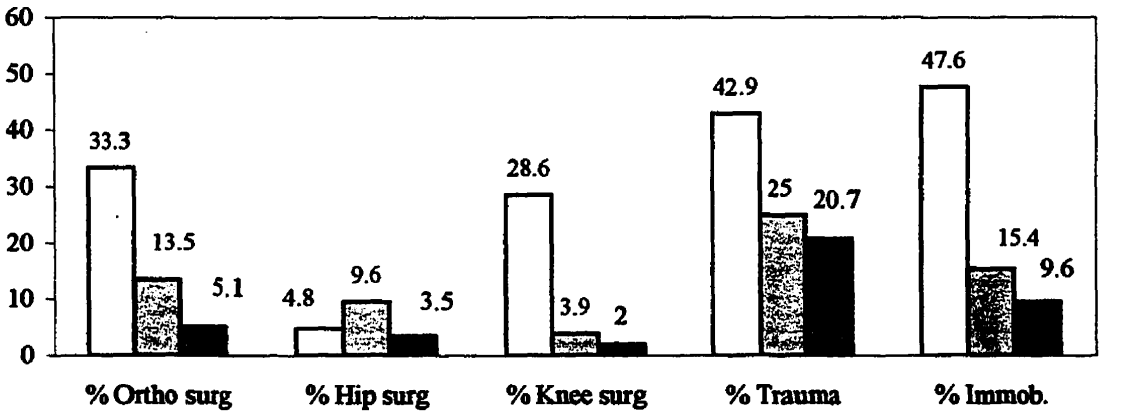
For the secondary analysis, the sample was restricted to cases with proximal DVT (n=52) vs. controls, ie. patients without DVT (n=198). These results will be presented in less detail, with the principle aim of highlighting differences between this and the primary analysis.

Patients with isolated calf DVT (n=21) were removed from this analysis. These patients had different characteristics than patients with proximal DVT (see Graphs 1,2 and 3 below). Patients with calf DVT had a higher proportion of male sex, inpatient location, knee surgery, trauma and immobilization than proximal DVT or control patients, and a lower prevalence of SVD, swelling, and hip surgery than proximal DVT patients.

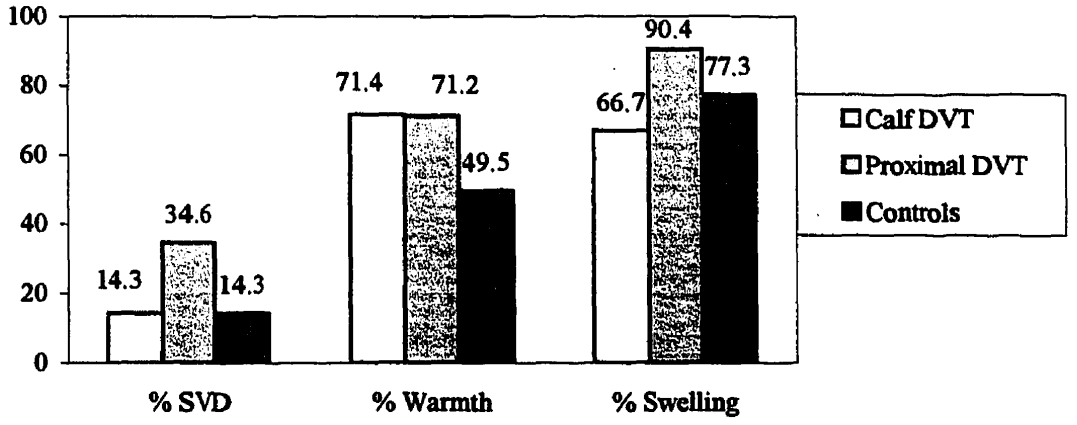
Graph 1. Demographic variables by DVT site



Graph 2. Surgery, trauma and immobilization by DVT site



Graph 3. Clinical signs by DVT site



4.2.1. Secondary analysis: Univariate analysis

Demographic variables (Table 4.23)

Mean age was similar in cases and controls. Once again, a higher proportion of cases (61.5%) than controls (43.4%) were male. The difference in patient location between cases and controls was still present but less apparent than for the primary analysis.

Table 4.23. Secondary analysis. Demographic data in cases and controls

VARIABLE	CASES n=52	CONTROLS n=198	95% CI FOR DIFFERENCE	P VALUE
Age (years), mean±SD	60.8±16.3	56.4±17.0	4.4 [-.63, 5.0]	0.10
Sex, % male	61.5	43.4	18.1% [3, 33]	0.02
Patient location, % inpatient	30.8	21.7	9.1% [5, 23]	0.17

Baseline health status

i. Active medical illness

Active illness (excluding cancer) was again significantly more common in controls than cases (41.4% vs. 25%, 95% CI for difference [-30, -3]).

ii. Immobilization and leg trauma (Table 4.24)

In contrast to the primary analysis, there were no differences in immobilization and leg trauma between cases and controls.

Table 4.24. Secondary analysis. Recent leg trauma and immobilization in cases and controls

VARIABLE	CASES N=52	CONTROLS N=198	DIFFERENCE [95% CI]	P VALUE
Immobilized >1 day in last 30 days (%)	15.4	9.6	5.8% [-5, 16]	0.23
Among patients immobilized in last 30 days, number of days immobilized, median(IQR)	7 (1-11.5) (n=12)	5 (1-10) (n=27)	----	0.43
Recent leg trauma (%)	25	20.7	4.3% [-9, 17]	0.50

iii. Cancer

Similar proportions of cases and controls had active cancer and a history of cancer.

iv. Surgery and orthopedic surgery (Table 4.25)

Neither surgery in general nor orthopedic surgery were significantly more prevalent in cases than controls, although for orthopedic surgery, primarily hip surgery, a trend in this direction was noted.

Table 4.25. Secondary analysis. Recent surgery, and type, in cases and controls

VARIABLE (%)	CASES N=52	CONTROLS N=198	DIFFERENCE [95% CI]	P VALUE
Surgery in last 6 months %	40.4	28.3	12.1% [-3, 27]	.093
Type of surgery %				
Orthopedic	13.5	5.1	8.4% [-1, 18]	.056
Hip	9.6	3.5	6.1% [-2, 15]	.14
Knee	3.9	2.0	1.9% [-4, 8]	.61

Clinical symptoms (Table 4.26)

Swelling was noted more often by cases than by controls (90.4% vs. 77.3% respectively, 95% CI for difference [3, 23]). There were no differences in the prevalence of other symptoms or the duration of any symptoms between cases and controls.

Table 4.26. Secondary analysis. Type and duration of symptoms in cases and controls

VARIABLE	CASES N=52	CONTROLS N=198	DIFFERENCE [95% CI]	P VALUE
Any symptom (%)	100	99.5	0.5% [-.5, 1]	1.0
Global symptom duration (days), median (IQR)	5 (2-14)	6 (3-20)		.49
Log mean symptom duration (days) ± SD	1.85±1.23	2.11±1.4	- .26 [-.66, .40]	.24
Pain (%)	76.9	78.1	- 1.9% [-15, 11]	.77
Number of days since pain onset (days), median (IQR)	2.5 (0-7.5)	4 (1-9)		.21
Swelling (%)	90.4	77.3	13.1% [3, 23]	.035
Number of days since swelling onset(days), median(IQR)	3 (1-9.5)	4 (1-11)		.64
Redness (%)	61.5	68.2	- 6.7% [-21, 8]	.24
Number of days since onset redness (days), median (IQR)	2 (0-7.5)	3 (0-7)		.55

Clinical signs (Table 4.27)

Table 4.27. Secondary analysis. Clinical signs in cases and controls

VARIABLE %	CASES n=52	CONTROLS n=198	DIFFERENCE [95% CI]	P VALUE
Any sign	100	97.5	2.5% [.3, 5]	1.0
Tenderness	78.9	78.3	0.6% [-12, 13]	.93
Pitting edema	44.2	39.9	4.3% [-11, 19]	.57
Non-pitting edema	82.7	69.7	13% [.9, 25]	.057
Erythema	48.1	35.9	12.2% [-3, 27]	.11
Superficial venous dilation	34.6	14.7	19.9% [6, 34]	.001
Warmth	71.2	49.5	21.7% [8, 36]	.005
Palpable cord	19.2	14.7	4.5% [-7, 16]	.42
Homan's sign	50	38.4	11.6% [-4, 27]	.13

Non-pitting edema was noted in 82.7% of cases and 69.7% of controls. Warmth was noted in 71.2% of cases and 49.5% of controls. SVD was seen in 34.6% of cases and 14.7% of controls.

Summary of test characteristics of univariate predictor variables (Table 4.28)

When compared to the corresponding table for the primary analysis (Table 4.10), note that the LR+ and LR- for male sex, active illness and warmth were unchanged, for

orthopedic surgery were less extreme, and for SVD were more extreme. Swelling and non-pitting edema are new entries on the table, but had poor LR+ (1.2 each). Inpatient location, immobilization, active cancer, recent surgery, hip and knee surgery, and Homan's sign were not significant predictors of case status in the secondary analysis.

Table 4.28. Secondary analysis. Accuracy of individual predictor variables for the diagnosis of DVT

Predictor	Sens	Spec	PPV	NPV	LR+	LR-
Male sex	.62	.57	.27	.85	1.4	.67
Active illness	.25	.59	.25	.80	.61	1.3
Orthopedic surgery	.13	.95	.41	.81	2.6	.92
Swelling	.90	.23	.24	.90	1.2	.43
Non-pitting edema	.83	.30	.18	.87	1.2	.57
SVD	.35	.85	.38	.83	2.3	.76
Warmth	.71	.50	.27	.87	1.4	.58

4.2.2. Secondary analysis: Bivariate analysis

Spearman's rank correlation (Table 4.29)

Spearman's nonparametric rank correlation was computed for all variables. Displayed below are the r values for variable pairs that were correlated with $|r| \geq .2$, as well as variable pairs that, in the primary analysis, were correlated with $|r| \geq .2$.

Table 4.29. Secondary analysis. Spearman's rank correlation.

	Case status	Outpatient	Age	Surgery	Trauma	Immobilization
Surgery		-.47			.22	.35
Orthopedic surgery	.14	-.22			.32	.32
Knee surgery	.05	-.09			.17	.11
Hip surgery		-.22			.29	.34
Symptom duration		.25				-.18
Immobilization	.07	-.35		.35	.35	
SVD	.21		.21			
Warmth	.18					

Compared to the primary analysis, the correlations between case status and orthopedic surgery, knee surgery, and immobilization were much weaker.

Secondary analysis: Stratified analyses

i. Confounding

To a degree similar to that seen in the primary analysis, inpatients had more surgery, orthopedic (knee and hip) surgery, age over 65 and immobilization than outpatients. Patients over 65 had more SVD than those under 65. Homan's sign was more common in males than females. Patients with trauma had more surgery, orthopedic surgery (hip) and immobilization than patients without trauma. Patients with orthopedic surgery were more often immobilized than patients without orthopedic surgery (data for above not shown). For the secondary analysis, many of these relationships were not relevant due to the lack of association between case status and patient location, surgery, trauma or immobilization.

ii. Effect modification (Table 4.30)

SEX*SVD was identified as a potential effect modifier, and was included in the variable matrix for the multivariate analysis. There was no evidence from the bivariate analysis for other effect modification.

Table 4.30. Secondary analysis. SVD in cases and controls, stratified by sex

Levels of stratification variable	Cases	Controls	P value	95% CI for difference
Males	37.5%	9.3%	.001	[6, 34]
n=118	n=32	n=86		
Females	30%	18.8%	.24	[-10, 33]
n=132	n=20	n=112		

4.2.3. Secondary analysis: Multivariate analysis

Using the bic.logic function of S-plus (80), as described in the primary analysis, the best model by BIC criteria contained the same variables as for the primary analysis. This model (model C) was logit(p) = -2.7308 + .8773 SEX + .8449 WARMTH + 1.409 ORTHO + 1.2242 SVD. Table 4.31 details the parameter estimates, standard errors and odds ratios for each term in model C.

Table 4.31. Model C: best model by BIC criterion, secondary analysis

VARIABLE[1]	PARAMETER ESTIMATE	STANDARD ERROR	ODDS RATIO[2]	95% CI FOR ODDS RATIO[3]
SEX	.88	.35	2.4	1.2, 4.8
WARMTH	.85	.35	2.3	1.2, 4.7
ORTHO	1.41	.56	4.1	1.4, 12.3
SVD	1.22	.37	3.4	1.6, 7.0
Intercept	-2.73	.38	---	---

Notes:
[1] Variable coding: SEX: male=1 female=0 WARMTH: warmth of limb present on exam yes=1 no=0
ORTHO: recent orthopedic surgery 1=yes 0=no SVD: superficial venous dilation 1=yes 0=no
[2] Odds ratio $= e^{\text{parameter}}$ [3] 95% CI $= e^{\text{parameter} \pm 1.96(\text{standard error})}$

In model C, the adjusted odds ratio for male sex as a predictor of proximal DVT was 2.4, for warmth on limb exam was 2.3, for orthopedic surgery was 4.1, and for SVD was 3.4.

Next, the interaction term SEX*SVD was added to the variable matrix. The following model (model D) was obtained: logit(p) = -2.5368 + 0.5037 SEX + .8802 WARMTH + 1.4474 ORTHO + .5197 SVD + 1.3609 SEX*SVD. The model parameters, their standard errors and odds ratios are shown in Table 4.32.

Table 4.32. Model D

VARIABLE[1]	PARAMETER ESTIMATE	STANDARD ERROR	ODDS RATIO[2]	95% CI FOR ODDS RATIO[3]
SEX	.50	.40	1.7	.75, 3.6
WARMTH	.88	.36	2.4	1.2, 5
ORTHO	1.45	.55	4.3	1.4, 12.8
SVD	.52	.56		
			FEMALES 1.7	.56, 5
			MALES 11.2	2.3, 18.9[4]
SEX*SVD	1.36	.78		
Intercept	-2.54	.38	---	---

Notes:
[1] Variable coding: SEX: male=1 female=0 WARMTH: limb warmth present=1 absent=0
ORTHO: recent orthopedic surgery 1=yes 0=no SVD: superficial venous dilation 1=yes 0=no
SEX*SVD: =1 if male and SVD; =0 if female; =0 if male and no SVD
[2] Odds ratio $= e^{\text{parameter}}$ [3] 95% CI $= e^{\text{parameter} \pm 1.96(\text{standard error})}$
[4] 95% CI for odds ratio of interaction term obtained using a Monte Carlo simulation(83)

In model D, the adjusted odds ratio for male sex as a predictor of DVT was 1.7, for warmth was 2.4, for orthopedic surgery was 4.3, for SVD in a female was 1.7, and for

SVD in a male was 11.2 ($e^{.50+.52+1.4}$). Note that for sex, and for SVD in women the 95% CI of the odds ratio included zero.

The difference in BIC between model C and model D was 2.3, indicting positive but not strong evidence in favor of model C. Model C was 3 times more probable than model D (posterior probabilities .01 and .003 respectively).

Goodness of fit

The covariate patterns were examined for both models to see how well the models predicted the probability of case status compared to that observed in the data.

Table 4.33. Secondary analysis. Goodness of fit: predicted probability of DVT vs. that observed, by covariate pattern

Covariate Pattern				Cell size	Predicted probability*		Observed probability
SEX	WARMTH	ORTHO	SVD	N=250	Model C	Model D	
1	0	0	0	40	.14	.12	.15
1	1	0	0	53	.27	.24	.21
1	1	1	0	4	.60	.45	.50
1	1	1	1	0	.84	.90	N/A
1	0	0	1	7	.35	.46	.43
1	0	1	1	0	.69	.79	N/A
1	1	0	1	13	.55	.67	.69
1	0	1	0	1	.39	.36	1.0
0	0	0	0	51	.06	.07	.04
0	1	0	0	45	.13	.16	.20
0	1	1	0	4	.38	.45	.25
0	1	1	1	1	.68	.58	1.0
0	0	0	1	9	.18	.12	.11
0	0	1	1	2	.47	.36	0
0	1	0	1	15	.34	.24	.27
0	0	1	0	5	.21	.25	.40

Notes:
N/A = not assessable due to cell size of zero
*predicted probability calculated from:

Model C: $p=\frac{\exp(-2.7308 + .8773(SEX) + .8449(WARMTH) + 1.409(ORTHO) + 1.2242(SVD)}{1 + \exp(-2.7308 + .8773(SEX) + .8449(WARMTH) + 1.409(ORTHO) + 1.2242(SVD)}$

Model D: $p=\frac{\exp(-2.5368 + .5037(SEX) + .8802(WARMTH) + 1.4474ORTHO) + .5197(SVD) + 1.36(SEX*SVD)}{1+\exp(-2.5368 + .5037(SEX) + .8802(WARMTH) + 1.4474ORTHO) + .5197(SVD) + 1.36(SEX*SVD)}$

Both models showed good ability to predict case status. For covariate patterns with large cell size, model D predicted slightly more accurately than model C. The data again supported effect modification between sex and SVD: in women, the increase in

observed probability of DVT when SVD was present was .07 (covariate patterns 0000 vs. 0001), but in men was .28 (covariate patterns 1000 vs. 1001).

Final model chosen

The final model chosen for the secondary analysis was model C, for reasons similar to those discussed in the primary analysis. Although effect modification between sex and SVD was substantively plausible and seemed to occur in the data at least on bivariate analysis, there was not enough evidence from the multivariate analysis to support the model that included the interaction term.

Summary of multivariate analyses, primary and secondary analysis

For both the primary and the secondary analysis, the final models contained the predictor variables male sex, warmth on exam, orthopedic surgery, and SVD on exam. The adjusted odds ratios and the corresponding 95% confidence intervals for these variables were 2.8 [1.5, 5.1], 5.4 [2.2, 13.6], 2.1[1.2, 3.9] and 2.9 [1.4, 5.7], respectively, for DVT (all sites), and 2.4 [1.2, 4.8], 4.1 [1.4, 12.3], 2.3 [1.2, 4.7] and 3.4 [1.6, 7.0], respectively, for proximal DVT.

4.3 Clinical prediction index

For the clinician faced with a symptomatic patient in whom DVT is suspected, a clinical prediction index for DVT that is easy to apply is likely to be more practical than the logistic regression equation it is based on, even though some predictive power is inevitably lost in the transition from model to index. I created a simple clinical prediction index for DVT by dividing the study population into categories of patients with 0, any 1, any 2, and any 3 or more predictors. The predictors were the variables included in the regression model, ie. male sex, warmth on exam, orthopedic surgery, and SVD on exam. I did not weight these predictors, even though the adjusted odds ratios of some variables were larger than others. The 95% CI for the odds ratios between variables overlapped, suggesting that the higher odds ratios for some predictors may have been due to sampling

variation rather than to true effects. Weighting would have produced a more complicated index with little gain.

Table 4.34 shows the number of patients with and without DVT and proximal DVT according to the number of predictors present.

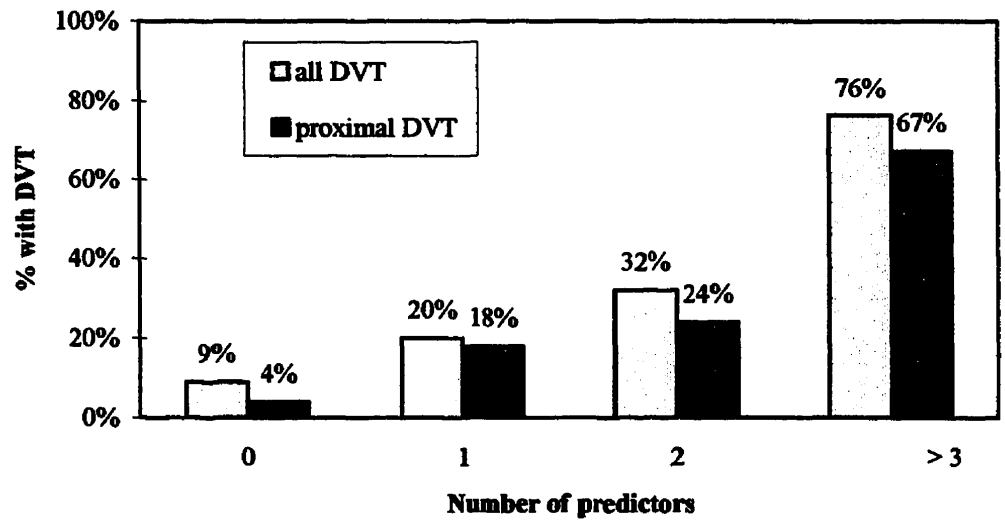
Table 4.34. DVT and proximal DVT vs. number of predictors present

DVT	Number of predictors							
	0		1		2		≥3	
	DVT	proximal DVT	DVT	proximal DVT	DVT	proximal DVT	DVT	proximal DVT
present (n)	5	2	20	18	29	20	19	12
absent (n)	49	49	81	81	62	62	6	6
TOTAL	54	51	101	99	91	82	25	18

Graph 4 shows that the probability of DVT and proximal DVT increases as the number of predictors present increases.

Graph 4

Percent with DVT and proximal DVT by number of predictors present



Next, to create the clinical prediction index, I chose cutoff points that might separate patients with high vs. low probability of DVT. The first cutoff was the presence

of *one or more vs. zero* predictors, the second cutoff was the presence of *two or more vs. one or less* predictors, and the third cutoff was the presence of *three or more vs. two or less* predictors. Table 4.35 depicts the changes in sensitivity, specificity, LR+, LR- , predictive accuracy and misclassification rate for each cutoff point of the predictive index. The results were similar for the prediction of proximal DVT (not shown).

As the number of predictors required for a positive "test" increased, specificity increased at the expense of sensitivity. The PPV, NPV, LR + and misclassification rate were most favorable for the cutoff *three or more vs. two or less*, however only 25 patients fell above this cutoff point, meaning that 54 patients with DVT would have been classified as having low probability of DVT. Using the cutoff *one or more vs. zero*, only 5 DVT patients would have been classified as low probability of DVT, but 81 patients without DVT would have been classified as high probability of DVT. Using the cutoff *two or more vs. one or less*, 25 DVT patients would have been classified as low probability of DVT and 68 patients without DVT would have been classified as high probability of DVT.

Table 4.35. Clinical prediction index: test characteristics of different cutoff points

	Cutoff point: number of predictors		
Test characteristic	1 or more[1]	2 or more[2]	3 or more[3]
Sensitivity	.93	.66	.26
Specificity	.25	.66	.97
Positive predictive value (PPV)	.31	.53	.76
Negative predictive value (NPV)	.91	.84	.78
Likelihood ratio-positive (LR +)	1.2	1.9	8.7
Likelihood ratio-negative (LR -)	.28	.52	.72
Misclassification rate*	57%	34%	22%

Notes:

* calculated by (# false positives + # false negatives) / N

[1] Reference category: patients with 0 predictors

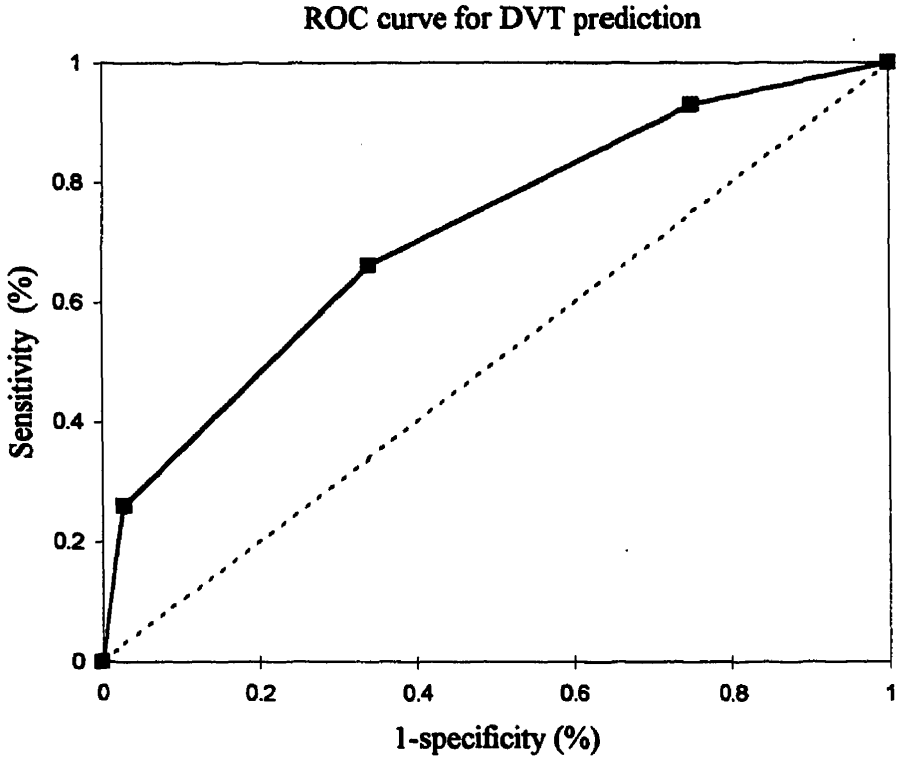
[2] Reference category: patients with 0 or 1 predictors

[3] Reference category: patients with 2 or less predictors

"Gold standard" used was number of patients with and without DVT in the study population

Graph 5 depicts the Receiver Operating Characteristics (ROC) curve for the clinical prediction index. The ROC curve for the diagnosis of proximal DVT was similar (not shown).

Graph 5



Notes:

Data points are, from left to right, 'more than 3 vs. two or less' predictors, 'more than two vs. one or less' predictors, and 'more than one vs. zero' predictors.

Diagonal represents prediction that is no better than pure chance (ie. sensitivity + specificity = 1)

5. Discussion

In medicine, there are certain symptoms and signs for which the diagnosis is virtually unmistakable: for example, the protruding eyes and stare of Grave's disease, and the characteristic vesicles of chicken pox. In these examples, the likelihood of the diagnosis is so high that ancillary information obtained by clinical history, clinical exam and diagnostic testing adds little to the diagnostic process. In contrast, for clinical presentations that carry a broad differential diagnosis with a wide spectrum of probabilities, this ancillary information is key to arriving at the correct diagnosis.

The clinical presentation of DVT falls into this latter category. As presented in the literature review, the clinical symptoms and signs of DVT are widely prevalent and non-specific, and both DVT and the conditions that mimic DVT are common enough in the general population that the two may occur together. For the clinician faced with a patient in whom DVT is suspected, any factor or combination of factors that improve estimation of the prior (ie. pre-test) probability of DVT beyond simply what might be known about the usual prevalence of DVT in similar populations should result in more appropriate selection of patients for diagnostic testing and more informed choice of the particular test used to diagnose DVT.

In this thesis, using logistic regression techniques and Bayesian model selection methods, I analyzed which variables and combinations of variables were associated with an increased probability of a confirmed diagnosis of DVT and proximal DVT in the study population. I then developed a clinical prediction index, and studied how accurately it estimated the pre-test probability of DVT in patients referred for suspected DVT.

5.1. Overview of important findings

5.1.1. Simple analyses

In the univariate analysis, many previously known risk factors for DVT were associated with case status. Compared to controls, cases were 2.5 times more likely to

have been immobilized for more than one day in the last 30 days, 1.5 times more likely to have had recent trauma or active cancer, 4 times more likely to have had orthopedic surgery and 5 times more likely to have had knee surgery in the last six months.

Patients recovering from abdominal (20) and gynecological (25) surgery have a 20-30% incidence of DVT, as diagnosed by ^{125}I fibrinogen leg scan (a screening tool for DVT). Although in this sample more cases than controls had abdominal and gynecological surgery, these differences did not achieve statistical significance, perhaps due to the small number of patients in each group, or perhaps because these surgeries confer an increased risk of asymptomatic, but not symptomatic DVT.

Hip surgery was not shown to be associated with case status, but knee surgery was. Although both are known to be major risk factors for DVT, the probable explanation for the observed difference in this sample is that at the time the IPG-CV study population was assembled, pre-operative thromboprophylaxis was routinely used only in hip surgery patients, hence many of the DVTs that might have developed in this group were averted. Alternately, since the question posed to patients was for surgery any time in the last six months, the exact timing of DVT in relation to surgery could not be ascertained-perhaps surgery was performed more recently in those with knee surgery than those with hip surgery. Although the highest risk period for DVT is immediately following surgery, it has been reported that this risk persists for two to three months post-operatively (20). In any case, although in this sample patients presenting with knee surgery and symptoms of DVT were more likely to have confirmed DVTs than those presenting with hip surgery and symptoms of DVT, I chose not to consider the site of orthopedic surgery to be relevant for the clinical prediction index.

Although many epidemiological studies have suggested a link between trauma and DVT, only one prospective study showed trauma to be an independent risk factor for DVT(21). In this study, however, trauma was only weakly associated with confirmed DVT in the univariate analysis, most of which was due to confounding by orthopedic surgery and immobilization.

Patients with medical illness such as stroke and myocardial infarction are at increased risk for DVT (20). In this sample, however, active medical disease, particularly

diabetes and "other" disease, was present twice as often in controls than cases, and was confirmed as an independent predictor of control status in the initial most likely model generated by the multivariate analysis (adjusted OR .45, 95% CI [.24, .86] ie. "protective"; complete model not shown in Results). There are a number of possible explanations for this finding. "Other" illness was not well-defined on the original data collection sheet, so it is not certain which diseases were in this category. It is unlikely, however, that any disease, except for rare bleeding disorders, would be truly protective against DVT. Since fewer patients with active disease than those without active disease had surgery (since they are generally at higher risk for surgical complications), perhaps active disease was more prevalent in controls simply because patients with active disease were not exposed to surgery, a strong risk factor for DVT. However, the stratified analysis showed that even among patients who did not have surgery, active disease was more common in controls than cases (data not shown).

The most likely explanation for the higher prevalence of active medical disease in controls than cases is that among patients with active disease, leg symptoms and signs are common and are more often caused by factors related to the underlying disease than by DVT. For example, in diabetics, leg pain is more likely to be caused by neuropathy or impaired arterial circulation than by DVT, and similarly, leg pain and swelling in patients with arthritis are more likely to be caused by joint inflammation than by DVT. When such patients present with leg symptoms, they are less likely than the average individual to have DVT as the cause of their symptoms. However, because active medical disease was not protective in an epidemiological sense, and because I could not be certain which diseases were included in the "other" category, the model containing the variable for active disease was not retained for further consideration, despite the fact that active disease was a good predictor of not having confirmed DVT.

Despite good evidence from incidence studies that the risk for DVT increases with age (15,16), in this patient sample there was no association between age and DVT. Neither mean age nor proportion of patients aged over 65 differed between cases and controls in the univariate and multivariate analyses. Although it is possible that age did not appear to be an independent predictor of DVT in this sample because it was strongly

correlated with other, more powerful risk factors (ie. collinearity), this was not borne out by either the correlation analysis or the stratified analysis. However, in the stratified analysis age was associated with active disease. Among patients with active disease, 74% of cases vs. 38% of controls were over 65. Perhaps the "protective" effect of active disease confounded the weaker risk for DVT conferred by age, which, due to the relatively small sample size, was not powerful enough to come out as an independent risk for DVT in the multivariate analysis.

It is unlikely that the increased risk of DVT with increasing age, as reported in DVT incidence studies, simply reflects the higher prevalence in the aged of other known risks for DVT (e.g. cancer, surgery, immobilization). Numerous multivariate analyses in different hospital populations have confirmed the independent effect of age on DVT risk (25,71,72,74). The lack of association between age and confirmed DVT in this sample will be addressed further in the section on the internal validity of this study.

Male sex was a predictor of DVT in both the univariate and multivariate analyses (adjusted OR 2.8 [1.5, 5.1]). Although a higher proportion of males than females had active cancer (11.9% vs. 8%), knee surgery (6% vs. 2.9%), and immobilization (17.2% vs. 10.2%), all of which were associated with DVT in the univariate analysis, these differences are unlikely to entirely explain the difference in DVT prevalence between males and females. There is no cogent reason why males should be at higher risk for DVT than females, and while some DVT incidence studies have also noted this risk difference (13,15), others have not (16,17).

Part of the explanation for the apparent sex difference in "risk" of DVT may be due to the following: studies have shown that in general, women are more likely than men to seek medical attention and, more often than men, visit physicians for chronic, non-fatal disease (84,85). In fact, of the 1034 patients with leg symptoms who were screened for the original study (IPG-CV study), 60% were female (because one of the exclusion criteria was pregnancy, the final CPOD study population had equal numbers of males and females), but only 18% of study women had DVT compared to 36% of study men. Perhaps it is not that male sex is a risk for DVT, but that female sex is "protective" much the same way that active disease was found to be "protective": if women tend to present

for DVT testing with leg symptoms that are more chronic and less severe than in men, it follows that DVT will be diagnosed less often in women than in men, and that male sex will thus be a predictor of DVT. There were no measures of symptom severity in this study, but the data on symptom duration show that women had longer mean symptom duration than men, which suggests that their symptoms were more chronic than in men (mean $\pm$ SE symptom duration (days) 23.8 $\pm$ 4.6 for women vs. 15.1 $\pm$ 5.2 for men, 95% CI for difference [-4.8, 22.2]).

Alternately, the apparent difference in risk between the sexes may have resulted from selection bias, which will be addressed in the section on the internal validity of this study. The external validity of this finding, ie. whether male sex is a predictor of DVT in other patient populations, needs to be studied.

Immobilization was associated with case status in the univariate analysis, but it was also correlated with surgery, orthopedic surgery, active cancer and trauma, and was not shown to be an independent predictor of DVT. This contrasts with autopsy and DVT screening studies that found an association between immobilization and DVT (30,31). Among the clinical prediction indices for DVT that used multivariate adjustment techniques, only one (76) found immobilization to be an independent predictor of DVT. Whether immobilization in itself is truly an etiologic risk factor for DVT has not been definitively established, however immobilization could still be a good predictor of DVT because of its association with other risk factors for DVT.

With regard to symptoms and signs, most had poor sensitivity and specificity, which is similar to the findings of previous studies reviewed in Chapter 2. No symptoms predicted case status, but several signs did: warmth, superficial venous dilation (SVD) and Homan's sign were detected significantly more often in cases than controls. For all three signs the negative predictive values were high (.83, .76, and .78 respectively) such that if these signs were absent, the patient was unlikely to have DVT. However, the corresponding positive predictive values were poor (.35, .42, .33 respectively): most patients who had these signs did not have DVT. The sign with the highest specificity was SVD (.85), which is similar to that seen in previous studies (12,63,65). In the multivariate analysis, only SVD and warmth were independently associated with case status.

For proximal DVT vs. controls, the results of the univariate analysis were similar, except that due to the exclusion of patients with calf DVT in whom immobilization and knee surgery were highly prevalent, these variables were no longer associated with case status.

5.1.2. Logistic regression analysis

In the logistic regression analysis, four variables were independently associated with DVT: male sex (adjusted OR 2.8), orthopedic surgery (adjusted OR 5.4), limb warmth (adjusted OR 2.1) and SVD (adjusted OR 2.9). Not surprisingly, because calf DVT patients were more likely than proximal DVT patients to be male and to have had orthopedic surgery, and were less likely to have warmth and SVD, the corresponding adjusted odds ratios for the secondary analysis, which excluded calf DVT patients, were slightly lower for male sex (2.4) and orthopedic surgery (4.1), and slightly higher for warmth (2.3) and SVD (3.4).

For both analyses, the 95% confidence intervals for the odds ratios were not unduly wide, suggesting that the estimated parameters were known with a reasonable degree of accuracy, and that the sample size must have been adequate for the analysis of these variables. However, this does not preclude that there may have been inadequate power to detect real differences between cases and controls for other measured predictor variables that did not appear in the final models.

Although the size of the effects were similar for both analyses, the 95% confidence intervals around the odds ratios were narrower for the proximal DVT analysis despite a smaller sample size, hence are likely to be known with a greater degree of accuracy. This suggests that for the variables in the final model, proximal DVT patients differ more from controls than DVT patients as a whole do. Although not shown in this study due to the small number of patients with calf DVT, one might infer that this is because calf DVT patients are more similar to control patients than are proximal DVT patients.

Effect modification between male sex and SVD was noted in the bivariate analysis and was substantively plausible. However, using non-informative priors, the posterior probability of the model that contained the interaction term SEX*SVD was three times

less likely than the model without the interaction term, and the BIC was less favorable. Since the BIC criterion was used for model selection because it is the criterion most likely to produce the best "out of sample" (ie. generalizable) model (80), this suggests that there was insufficient evidence in the data set alone to include the interaction term. This issue could be more definitively resolved in a second data set with a higher sample size.

As shown in the Results chapter, many of the models generated had approximately equal probability according to the BIC model selection criterion, and no one model showed strong or very strong evidence of being more likely than another (80). This can be a sign that no model predicts all that well, and may be due either to inadequate power to detect differences between groups for the variables measured, a true lack of differences between groups for these variables, or similar predictive power of several variables which, because of the limited sample size, could not all be included in the model. Due to the relatively small sample size of this study, it is certain that inadequate power was a factor, especially since there were a large number of competing variables in the initial variable matrix (1 variable per 10 subjects). With regard to whether or not there were true differences between cases and controls for the variables measured, the goodness of fit analyses showed that for covariate patterns with reasonably large cell size, the models predicted well for both DVT and proximal DVT.

The IPG-CV study was not originally designed to address the research questions of this thesis. As a consequence, many known risk factors or predictors of DVT were simply not measured, and were thus unavailable for consideration in the model selection process. Examples of these include measured difference in leg circumference, history of prior DVT, oral contraceptive use, pregnancy and postpartum, and inherited disorders of blood coagulation. It is possible that if these variables had been available for analysis, with a correspondingly larger sample size, the choice of the most likely model would have been more evident.

5.1.3. Clinical prediction index

Using the four predictors that were in the final logistic regression model, I developed a clinical prediction index to estimate the probability of DVT at any site in patients presenting with suspected DVT.

I did not develop an index for prediction of proximal DVT. One of my study goals was to identify unique predictors of proximal DVT, since proximal DVT carries a higher risk of PE and long-term disability than calf DVT (42). However, the final model of the secondary analysis (proximal DVT vs. controls) contained the same variables as that of the primary analysis (all DVT vs. controls). This indicates either that in this sample there were no differences in predictor variables between proximal DVT and DVT at any site, or that the study had inadequate power to detect these differences. In any case, patients do not arrive at the emergency room declaring the site of their DVT. Instead of comparing proximal DVT to controls, a more clinically relevant question that could have led to a discriminating index for site of DVT is: do predictors of calf DVT differ from predictors of proximal DVT, and do both differ from controls? As shown in the Results chapter, there were differences in many variables between calf DVT and proximal DVT patients. Unfortunately, I could not address this more relevant question using multivariate techniques because there were only 21 patients with calf DVT. I thus developed a clinical prediction index for predicting DVT at any site, since even the most discriminating index that differentiates proximal DVT patients from controls overlooks the fact that to get at patients with proximal DVT, patients with calf DVT must have already been identified and excluded.

Development of the logistic regression model to predict DVT was a necessary step to creating the clinical prediction index. However, the index is simply made up of combinations of the variables in the regression model without considering their weight or the degree of accuracy to which the parameters are known. As a result, information was lost in the transition from model to index. This is somewhat compensated for by the fact that an index is easier to use in the clinical setting than a logistic regression equation.

As shown in Table 4.34, there was sensitivity to change in the probability of DVT as the number of predictors variables increased: among patients with zero predictors, 9% had DVT, among patients with any one predictor, 20% had DVT, among patients with any two predictors, 32% had DVT, and in the group with any 3 or all predictors, 76% had DVT.

The predictors were then grouped in various combinations to create different cutoff points for the clinical prediction index, such that patients below the cutoff were at low risk for DVT, and patients above the cutoff were at high risk for DVT. A sensitivity analysis was done for these cutoff points.

Before discussing the test characteristics of the clinical prediction index, it is important to acknowledge that for most diagnostic tests, prediction indices included, the perfect cut-off point (ie. one that achieves 100% sensitivity and specificity) is unattainable, or, even if attainable under study conditions, not achievable in the "real world". This is because the disease under study can manifest differently in different people (i.e. human variability in expression of disease due to genetic, environmental, social or cultural factors) and tests may not always perform as anticipated (i.e. lab error or human error in applying or interpreting a test or index). Hence, there will almost always be some tradeoff between sensitivity and specificity, and it is up to the clinician to evaluate how much of one can be sacrificed for the other. For some diseases, the choice is obvious: e.g., for a clinical index designed to predict the probability of breast cancer and the subsequent need for biopsy in a woman presenting with a breast lump, the most important requirement would be high sensitivity, in order to avoid a false negative diagnosis in a woman who truly has breast cancer. The anxiety associated with false positive labeling of a woman without breast cancer, although undesirable, becomes secondary.

For the diagnosis of DVT, the choice is less clear. A false negative diagnosis is likely to be more adverse than a false positive diagnosis, because untreated DVT can lead to fatal pulmonary embolism and the post-phlebitic syndrome. The risks associated with anticoagulating a patient who does not have a DVT are smaller, but, rarely, can be catastrophic. As Wheeler points out, there are few studies on the natural history of untreated DVT (50). In many of the studies which documented post-operative DVT using

^{125}-I fibrinogen leg scan, patients with positive test results were not treated and did not suffer adverse outcomes, probably because these small thrombi lysed in situ or failed to propagate. In the few patients in whom PE was documented, it usually was not life-threatening (22). Although IPG can fail to detect calf thrombi, non-occlusive proximal thrombi, and older occlusive thrombi with well-developed collateral circulation, prospective studies have shown a very low rate of complications and no fatal PE in patients with suspected DVT in whom treatment was withheld because of negative IPG (3,86,87) or CUS (46).

It is unlikely that more epidemiological information will become available on the natural history of untreated DVT. Therefore, the most conservative approach is to acknowledge that DVT has a wide spectrum of severity and outcomes, but that as yet it cannot be predicted where an individual patient lies on this spectrum. It is safe to say that low sensitivity is almost certainly more dangerous than low specificity, since the potential problems associated with low specificity relate more to the costs associated with unnecessary treatment and to the adverse impact of false labeling on quality of life than to an important risk of adverse patient outcome.

In the clinical prediction index, for one or more vs. zero predictors, the estimated sensitivity was .93. For two or more vs. one or less predictors, the estimated sensitivity was .66. For three or more vs. two or less predictors, the estimated sensitivity was .26. The corresponding specificities were .25, .66, and .97. In simple terms, as the cutoff for a positive "test" moved from 3 or more variables to one or more variables (ie. as one required that fewer variables be positive to consider the result of the index positive), the chance of falsely classifying a case as a control decreased, but the chance of falsely classifying a control as a case increased.

The ROC curve depicts the relationship between sensitivity and 1-specificity for the different cutoff points of the predictive index. The magnitude of the area between the diagonal and the ROC curve shows that the clinical index, overall, predicts considerably better than pure chance. A perfect cutoff point, ie. one that is able to differentiate all cases from all controls, would have its data point located in the top left corner of the graph.

None of the three cutoff points provides the ideal balance between high sensitivity and reasonably preserved specificity. Although for the reasons discussed above it is tempting to choose the cutoff point with the highest sensitivity (one or more vs. zero predictors), the PPV for this cutoff is only 31%, which is prediction that is no better than chance, since the pre-test probability of DVT based simply on the known prevalence of DVT in symptomatic patients in general is 25-30%, and was 27% in this sample. Overall, the cutoff point of more than three vs. two or less predictors had the best balance of good PPV and NPV (76% and 78% respectively) and the best LR+ (8.7). However, only 19 cases were in this risk category, leaving 54 cases who fell below the cutoff point and who would have been classified as low risk for DVT. The misclassification rate was 57% for the cutoff point more than one vs. zero predictors, 34% for more than two vs. one or less predictor, and 22% for more than 3 vs. two or less predictors.

In clinical practice it is the predictive value of a diagnostic test that is most important in treatment decisions, and the predictive value is greatly influenced by the probability of disease. This can influence test selection and management in the following way: in patients with a low probability of DVT, a negative non-invasive test is adequate to rule out DVT, based on published estimated sensitivities and specificities of such tests. Similarly in patients with a high probability of DVT, a positive non-invasive test is adequate to rule in DVT. When the test result is discordant with the estimate of pre-test probability, a different non-invasive test or contrast venography (CV) should be performed, due to the high false positive rate in patients with low probability of DVT and the high false negative rate in patients with high probability of DVT. In patients with a moderate pre-test probability of DVT, since the PPV is poor, a positive test should be confirmed with a different non-invasive test or venography, but a negative non-invasive test, if known to have high local sensitivity, is probably enough to rule out DVT, especially if serial testing is performed (50).

Applying this strategy to my sample, I classified 0 predictors as representing "low probability" for DVT (n=54; 20% of study population; 5 had DVT), 1 or 2 predictors as "moderate probability" for DVT (n=192; 71% of study population; 49 had DVT) and 3 or more predictors as "high probability" for DVT (n=25; 9% of study population; 19 had

DVT). Non-invasive tests would likely have sufficed and CV avoided in 49 low probability patients, 143 moderate probability patients and 19 high probability patients ie. a total of 211 (78%) of the 271 study patients. CV would still have been be required in the 5 low probability patients who had a positive non-invasive test, the 49 moderate probability patients with a positive non-invasive test, and the 6 high probability patients with a negative non-invasive test ie. a total of 60 (22%) of the 271 study patients.

Hence, although the clinical prediction index did not perform well enough to be used as a diagnostic "test" for DVT, it could prove to be useful for the selection of the most appropriate diagnostic test and its interpretation in a similar study population with comparable distributions of predictors. Ultimately, the performance of this index should be tested prospectively in a larger data set.

Two other clinical prediction indices have been developed for use in patients with suspected DVT, as discussed in the Literature Review chapter. In the first, using a sample size similar to my study, swelling below the knee, swelling above the knee, recent immobility, cancer and fever were independent predictors of proximal DVT (76). Data collection was via retrospective chart review. This index performed poorly compared to mine: although the area under the ROC curve was similar, the likelihood ratios and PPVs of the various cutoff points were more modest, and patients with two predictors had a higher proportion of DVT than patients with 3-5 predictors, which suggests an overall lack of internal validity of the study.

The second index, developed *a priori* by Wells based on a literature review of DVT risk factors and clinical expertise, was applied to 529 outpatients with first episode of suspected DVT (77). The PPVs for the high, moderate and low risk groups were .85, .33, and .05 respectively, which mirror almost exactly the findings in my study for patients with 3 predictors, 1 or 2 predictors, and 0 predictors (PPV .76, .34, and .07 respectively). However, a greater proportion of Wells' study patients than mine were in the low or high pretest probability categories.

This study had several limitations which will be discussed under the headings of internal validity and external validity.

5.2.1. Internal validity

Internal validity refers to the validity of the analytic inferences as they pertain to the actual subjects in the study (88). Selection bias, misclassification bias, and confounding may have impacted on the internal validity of this study.

Selection bias

This study was not population-based. During the recruitment of the study population, there were undoubtedly selection pressures in effect, some of which can be surmised but most of which remain unknown. Although all patients who were referred for suspected DVT during the study period were eligible for the study, the actual process of referral, i.e. how patients came to be referred in the first place, is not known. For example, it is not known by whom and from where patients were referred, and whether physicians referring patients generally had a high or low threshold for referral.

Because the Montreal General Hospital (MGH) is a tertiary care hospital, patients selected into the study had a high rate of recent surgery. Having had recent surgery can have an impact both on the way DVT manifests and on a physician's threshold for deciding that a clinical finding might be due to DVT (detection bias). During the time the IPG-CV study was recruiting patients, the principal investigator was considered to be a city-wide expert in venous thromboembolic disease. It is possible that unusual, complex cases were preferentially selected into the study simply because more of these patients were referred to the MGH.

Patients excluded from the study were older than those included (mean age 60.3 years vs. 56.7 years). Overall, the majority of exclusions were for refusal or failure to give informed consent, which can be a marker of severe illness or dementia. Thus, it is likely

that mostly "healthy" elderly were selected into the study, which might explain why the expected association between age and DVT was not found.

In this study, male sex was found to be an independent predictor of DVT. In women, the postpartum period carries a high risk for DVT (estimated risk 2/1000 (38)). The MGH has no obstetrics department, so postpartum patients were extremely unlikely to be referred for the IPG-CV study. By extrapolation from comparably-sized Montreal hospitals that handle about 4000 deliveries per a year, over the 2 year IPG-CV study period, up to 16 additional women with DVTs might have entered the study, which could have attenuated or eliminated the estimated risk difference between men and women. The risk difference between men and women could also have resulted from selecting into the study men who were at higher than average risk for DVT; for example, hypothetically, if all men with sports injuries who required knee surgery were self-referred to one sports surgeon who had a high post-operative DVT rate, it would appear that male sex was a predictor of DVT.

Misclassification bias

Great care was taken to correctly classify cases and controls, hence misclassification of outcome, although possible, was unlikely. However, since the study was not originally designed to analyze predictors of DVT, misclassification of "exposure" (ie. presence or absence of predictor variables) could have occurred if the nurses collecting baseline information were not adequately diligent in extracting and recording information. This misclassification, if it occurred, is likely to have been non-differential, since the diagnostic test for DVT was performed only after the baseline data was collected.

In general, the predictor variable measures were crude, most requiring simple yes/no answers, hence errors of documentation could have occurred. For the symptoms and physical exam data, information could only be recorded as 'present' or 'absent', i.e. there were no severity ratings. Also, the reliability of the examination technique was not measured. Finally, information on other potential predictor variables was simply never collected (for example, measured leg circumference was listed on the original data

collection sheet but was missing for most patients, the question on surgery in the last six months did not allow for documentation of time since surgery, etc.).

Confounding bias

Although the stratified analysis did reveal confounding that was adjusted for in the multivariate model, both residual confounding and undetected confounding by unmeasured variables might have occurred.

5.2.2. External validity

External validity refers to the validity of the analytic inferences as they pertain to people outside the study population (88).

The generalizability of this study is influenced by the following points:

- The study was conducted at a single tertiary care hospital that has no obstetric department.
- Only symptomatic patients with a first episode of suspected DVT were eligible.
- Pregnant women were excluded.
- Compared to other studies of DVT risk factors in patients with leg symptoms (76,77), this study population had a high rate of recent surgery (33% overall) and leg trauma (23% overall).
- The prevalence of DVT in this sample (27%) was similar to that seen in other studies of symptomatic patients, however the ratio of proximal DVT to calf DVT (2.5 to 1) was lower than in other studies (76,77).

Since the prevalence of a disease has a major impact on the predictive accuracy of a test used to diagnose the disease, the results of the this study could only be generalized to similar populations with similar disease prevalence.

In summary, assuming adequate internal validity, the results of this study could likely be generalized to symptomatic patients with a first episode of suspected DVT

coming from a population with a similar DVT prevalence. However, the external validity of this study could best be assessed by prospectively applying the regression model and the clinical prediction index to a large population of patients with a wide clinical spectrum of DVT.

5.3. Conclusions

In this study I demonstrated that male sex, orthopedic surgery, warmth and superficial venous dilation on exam were independent predictors of symptomatic DVT and symptomatic proximal DVT. Using various groupings of these predictors, I developed a clinical prediction index for symptomatic DVT.

My clinical prediction index meets most of the suggested methodological standards for clinical prediction rules (67,70). First, it uses predictors that are clinically relevant and whose presence or absence are easy to determine. The outcome to be predicted, DVT, was clearly defined, using an accepted reference test. Assessment of outcome was prospective and was blinded to the information collected at baseline. The demographic characteristics of the study population and the study site were described, as were the mathematical techniques used to develop the index. Finally, a measure of the misclassification rate was provided.

While the index's ROC curve showed moderate overall prediction, there was no single cut-off point that gave the best balance of desired sensitivity and specificity. The index showed adequate predictive accuracy and high LR+ for patients with 3 or more predictors (ie. high-risk patients), however these patients made up only a small proportion of the study population. Nevertheless, by grouping patients into low, moderate and high probability for DVT, the index was useful in a strategy aimed to limit the need for CV to diagnose DVT, in favor of less invasive tests.

The external validity of this index has not been assessed. Ultimately, the most important criterion of usefulness of a clinical prediction index is evidence that applying the

index to its target population results in a positive impact on measures of health.

This was not an etiological study. It did not seek to identify causative factors ("risk factors") for DVT, but rather sought to identify predictors of DVT that might be useful to the clinician faced with a patient in whom DVT is suspected. In the end, the combination of variables that best predicted DVT in this study was a mix of possible causative factors (male sex, orthopedic surgery) and effects or "markers" of DVT (warmth, superficial venous dilation).

As discussed in the Methods chapter, because of the limitations imposed by the design of the original IPG-CV study, the selection of cases and controls in this study would have been inadequate for a true case-control study seeking to explore DVT causality.

With regard to DVT causality, at present there is limited understanding of this disease. To borrow from cancer causality terminology, deep vein thrombosis is a disease with many promoters but no clear initiators: numerous risk factors have been identified and can be measured, but probably many more, and perhaps the most salient ones, have not been measured, because they are as yet unknown to us, and even if they become known to us, may not be readily measurable. Deep vein thrombosis may be more than one disease, given the wide range of patients it affects and venous sites it inhabits. Until more is understood about the pathophysiology of DVT and why its clinical spectrum is so different in different patients, identifying true etiologic risk factors for DVT might not be readily achievable.

Nonetheless, on a daily basis, clinicians are put into the position of having to make diagnostic and therapeutic decisions for patients presenting with complex groupings of symptoms, signs and baseline characteristics. As such, clinical prediction rules that have proven predictive accuracy are important tools for the clinician, regardless of whether the predictors are truly etiologic, are confounders of the risk factors they represent, or are markers of the disease itself.

6. Bibliography

1. Kakkar VV, Howe CT, Flanc C, et al. Natural history of post-operative deep vein thrombosis. Lancet 1969;(2):230-3.

2. Prevention of fatal post-operative pulmonary embolism by low doses of heparin. An international multicentre trial. Lancet 1975;2:45-51.

3. Hull RD, Hirsh J, Carter CJ, et al. Diagnostic efficacy of impedance plethysmography for clinically suspected deep vein thrombosis: a randomized trial. Ann Intern Med 1985;102:21-8.

4. Huisman MV, Buller HR, ten Cate JW, et al. Unexpected high prevalence of silent pulmonary embolism in patients with deep vein thrombosis. Chest 1989;95:498-502.

5. Moser KM, LeMoine JR. Is embolic risk conditioned by location of deep venous thrombosis? Ann Intern Med 1981;94:439-4.

6. Barritt DW, Jordan SC, Brist MB. Anticoagulant drugs in the treatment of pulmonary embolism. A controlled trial. Lancet 1960;18:1309-12.

7. Salzman EW, Deykin D, Shapiro RM, Rosenberg R. Management of heparin therapy. A controlled prospective trial. N Engl J Med 1975;292:1046-50.

8. Hull R, Delmore TJ, Genton E, et al. Warfarin sodium versus low dose heparin in the long-term treatment of venous thrombosis. N Engl J Med 1979;301:855-8.

9. Levine MN, Raskob G, Landefeld S, Hirsh J. Hemorrhagic complications of anticoagulant treatment. Chest 1995;108(supplement):276S-90S.

10. David M. Venous thromboembolism during pregnancy and postpartum. in Leclerc JR. Venous thromboembolic disorders. Philadelphia: Lea & Febiger; 1991; p. 346-65.

11. Leclerc JR, Illescas F, Jarzem P. Diagnosis of deep vein thrombosis. in Leclerc JR. Venous thromboembolic disorders. Philadelphia: Lea & Febiger; 1991; p. 176-228.

12. Haeger K. Problems of acute deep venous thrombosis. I. The interpretation of signs and symptoms. Angiology 1969;20:219-3.

13. Coon WW, Willis PW, Keller JB. Venous thromboembolism and other venous diseases in the Tecumseh Community Health Study. Circulation 1973;48:839-46.

14. Gillum RF. Pulmonary embolism and thrombophlebitis in the United States, 1970-1985. Am Heart J 1987;114:1262-4.

15. Anderson FA, Wheeler HB, Goldberg RJ, et al. A population-based perspective of the hospital incidence and case-fatality rates of deep-vein thrombosis and pulmonary embolism. The Worcester DVT study. Arch Intern Med 1991;151:933-8.

16. Nordstrom M, Lindblad B, Bergqvist D, Kjellstrom T. A prospective study of the incidence of deep-vein thrombosis within a defined urban population. J Intern Med 1992;232:155-60.

17. Kniffin WD, Baron JA, Barrett J, et al. The epidemiology of diagnosed pulmonary embolism and deep venous thrombosis in the elderly. Arch Intern Med 1994;154:861-6.

18. Hull R, Hirsh J, Sackett DL, Stoddart G. Cost effectiveness of clinical diagnosis, venography, and noninvasive testing in patients with symptomatic deep-vein thrombosis. N Engl J Med 1981;304:1561-7.

19. Strandness DE, Langlois Y, Cramer M, et al. Long-term sequelae of acute venous thrombosis. JAMA 1983;250:1289-92.

20. Clagett GP, Anderson FA, Jr., Heit J, et al. Prevention of venous thromboembolism. Chest 1995;108:312S-34S.

21. Geerts WH, Code KI, Jay RM, et al. A prospective study of venous thromboembolism after major trauma. N Engl J Med 1994;331:1601-6.

22. Collins RB, Scrimgeour A, Yusuf S, Peto R. Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin. N Engl J Med 1988;318:1162-73.

23. Belch JJF, Lowe GDO, Pollock JG, et al. Low dose heparin in the prevention of deep-vein thrombosis after aortic bifurcation graft surgery. Thromb Hemost 1979;42:1429-33.

24. Cass AJ, Jennings A, Greenhalgh RM. Leg swelling after aortic surgery. International Angiology 1986;5:207-8.

25. Clarke-Pearson DL, DeLong ER, Synan IS, et al. Variables associated with post-operative deep venous thrombosis: a prospective study of 411 gynecology patients and creation of a prognostic model. Obstet Gynecol 1987;69:146-50.

26. McKenzie PJ. Effect of anaesthetic technique on deep vein thrombosis: A comparison of subarachnoid and general anaesthesia. Br J Anaesth 1985;57:853-7.

27. Poikolainen E, Hendolin H. Effects of lumbar epidural analgesia and general anaesthesia on flow velocity in the femoral vein and post-operative deep vein thrombosis. Acta Chir Scand 1983;149:361-4.

28. Levine MN, Gent M, Hirsh J et al. The thrombogenic effect of anticancer drug therapy in women with stage II breast cancer. N Engl J Med 1988;318:404-7.

29. Hirsch DR, Ingenito EP, Goldhaber SZ. Prevalence of deep venous thrombosis among patients in medical intensive care. JAMA 1995;274:335-7.

30. Gibbs NM. Venous thrombosis of the lower limbs with particular reference to bedrest. Br J Surg 1957;45:209-36.

31. Kierkegaard A, Norgren L, Olsson C-G, et al. Incidence of deep vein thrombosis in bedridden non-surgical patients. Acta Med Scand 1987;222:409-14.

32. Kierkegaard A. Incidence and diagnosis of deep vein thrombosis associated with pregnancy. Acta Obstet Gynecol Scand 1983;62:239-43.

33. Ginsberg JS, Brill-Edwards P, Burrows RF, et al. Venous thrombosis during pregnancy: Leg and trimester of presentation. Thromb Hemost 1992;67:519-20.

34. Thromboembolic Risk Factors (THRIFT) Consensus Group. Risk of and prophylaxis for venous thromboembolism in hospital patients. Br Med J 1992;305:567-74.

35. Realini JP, Goldzieher JW. Oral contraceptives and cardiovascular disease: A critique of the epidemiologic studies. Am J Obstet Gynecol 1985;152:729-98.

36. Porter JB, Hunter JR, Jick H, Stergachis A. Oral contraceptives and non-fatal vascular disease. Obstet Gynecol 1985;66:1-4.

37. Spitzer WO, Lewis MA, Heinemann LA, et al. Third generation oral contraceptives and risk of venous thromboembolic disorders: an international case-control study. Br Med J 1996;312:83-8.

38. Paltiel O. Epidemiology of venous thromboembolism. in Leclerc JR. Venous thromboembolic disorders. Philadelphia: Lea & Febiger; 1991; p. 141-65.

39. Dahlback B. Inherited thrombophilia: resistance to activated protein C as a basis for venous thrombosis. Blood 1995;85:607-14.

40. Salzman EW, Hirsh J. Prevention of venous thromboembolism. in Colman RW, Hirsh J, Marder VJ, et al.(editors).Hemostasis and thrombosis: basic principles and clinical practice. 3rd ed. Philadelphia: JB Lippincott; 1994; p. 1332

41. Sandler DA, Martin JF. Autopsy proven pulmonary embolism in hospital patients: are we detecting enough deep vein thrombosis? J Royal Soc Med 1989;82:203-5.

42. Philbrick JT, Becker DM. Calf deep venous thrombosis. A wolf in sheep's clothing? Arch Intern Med 1988;148:2131-8.

43. Baker WF, Jr., Bick RL. Deep vein thrombosis. Diagnosis and management. Med Clin N Am 1994;78:685-712.

44. Comerota AJ, Katz ML, Grossi RJ, et al. The comparative value of noninvasive testing for diagnosis and surveillance of deep vein thrombosis. J Vasc Surg 1988;7:40-9.

45. Heijboer H, Cogo A, Buller HR, et al. Detection of deep vein thrombosis with impedance plethysmography and real-time compression ultrasonography in hospitalized patients. Arch Intern Med 1992;152:1901-4.

46. Heijboer H, Buller HR, Lensing AWA, et al. A comparison of real-time compression ultrasonography with impedance plethysmography for the diagnosis of deep-vein thrombosis in symptomatic outpatients. N Engl J Med 1993;329:1365-9.

47. Lensing AWA, Buller HR, Prandoni P, et al. Contrast venography, the gold standard for the diagnosis of deep vein thrombosis: Improvement in observer agreement. Thromb Hemost 1992;67:8-12.

48. Hull R, Hirsh J, Sackett DL, et al. Clinical validity of a negative venogram in patients with clinically suspected venous thrombosis. Circulation 1981;64:622-5.

49. Lensing A, Prandoni P, Buller HR, et al. Lower extremity venography with iohexol: results and complications. Radiology 1990;177:503-5.

50. Wheeler HB, Hirsh J, Wells P, Anderson FA, Jr. Diagnostic tests for deep vein thrombosis. Clinical usefulness depends on probability of disease. Arch Intern Med 1994;154:1921-8.

51. Koopman MM, van Beek EJ, ten Cate JW. Diagnosis of deep vein thrombosis. Prog Cardiovasc Dis 1994;37:1-12.

52. Agnelli G, Cosmi B, Ranucci V, et al. Impedance plethysmography in the diagnosis of asymptomatic deep vein thrombosis in hip surgery. A venography-controlled study. Arch Intern Med 1991;151:2167-71.

53. Ginsberg JS, Wells PS, Hirsh J, et al. Reevaluation of the sensitivity of impedance plethysmography for the detection of proximal deep vein thrombosis. Arch Intern Med 1994;154:1930-33.

54. Anderson DR, Lensing AW, Wells PS, et al. Limitations of impedance plethysmography in the diagnosis of clinically suspected deep-vein thrombosis. Ann Intern Med 1993;118:25-30.

55. Cogo A, Prandoni P, Bernardi E, et al. Distribution of proximal vein thrombi in symptomatic patients: diagnostic implications. [Abstract] Thromb Hemost 1993;69:619

56. Lensing AWA, Prandoni P, Brandjes DPM, et al. Detection of deep vein thrombosis by real-time B-mode ultrasonography. N Engl J Med 1989;320:342-5.

57. Hull R, Hirsh J, Sackett DL, et al. Combined use of leg scanning and impedance plethysmography in suspected venous thrombosis. N Engl J Med 1977;296:1497-500.

58. Johnson WC. Evaluation of newer techniques for the diagnosis of venous thrombosis. J Surg Res 1974;16:473-81.

59. Cranley JJ, Canos AJ, Sull WJ. The diagnosis of deep venous thrombosis. Fallibility of clinical symptoms and signs. Arch Surg 1976;111:34-6.

60. Singer J. Value of clinical signs in the diagnosis of deep vein thrombosis. Lancet 1980;1186

61. Vaccaro P, Van Aman M, Miller S, et al. Shortcomings of physical examination and impedance plethysmography in the diagnosis of lower extremity deep venous thrombosis. Angiology 1987;38:232-5.

62. Cooperman M, Martin EW, Satiani B, et al. Detection of deep vein thrombosis by impedance plethysmography. Am J Surg 1979;137:252-4.

63. Lindqvist R. Ultrasound as a complementary diagnostic method in deep vein thrombosis of the leg. Acta Med Scand 1977;201:435-8.

64. Simpson FG, Robinson PJ, Bark M, Losowsky MS. Prospective study of thrombophlebitis and "pseudothrombophlebitis". Lancet 1980;1:331-33.

65. Richards KL, Armstrong JJD, Tikoff G, et al. Noninvasive diagnosis of deep vein thrombosis. Arch Intern Med 1976;136:1091-6.

66. Sandler DA, Martin JF, Duncan JS, et al. Diagnosis of deep-vein thrombosis: comparison of clinical evaluation, ultrasound, plethysmography, and venoscan with X-ray venogram. Lancet 1984;2:716-9.

67. Braitman LE, Davidoff F. Predicting clinical states in individual patients. Ann Intern Med 1996;125:406-12.

68. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974;185:1124-31.

69. Sox HC. Probability theory in the use of diagnostic tests. Ann Intern Med 1986; 104:60-6.

70. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. N Engl J Med 1985;313:793-9.

71. Crandon AJ, Peel KR, Anderson JA, et al. Postoperative deep vein thrombosis: identifying high-risk patients. Br Med J 1980;281:343-4.

72. Lowe GDO, McArdle BM, Carter DC, et al. Prediction and selective prophylaxis of venous thrombosis in elective gastrointestinal surgery. Lancet 1982;i:409-12.

73. Mellbring G, Dahlgren S. Prediction of postoperative venous thrombosis. Lancet 1983;1:1382

74. Sue-Ling HM, Johnston D, McMahon MJ, et al. Pre-operative identification of patients at high risk of deep venous thrombosis after elective major abdominal surgery. Lancet 1986;1:1173-6.

75. Vine HS, Hillman B, Hessel SJ. Deep vein thrombosis: predictive value of signs and symptoms. A J Roentgen 1981;136:167-71.

76. Landefeld CS, McGuire E, Cohen AM. Clinical findings associated with acute proximal deep vein thrombosis: a basis for quantifying clinical judgment. Am J Med 1990;88:382-8.

77. Wells PS, Hirsh J, Anderson DR, et al. Accuracy of clinical assessment of deep-vein thrombosis. Lancet 1995;345:1326-30.

78. Hull R, van Aken WG, Hirsh J, et al. Impedance plethysmography using the occlusive cuff technique in the diagnosis of venous thrombosis. Circulation 1976;53:696-700.

79. Kleinbaum DG; Kupper LL; Muller KE. Applied regression analysis and other multivariable methods. Belmont, California: Duxbury Press; 1988.

80. Kass RE and Raftery AE. Bayes Factors. J Am Stat Assoc 1995;90(430):773-95.

81. Hosmer DW; Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.

82. Sackett DL; Haynes RB; Tugwell P. Clinical epidemiology. A basic science for clinical medicine. Boston: Little, Brown and Company; 1985.

83. Thisted R. Elements of statistical computing. New York: Chapman and Hall; 1988.

84. van der Waals FW. Differences in leading causes of death, hospitalization and general practice visits among Dutch women and men. Women & Health 1991;17(3):101-23.

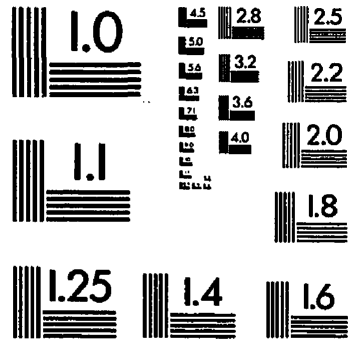
85. Briscoe ME. Why do people go to the doctor? Sex differences in the correlates of GP consultation. Soc Sci & Med 1987;25(5):507-13.

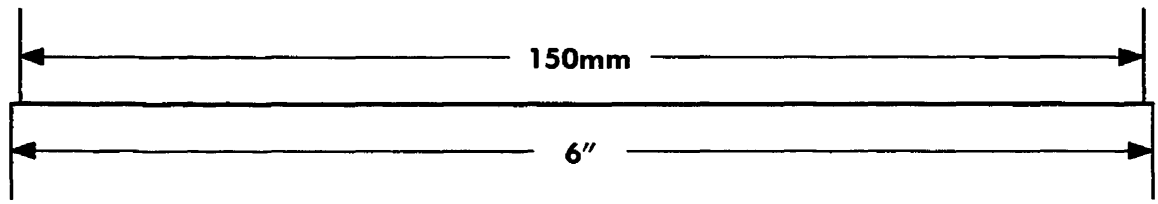
86. Huisman MV, Buller HR, ten Cate JW, Vreeken J. Serial impedance plethysmography for suspected deep venous thrombosis in outpatients: the Amsterdam General Practitioner Study. N Engl J Med 1986;314:823-8.

87. Huisman MV, Buller HR, ten Cate JW, et al. Management of clinically suspected acute venous thrombosis in outpatients with serial impedance plethysmography in a community hospital setting. Arch Intern Med 1989;149:511-3.

88. Rothman KJ. Modern Epidemiology. Boston, Little, Brown and Co.; 1986. p. 82.

TEST TARGET (QA-3)





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989