von Willebrand Disease and Exercise - Is it Worth the Sweat?

The Effect of Intense Physical Exercise on von Willebrand Factor and on Menstrual Blood Loss in Women with von Willebrand Disease

Rochelle Winikoff
Department of Epidemiology and Biostatistics
McGill University, Montreal
March 2003

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

TABLE OF CONTENTS

ABSTRACT:

The principal objective of this study was to examine the effect of intense physical exercise on menstrual blood loss in women with Type I von Willebrand disease (vWD). First, we investigated the effect of exercise on the level of the von Willebrand protein (which is deficient in the disease) in a pre-test post-test quasi-experiment conducted on a single group of 40 healthy adult pre-menopausal female volunteers recruited from Sainte-Justine Hospital in Montreal between October and December, 2001. The von Willebrand protein (vWF:Ag), coagulation Factor VIII (FVIII:C), bleeding time (BT), coagulation time (aPTT) and several markers of exercise intensity (sweat sodium, lactate, noradrenaline, adrenaline) were measured before and after a standardized exercise session. The significance of the change in these values with exercise was assessed using a paired Student's t-test. The exercise markers were explored as potential predictors of the exercise-related change in vWF:Ag using multiple linear regression. Results showed that there was an absolute mean increase of 0.30 (95% confidence interval (95% CI) 0.23-0.37) and 0.60 (95% CI 0.44-0.76) in vWF:Ag and FVIII:C, respectively, and a significant shortening of the BT and aPTT due to exercise. The change in the sweat sodium collected from patches applied to the forearm during exercise (a marker of exercise intensity) was found to be a significant predictor of the change in vWF:Ag induced by exercise (regression coefficient = 0.05 (95% CI 0.01-0.09). Changes of 1, 5 and 10 units in sodium were associated with average changes of 0.05, 0.26 and 0.52, respectively, in vWF:Ag from baseline (mean 0.83 U/ml). Next, we set out to assess the feasibility and acceptability of a 4-period randomized crossover trial in order to evaluate the effectiveness of exercise in reducing the menstrual blood flow in women with Type I vWD. The methods and protocol of this feasibility study are outlined in this thesis and issues related to patient recruitment, compliance and withdrawals are addressed. The strengths and pitfalls of the crossover design feasibility study are discussed and revisions for the definitive trial are recommended.

RÉSUMÉ

L'objectif principal de cette étude était d'éxaminer l'effet de l'exercise physique intense sur les pertes menstruelles abondantes chez les femmes atteintes de la maladie von Willebrand de Type I. Dans la première étude, nous avons investigué l'effet de l'exercice sur le niveau sanguin de la protéine von Willebrand (celle qui est déficiente dans cette maladie) par le biais d'une étude quasi-expérimentale de type pré-test post-test réalisée auprès d'un seul groupe des femmes bénévoles en bonne santé et préménopausées. Elles furent recrutées à l'Hôpital Sainte-Justine de Montréal entre octobre et décembre 2001. La protéine von Willebrand (vWF:Ag), le facteur VIII de la coagulation (FVIII:C), le temps de saignement (TS), le temps de coagulation (aPTT), ainsi que quelques indicateurs de l'intensité de l'exercice ont été mesurés avant et après un exercice standardisé. Un test t pairé de Student a été utilisé pour évaluer la signification des changements dans ces valeurs suite à l'exercice. Des indicateurs de l'intensité de l'exercice physique ont étés investigués comme variables prédictives potentielles du changement dans le vWF:Ag induit par l'exercice. Une analyse de régression linéaire multiple a été utilisée. Les résultats de cette étude démontrent une augmentation moyenne absolue de 0.30 (intervalle de confiance (IC) de 95% 0.23-0.37) et 0.60 (IC 95% 0.44-0.76) dans les facteurs de vWF:Ag et FVIII:C, respectivement, et un raccourcissement significatif du TS et de l'aPTT suite à un exercice. L'augmentation de niveaux de chlorure de sodium dans la transpiration produite lors de l'exercice fut demontrée comme étant de facteur prédictif significatif du changement dans le vWF:Ag (coefficient de régression = 0.05 IC 95% 0.01-0.09). Des changements de 1, 5 et 10 unités de chlorure de sodium sont associés avec des changements moyens respectifs de 0.05, 0.26 et 0.52 unités dans le niveau de base du vWF:Ag (moyenne 0.83 U/ml). Par la suite, nous avons examiné la faisabilité et l'acceptabilité d'une étude "crossover" aléatoire de quatre périodes, pour évaluer l'impact clinique de l'exercise sur les pertes menstruelles excessives chez les femmes atteintes de la maladie von Willebrand de Type I. La méthodologie et le protocole de cette étude pilote sont décrits dans la présente thèse et les questionnements reliés au recrutement, à l'assiduité et aux abandons sont discutés. Les forces et faiblesses de cette étude sont discutées et des recommendations sont formulées pour la révision de l'étude définitive.

ACKNOWLEDGMENTS:

I am especially grateful to my thesis supervisor, Dr. Claire Infante-Rivard, for her guidance and support during the preparation of this thesis. I would like to thank her for her valuable, constructive and timely comments as well as her reassurance at times when it was much needed. I would like to thank Dr. Stan Shapiro, a member of my thesis subcommittee, for his kind and helpful advice. I would also like to give special thanks to Igor Karp for his generous help, his constant availability and his incredible patience.

I would also like to thank Dr. Georges-Etienne Rivard, my clinical fellowship mentor in hemostasis at Sainte-Justine Hospital, without whom the realization of these two research projects would not have been possible. His dedication to and enthusiasm in clinical medicine and research have been a major inspiration to me and he has been and continues to be a role model for my personal achievements. I thank him as well for his constant support in my clinical research endeavors and in my pursuit of my Masters degree in Epidemiology and Biostatistics.

Finally, I owe endless thanks to my parents (Helen and Robert Winikoff) and to my sister (Lyanne) for their unconditional support and encouragement during my training and the preparation of this thesis, and to my Grandmother (Fay Randolph) and my dearest friends (you know who you are) for their faith in me and my ability to complete this tremendous task.

DEDICATION:

I dedicate this thesis to my late grandmother, Etta Moscovitch, who continues to inspire me even in her absence. I am indebted to her for her wisdom in all aspects of life and for teaching me that hard work is the key to success. She has also taught me how to be a well-balanced person, to respect others, to exercise sound judgement, to set high standards for myself and to perceive no limits as to what I can accomplish. I need not justify the importance of this thesis to her because the pursuit of higher knowledge is the crux of her legacy. I have no doubt that she is proud.

INTRODUCTION:

Menorrhagia (excessive menstrual blood loss) is the most prevalent symptom in women with von Willebrand disease (vWD) and has a major negative impact on the quality of life of those affected. All current treatment options to curtail menstrual blood loss in women with this condition require either the ingestion or injection of a medication or a surgical intervention. Natural approaches to treat menorrhagia in these women are virtually absent from current medical practice. Such approaches, if proven effective, could be especially appealing to some women with this condition and could also circumvent the potential morbidity, undesirable side effects and costs associated with current treatments.

The published literature suggests that intense physical exercise can bring about many changes in the coagulation system. However, the precise effect that exercise has on von Willebrand factor protein (vWF:Ag), which is typically deficient in individuals with vWD, has been the focus of only a few small uncontrolled studies in men. In this study, we set out to determine to what extent intense physical exercise could reduce menstrual blood loss in women with menorrhagia related to vWD. First, we assessed the change in vWF:Ag induced by exercise in normal pre-menopausal women volunteers recruited from Sainte-Justine Hospital, in a one-group pre-test post-test quasi-experiment. The exercise protocol from the first part of the study was then used as the intervention in a 4-period crossover feasibility study where we examined the effects of exercise on menorrhagia in women with Type I vWD identified from our clinic database. The analyses presented in this thesis are based on the data that was collected by our investigating team during the study.

BACKGROUND:

i. Burden of Illness

vWD is the most common inherited bleeding disorder that affects approximately 1% of the general population, or 1 in every 100 persons, when random screening is carried out.[1] Based on this estimated prevalence, it is expected that approximately 300,000 Canadians are affected by this bleeding disorder. vWD is a genetic disorder that it transmitted in an autosomal dominant fashion and, therefore, it affects men and women equally.[2]

ii. The Role of von Willebrand Protein (vWF:Ag)

vWF:Ag, a circulating blood protein, has two major functions in hemostasis (the arrest of bleeding). First, it mediates platelet adhesion and the formation of the platelet plug at the site of vascular injury (primary hemostasis), and in so doing, permits the initial arrest of bleeding. Second, vWF:Ag carries and protects the coagulation protein Factor VIII (FVIII:C), which is essential to the formation of a more definitive blood clot at the site of injury (secondary hemostasis).[3] vWD is characterized by defective or deficient blood levels of vWF:Ag caused by genetic mutations in the vWF:Ag gene located on chromosome 12.[4]

iii. Clinical Presentation

vWD is associated with an abnormal bleeding tendency. It manifests as a propensity to bruise easily and often spontaneously, have nose bleeds, bleed from the gums, and can result in sometimes serious hemorrhaging after surgery or dental extractions. In addition, heavy and prolonged menstrual periods (menorrhagia) and bleeding after childbirth are common manifestations in women afflicted with this disorder.[5] Even though the frequency of genetic transmission of this disorder is equal in males and females, it is likely that the additional challenges to the hemostatic system posed by monthly menstruation and childbirth account for the higher reported prevalence of females with

vWD.[6,7,8] Overall, the impact of vWD on the health of women worldwide is considerable,[7,9] which has lead to an expansion of interest and research studies in women with this bleeding disorder.

iv. The Classification of von Willebrand Disease

The current classification of vWD includes 3 main subtypes, referred to in the medical literature as Type I, Type II and Type III. Type I and Type III are caused by a mild and severe quantitative deficiency, respectively, of circulating vWF:Ag. Type II is caused by a qualitative defect in the vWF:Ag.[10] Type I constitutes approximately 75% of all reported cases and results in a mild to moderate bleeding tendency.[5] Given that this subtype is the most commonly encountered in clinical practice, it is also the most amenable and appropriate target for clinical research.

v. The Diagnosis of von Willebrand Disease

The laboratory diagnosis of vWD is established by demonstrating lower than normal levels of vWF:Ag, von Willebrand factor activity (Ristocetin cofactor, RcoF) and a prolonged bleeding time (BT). The traditional BT measures, in seconds, the length of time that an individual bleeds following a standardized skin incision. However, a newer method to measure the bleeding time non-invasively is now commonly used. It uses the Platelet Function Analyzer ® (PFA-100®)[11] instrument, which measures the closure time (CT) in seconds, and has a higher sensitivity and specificity (95-100% and > 95%, respectively)[12,13,14] than the traditional bleeding time (65%)[15] when used as a screening tool to diagnose vWD. Decreased Factor VIII activity (FVIII:C) and prolongation of the coagulation time (aPTT) are occasionally seen in severe variants of Type I vWD but are more typical of the other subtypes.[7,16]

Genetic defects in the vWF:Ag gene are only partially responsible for the variability in the blood levels of vWF:Ag.[17,18] There are, in addition, several other genetic modulators of vWF:Ag levels; younger individuals, blacks, males and individuals of blood type O

have lower levels of vWF:Ag in comparison to their older, non-black, female and blood type A, AB, B counterparts.[19,20,21,22,23] Furthermore, many acquired conditions, either physiological (pregnancy, follicular phase of the menstrual cycle, stress) or pathological (inflammation, thyroid disorder), medication intake (oral contraceptives, steroids) and intake of caffeine or nicotine can increase the blood levels of vWF:Ag.[20,24,25] The heterogeneity in blood vWF:Ag levels attributable to many of these concomitant factors introduces complexity in diagnosing vWD, especially since genetic tests for this disorder are still not routinely available. At the present time, the diagnosis of vWD is based on a combination of clinical history and laboratory criteria. Laboratory testing for vWD often needs to be repeated several times before definitive conclusions about the presence or absence of the disease can be firmly made. In recognition of this difficulty, the International Society of Thrombosis and Hemostasis (ISTH) devised a clinical standard for the diagnosis of vWD that includes definitions for definite and probable disease.[26] (The current definition of definite Type I vWD is outlined in the *Subject Eligibility Criteria* section of the crossover feasibility study on page 56 of this thesis.)

vi. Menorrhagia in von Willebrand Disease

Menorrhagia is subjectively reported in over 90% of women with vWD, making it the single most common bleeding manifestation in women with this disorder.[27] Menorrhagia is associated with considerable impairment in quality of life, work absenteeism, and psychological impairment. It can also lead to the development of iron deficiency anemia and, in some circumstances, can lead to women having to undergo unnecessary hysterectomies to control bleeding.[28,29] Given the high prevalence of menorrhagia in women with vWD, menorrhagia has become a major target of clinical research.

Menorrhagia is defined as excessive menstrual blood loss of greater than 80 ml per menstrual cycle.[30] The ability of both patients and physicians to diagnose true menorrhagia is notoriously poor.[31,32,33] The clinical assessment of menorrhagia is often limited by the patient's cultural and/or social background.[34] Useful clinical indicators of menorrhagia include a history of menses (menstrual period) of long duration, excessive

tampon and/or pad use, frequent soiling of clothing or bedding, excessive fatigue, significant impairment of quality of life, history or current iron deficiency anemia, a positive transfusion history, signs of volume depletion or a previous hospitalization or hysterectomy to control menstrual bleeding.[28,34]

The gold standard for objectively quantifying menstrual blood loss is the measurement of hemoglobin on blood tinged sanitary napkins by the alkaline hematin method.[35] However, since this is a highly specialized and laborious technique, it is not often used. Instead, a pictorial blood assessment chart (PBAC) has been developed as an objective tool to diagnose menorrhagia and provides a semi-quantitative measure of blood loss. Each cycle score is tabulated by adding together separate pre-assigned scores that reflect the number of pads or tampons used, the degree to which each of these is soiled, and the presence of flooding and leakage.[36] Objective blood loss measured by the gold standard is highly correlated with the menstrual score determined using the PBAC (Pearson correlation coefficient of 0.85).[36] It has been shown that the PBAC score has an 86% sensitivity and 89% specificity to diagnose menorrhagia defined as ≥100 ml of blood loss per cycle by the alkaline hematin technique.[36] A score of ≥100 using the PBAC is equivalent to ≥80 ml blood loss. Further, a score of ≥185 has a high positive and negative predictive value to diagnose menorrhagia.[37] The validity of this instrument to diagnose menorrhagia is, however, limited by the fact that it has been established only for women using a single brand of tampon or sanitary napkin supplied by the investigators.[36,37] The sensitivity and specificity of this tool to diagnose menorrhagia in users of different brands of sanitary devices has not been studied. Since menstrual blood loss and PBAC scores of a given individual are usually consistent between 2 consecutive menstrual cycles, a single PBAC score of > 100 is sufficient for the diagnosis of menorrhagia.[37] The PBAC is a simple, inexpensive and reasonably accurate method for assessing menstrual blood loss, and as such, has become a commonly used and useful tool both to diagnose menorrhagia and to evaluate responses to therapy for this condition. This scoring system is now used on a routine basis in gynecology clinics and many hemostasis clinics within North America.

Menstrual blood loss is controlled by hormonal, mechanical and hemostatic mechanisms. The formation of hemostatic plugs in the first 20 hours of a menstrual bleed is important for uterine hemostasis.[38] Thus, any disorder of blood coagulation, including disorders of blood vessels, platelet abnormalities, coagulation factor deficiencies and vWD can cause menorrhagia. In one study, 17% of women with menorrhagia were found to have a bleeding disorder.[39] Menorrhagia is the most common bleeding symptom reported in adult females with vWD who have reached the age of menarche, and occurs in 93.1% of all cases.[40] Using the PBAC score, 74% of the women with vWD scored > 100, as compared to 29% of the women in the control group.[41] The median PBAC score was 122 (range 38-482) in women with inherited bleeding disorders, as compared to 73 (range 9-310) in the control group.[41] Therefore, the PBAC score has proven to be useful to diagnose menorrhagia when applied to menstruating women with vWD.

vii. Treatment of Menorrhagia in Patients with von Willebrand Disease

There is currently no consistent approach to the management of menorrhagia in women with bleeding disorders.[8] Rather, there are a variety of both specific and non-specific approaches that are used to treat menorrhagia. Non-specific approaches include surgery (endometrial ablation, dilatation and curettage, hysterectomy), hormonal therapy (contraceptive pill, estrogen, progesterone) and the use of antifibrinoloytic agents. Surgical procedures offer a definitive reversal of menorrhagia but are associated with significant morbidity and are not ideally suited to treat women of childbearing age who comprise the large majority of women suffering form menorrhagia.[29] Although oral contraceptives and other hormonal therapies are effective in treating idiopathic menorrhagia (menorrhagia of an unknown cause), less than 50% of women with menorrhagia related to vWD respond to these interventions.[29] Women who do not desire contraception or who have a personal or family history of breast cancer or thromboembolic disease (propensity to form venous clots) are often not eligible for hormonal therapy. Antifibrinolytic agents work by solidifying the blood clot once it is formed. Tranexamic acid, a commonly used and inexpensive anti-fibrinolytic agent, has been shown to be effective in approximately 60% of women with menorrhagia of an

unknown cause.[42,43,44] Although several published case reports[45,46] and clinical experience suggest that this therapy is efficacious in treating menorrhagia related to vWD, this finding has not yet been studied prospectively. Antifibrinolytic agents are, in general, well tolerated but can cause nausea, headaches and dizziness. There is also a potential risk of intra-cranial thrombosis (a blood clot in the brain) associated with tranexamic acid, although this finding has not been confirmed in large-scale studies.[47]

Specific approaches aimed to correct the vWF:Ag deficit in vWD have become of increasing interest. The discovery of a drug called desmopressin *(1-deamino-8-D-arginine vasopressin, DDAVP)* has revolutionized the treatment of vWD over the past decade by virtue of its targeted effect on vWF:Ag.[48] DDAVP works by causing the release of endogenous vWF:Ag (and probably FVIII:C) from storage sites in vascular endothelial cells (cells that line blood vessels),[49,50] although the precise mechanism by which DDAVP induces this effect is still unknown. The intravenous administration of 0.3mcg/kg of DDAVP results in a 2 to 3-fold (and as high as a 5-fold) increase in the blood level of vWF:Ag and FVIII:C.[51,52,53] DDAVP also increases platelet function and shortens the bleeding time independently of its effect on raising vWF:Ag levels.[54,55,56,57] The *in vitro* bleeding time (PFA-100 ®) is useful for the therapeutic monitoring of patients with vWD who are treated with DDAVP.[58,59,60] Peak levels of vWF:Ag and FVIII:C are achieved 30 to 60 minutes after the DDAVP administration and begin to decline after 2 hours.[50] This response has been observed in healthy individuals,[51,61] as well as in patients with mild vWD and mild hemophilia A,[48,51,52,62,63] in whom DDAVP results in decreased clinical bleeding. Despite the genetically impaired production of vWF:Ag in Type I vWD, normal vascular storage pools permit a response to DDAVP that is equivalent to that of normal individuals.[64]

There are no published randomized studies of the effect of intravenous DDAVP in the treatment of menorrhagia. Published observational studies have shown that this agent is very effective to treat menorrhagia in women with underlying bleeding disorders, including vWD, and as a result, it has now become the accepted front-line targeted therapy for this indication.[65,66,67] Intranasal DDAVP has been used successfully to treat

menorrhagia in women diagnosed with mild or moderate type I vWD.[68,69] In one study, intranasal DDAVP decreased the mean (+/- SD) number of days of menstrual bleeding from 11 (+/- 9) to 6 (+/- 2) days in women with inherited bleeding disorders.[68] In another study, 92% of the female subjects with bleeding disorders and menorrhagia reported good to excellent responses using this treatment, which shortened and/or lightened 88% of their menstrual cycles.[69] In another recently published randomized double-blind clinical crossover trial, intranasal DDAVP failed to show any significant positive effect on the menstrual blood loss in women with mild types of vWD as determined by the PBAC.[70] The mean difference in PBAC scores between the DDAVP and the control group was 8 (95% CI–15.5-31.6). However, the possibility that blood levels of FVIII:C and vWF:Ag attained after delivery of the drug were subtherapeutic (since these were not measured), and the low power of this study due to the small sample size, limit any definitive conclusions on the effectiveness of intranasal DDAVP on menorrhagia.

Adverse effects have been reported in 29% to 68% of patients treated with DDAVP.[70,71] The most common side effects are mild, such as facial flushing, dizziness, headache, rhinitis (runny nose) and conjunctivitis (watery eyes). In more serious cases, chest tightness, severe headache, abdominal cramping and vomiting have been reported.[68,69,71] Seizures,[72] related to a significant decrease in the level of sodium induced by DDAVP, as well as isolated cases of myocardial infarction,[73,74] mostly in older individuals with established coronary disease, have also been reported. Another drawback of DDAVP therapy, especially if it is used on a regular monthly basis, is that it is expensive and, in its intravenous formulation, must be administered in the hospital.

The limited applicability and efficacy of certain of these therapeutic alternatives, in addition to their various contraindications, adverse side effects and high cost justify the search for new approaches to treat menorrhagia in vWD.

viii. Stress as a modulator of von Willebrand Factor Antigen

The limitations in the therapeutic alternatives discussed above have created an incentive to discover alternative, less toxic and expensive, mechanisms/interventions to DDAVP that can induce similar changes in vWF:Ag and improve hemostasis. Physiologic stress, either mental or physical, is an important but less commonly recognized modulator of vWF:Ag levels and is amenable to manipulation for this purpose. Published studies demonstrate that physical[75,76] and mental[77,78] stress can induce several changes in the coagulation system, most notably an increase in the level of circulating vWF:Ag and FVIII:C. The main explanation for this phenomenon is the release of these proteins from storage sites in vascular endothelial cells (cells that line blood vessels),[75] although there is also additional activation of inactive FVIII that is already present in the blood circulation.[79,80] Although the precise mediators of this release have not been identified, the effect of an acute rise in adrenaline and noradrenaline related to either physical or mental stress is highly plausible.[81]

The endogenous release of catecholamines, namely adrenaline and noradrenaline, mediate the sympathetically driven human "alarm reaction" (palpitations, tremor, sweating) in response to psychological stress.[82] Significant increases in noradrenaline and adrenaline have also been shown to occur in response to an intense cycling exercise and correlate strongly with other measures of exercise intensity, namely lactate and measures of maximum oxygen uptake.[83] Further, several studies in humans have shown that the direct intravenous infusion of adrenaline induces a 2-fold or greater increase in the level of circulating vWF:Ag[77,84] and FVIII:C[77,84,85,86,87], as well as a shortening of the BT (bleeding time)[86] and aPTT (coagulation time)[86,88] in both normal subjects[78,84,85,86,87,88] and patients with non-severe types of vWD[86,89,90] and hemophilia.[87,90] In these studies, responses were immediate, peaked at 1 hour and fell to baseline levels after 6 hours (the reported half-life of this effect ranges anywhere from 0.8-3.4 hours). These coagulation changes were prevented when these same individuals were treated with an anti-adrenergic agent (adrenaline receptor blocking agent) prior to the infusion of adrenaline.[85] Other studies have shown that vWF:Ag also increased in the culture milieu

of endothelial cells following an adrenaline infusion.[81,91,92] Further, it has been shown that blocking receptor sites for adrenaline and other catecholamines could abolish alterations in blood coagulation induced by physical exercise.[93,94] Seeing as adrenaline is known to increase significantly in situations of sympathomimetic discharge, in situations of high stress, either physical or mental, adrenaline release most likely mediates (at least in part) the changes in VWF:Ag and FVIII:C in response to stress.

In a study by Jern *et al*, 10 males between the ages of 22 to 30 (mean 25) were assigned to 3 different stress tests, a mental stress challenge, an intense physical exercise and an intravenous adrenaline infusion, on different days and in random order.[77] Mental stress caused a 29.5% and 73.5% relative increase in vWF:Ag and FVIII:C from baseline, respectively. Comparable changes were observed among the men when they were infused with adrenaline. Physical exercise at maximum capacity caused a 133% and 69% relative increase in vWF:Ag and FVIII:C from baseline, respectively. The absolute change in vWF:Ag and FVIII:C caused by mental stress relative to that caused by physical exercise was 31% and 38%, respectively.

Changes in circulating levels of vWF:Ag[77,79,84,95,96,97,98,99,100,101,102,103] and FVIII:C,[77,79,84,86,95,96,97,98,100,101,102,103,104,105,106,107,108,109,110,111,112] comparable to those seen after the administration of DDAVP, occur after strenuous physical exercise and range anywhere from 20% to 300%. Changes in FVIII:C are accompanied by a shortening of the aPTT in most studies where it was measured.[86,95,97,98,100,103,104,105,106,107,111] There is also evidence of platelet activation[79,96,113] (a main component of the initial blood clot) and increased thrombin formation (a main component of the definitive blood clot)[108,114,115,116,117] after vigorous exercise. The traditional bleeding time, a global functional measurement of vWF:Ag and platelet activity, was only measured in one study, in which it decreased significantly from baseline in 4 male and 2 female adult volunteers after both a 3 minute and a 2 hour vigorous exercise.[86]

The effect of exercise on the coagulation system among the various published studies is inconsistent. The discrepancies may be due in part to the use of different exercise

protocols, study populations and sample sizes in various studies. Earlier studies examined the impact of a short duration high intensity exercise on various coagulation parameters. In one study by Prentice *et al*, 2 untrained males had a 200% increase in both FVIII:C and vWF:Ag after running a ½ of a mile as fast as possible.[84] Rizza's study showed that there was a considerable variation in response among individuals doing a similar exercise.[112] The post-exercise FVIII:C change ranged from 139% to 333% among 15 untrained healthy subjects (4 women) who ran ¾ of a mile as fast as possible (average increase in FVIII:C of 219% from baseline). Increases of 180% to 300% and 143% to 200% were observed among 2 different individuals who repeated the same exercise on different days. When 5 of the individuals walked for 3 or 6 miles at a pace of 5 miles per hour, no change in FVIII:C was observed. Similar findings have been shown for more physically fit subjects. Changes of 187% and 115% in FVIII:C and vWF:Ag, respectively, were seen among 10 runners (5 female) between the ages of 21 and 56, after running ¾ of a mile as fast as possible.[102] When 3 of these individuals jogged in place for 30 minutes, comparatively lower increases of 35% and 32% in FVIII:C and vWF:Ag occurred. Mean increases of 100% and 38% in FVIII:C and vWF:Ag, respectively, were noted in a study by Stibbe on 10 untrained volunteers (4 females) ranging in age from 19 to 38, after running up and down a staircase at maximum speed for 3 minutes.[79] In this study, 2 of the 4 women ran for only 2 minutes because they were exhausted. When 5 healthy untrained men (ages 19-22) performed the same exercise, a mean increase in FVIII:C of 2 to 3 times over baseline and a concomitant shortening of the aPTT was seen despite the fact that none were exhausted after the exercise.[110]

Several studies have examined the effect of a graded treadmill or cycle exercise performed to exhaustion. vWF:Ag increased by 62% from baseline and was accompanied by a statistically significant shortening of the aPTT among 20 men who performed a treadmill test to exhaustion.[103] In another study, 20 trained males exercised to exhaustion and displayed changes of 0.51 in FVIII:C.[80] Responses were variable among individuals in that study, including 3 individuals who had less than or equal to 10% increments in FVIII:C. These 3 individuals also had the lowest maximum heart rates in the group. Individual differences in the subjective sense of exhaustion and its less than perfect

reflection of individual maximum exercise capacity[118] are limitations of that study and other studies that use this criterion as the endpoint. A similar observation is made by Iatridis *et al* in which 8 out of 59 untrained male subjects had equivocal or no increment in FVIII:C following an exhaustive exercise, whereas there was a 188% mean increase for the group and a significant shortening of the aPTT.[111] The variable response to exercise in this study may be related to the fact that exercise times ranged from 6 to 32 minutes before exhaustion was reached.

There are several studies that show more directly that exercise intensity, as measured by the post-exercise lactate, is an important determinant of the coagulation response to exercise. Lactate levels were moderately correlated with thrombin generation ($r = 0.62$ $p = 0.001$) among 30 individuals after running ($n = 10$), jogging ($n = 10$) or cycling ($n = 10$).[114] Post-exercise lactate levels in these 3 groups were 4.3 +/- 2.3, 1.4 +/- 1.0 and 2.3 +/- 1.1, respectively, which reflects the great variation in intensity achieved among the individuals performing these 3 exercises. A moderate correlation between post-exercise lactate and FVIII:C ($r = 0.65$ $p < 0.002$) was also present in a single group of 19 untrained males (ages 21-49) after a maximum treadmill exercise. In this same group of males, there was a 159% increase in FVIII:C from baseline and a 109% increase in vWF:Ag from baseline.[101] In another study, the aPTT shortened significantly when 10 untrained males (mean age 23) performed a cycle exercise to their predetermined lactate threshold, but no change was observed in the aPTT when they cycled to half of their maximum lactate threshold on a different day.[119]

In order to better define coagulation responses to different exercise protocols and intensities, Andrew *et al* studied 5 moderately active males between the ages of 25 and 44, each of whom performed 3 different exercises on separate occasions.[100] Direct comparisons of the effect of the 3 different exercises performed by the same group of individuals helped control for patient-related confounders that may have also influenced their response to exercise. Each participant performed a graded exercise to maximum intensity, a steady state low intensity exercise and a 30-second maximal exercise on a cycle ergometer. Various coagulation and fibrinolysis parameters were measured at

baseline and at several time points both during and after each exercise. During the graded exercise, at maximum effort, FVIII:C and vWF:Ag increased by 200% and 150%, respectively, from baseline. The mean blood lactate level was significantly higher (10 +/- 1.08) at maximum effort compared to rest and submaximal effort. Similar increments in FVIII:C and vWF:Ag (200% and 100%, respectively) were observed among the same individuals after doing a 30 second maximum cycle exercise, after which the mean lactate level was also elevated (13.2 +/- 1.31). However, when the same subjects performed a steady state low intensity exercise, after which lactate levels were barely increased (3.5 +/- 1.53), the FVIII:C and vWF:Ag levels were unchanged from baseline.

The duration of exercise at maximum intensity may be an important determinant of the peak coagulation changes due to exercise as well as the persistence of such changes. Changes of 194.5%, 96% and 144% in vWF:Ag were observed among 7 trained males between the ages of 24 and 45 who ran short (1.7 km), medium (4.8 km) and long distances (10.5 km), respectively.[99] The vWF:Ag increments were significantly higher in the short and long distance runs than in the medium distance run. These changes persisted for up to 10 hours. Other studies have shown that increased levels of FVIII:C persist anywhere from 6 to 8 hours after strenuous but short duration exercise (mean peak level of 200-300%)[110] and for as long as 24 hours after a marathon run (mean peak levels of 250-300%)[96] or triathlon race (mean peak level of 275%)[98]. These studies suggest that the peak coagulation changes and the persistence of such changes induced by exercise are also related to the duration of the exercise at maximum capacity.

Maximum exercise intensity for a given individual is measured by age-adjusted maximum heart rate, maximum oxygen consumption (VO_2 max), threshold respiratory oxygen/carbon monoxide ratio (RER > 1.0), or various combinations of these markers, in addition to exhaustion. When standardized intensities are used, more uniform coagulation responses occur with exercise. In one study by Lin *et al*, 11 male volunteers were subjected to a standardized exercise on a cycle ergometer until at least 2 criteria of those listed above were met.[104] The mean baseline level of the group was 108.6 (+/- 6.7) for FVIII:C, as compared to 274.1 (+/- 32.3) after exercise. The elevated levels of FVIII:C

remained statistically elevated for up to 6 hours after the exercise. The mean aPTT shortened from 27.6 (+/- 0.7) at baseline to 22.3 (+/- 0.5) immediately after exercise, although this shortening was no longer significant 2 hours after exercise.

Standardized protocols have also enabled a better understanding of the importance of exercise intensity on the relative balance between coagulation (clot formation) and fibrinolysis (clot dissolution) during exercise. Among 29 sedentary men, significant changes in FVIII:C were not observed until exercise was maximal; FVIII:C was 153 (+/- 10) at 100% VO_2 max, as compared to 85 (+/- 4) at rest.[107] Peak changes occurred in the early post-exercise period and were still significant after 25 minutes of rest (211 +/- 27). Significant shortening of the aPTT also occurred at maximal exercise and persisted into the recovery period. In contrast, fibrinolytic activity peaked at maximum heart rate and underwent a rapid decline after exercise. A similar pattern was noted among 10 healthy male volunteers who exercised to exhaustion on a cycle ergometer; their mean (+/- SE) FVIII:C measured at baseline was 107 (+/- 13), as compared to 269 (+/- 9) immediately after exercise (a 151% increase from baseline).[109] In this group, changes in FVIII:C occurred between 95%-100% of maximum heart rate, whereas significant changes in fibrinolysis were already present at 80% of maximum heart rate. It is evident from these studies that the achievement of maximal effort during exercise is essential in order to maximize coagulation changes during exercise and to isolate its effect from that of fibrinolysis in the post-exercise recovery period.

A single individual in the latter study had no change in his FVIII:C level despite having achieved maximum exercise as determined by objective criteria.[109] His maximum oxygen uptake (VO_2 max), physical activity and mental anguish at baseline were comparable to the rest of the group, as was his plasma volume loss from sweating during exercise. This observation suggests that there are other unknown modulators of the response to exercise, aside from intensity.

There are several studies that have examined the effect of physical conditioning on the response to acute exercise, although these yield conflicting results. Fergusson *et al*

showed that there was no difference in FVIII:C and aPTT among sedentary individuals, joggers and marathon runners, when they exercised to their maximum potential.[120] These results are corroborated by a study by El-Sayed *et al* in which 25 moderately active but untrained individuals were randomized to a physical conditioning program or to a control group.[106] Approximately ½ of the subjects in each group were female and mean ages were 33.4 (+/- 5.4) and 32.1 (+/- 6.4), respectively. Randomization was stratified according to age and baseline cardiorespiratory status. Both groups were subjected to a standardized cycle ergometer test to VO_2 max or exhaustion. VO_2 max increased significantly after conditioning in the experimental group, as compared to baseline. Comparable elevations in lactate occurred with exercise in both groups. Post-exercise levels of FVIII:C were 230 (+/- 100) and 250 (+/- 100) in the control and experimental groups, respectively, as compared to a baseline of 100 (+/- 20) in both groups before conditioning. The results after conditioning were not significantly different. Furthermore, the mean aPTT decreased from baseline and was not significantly different in each of the 2 groups.

These findings, however, have recently been challenged by the work of van den Berg *et al*, who showed that a higher FVIII:C and a lower aPTT were achieved in endurance trained individuals, as compared to sedentary individuals, following an exercise performed to maximum effort.[105] Another study published in the same year showed that untrained controls had significantly higher post-exercise increments in FVIII:C and greater shortening of the aPTT than conditioned individuals.[121] The precise effect of an individual's baseline physical condition on his/her coagulation response to maximal exercise remains to be determined.

Baseline vWF:Ag and FVIII:C increase with advancing age. However, a study conducted by van den Burg *et al* showed that age did not affect the coagulation response to exercise performed at maximum effort.[97] In this study, the increase in FVIII:C and the shortening of the aPTT after a standardized cycle ergometer test, to REQ > 1.15 or age predicted maximum heart rate, was comparable among sedentary males (total n = 38) in 3 different age categories (20 to 30, 35 to 45, and 50 to 60). The authors demonstrated that changes

observed with exercise were merely superimposed on baseline levels, which suggests that the vascular endothelium response to physiological stress was not altered by age in the ranges that they considered, despite considerable age-related decreases in work capacity and maximum heart rate. In a more recent study, van den Burg *et al* suggest that there is a differential effect of physical conditioning across ages in response to acute exercise; younger individuals display a higher FVIII:C and a shorter aPTT after physical conditioning in reference to baseline, as compared to older individuals, when allowed to exercise to their maximum potential.[95]

The effect of baseline levels of vWF:Ag and FVIII:C on the post-exercise increments has not been evaluated in healthy individuals. However, the effect of intense exercise on FVIII:C, vWF:Ag, aPTT and BT in individuals with vWD and hemophilia (genetic diseases that lead to a decreased production of blood levels of vWF:Ag and FVIII:C, respectively) has only been looked at in two studies. The first study by Edgeberg showed a normalization of FVIII:C, bleeding time and aPTT in 2 women with mild vWD who had a prolonged bleeding time and aPTT at baseline, and 50% of normal vWF:Ag levels after running ¾ of a mile as fast as possible.[110] In contrast, 2 individuals with severe vWD had no response to the same exercise. In the second study by Stibbe, an individual with mild vWD who ran up and down a staircase for 2 minutes normalized her bleeding time and FVIII:C.[79] A mean change of 250% in FVIII:C, comparable to that seen in the healthy individuals in the same study, took place in 2 mild hemophiliacs (deficient in FVIII:C) after running ¾ of a mile as fast as possible. However, the effect of a standardized exercise on individuals with either vWD or hemophilia has not been published.

ix. Sweat as a Marker of Exercise Intensity

Sweating related to intense physical exertion is also sympathetically driven and is likely related to the surge of adrenaline and noradrenaline (catecholamines) that occurs during intense exercise.[122,123] Sweating patterns in response to intense physical activity show that sweating thresholds vary according to body site. In males, the forehead and the lower

back consistently produce the highest volume of sweat and have the highest sweating rate in response to exercise.[124] The threshold for sweating on the forehead and chest remains constant whereas the threshold for sweating on the forearm increases with increasing exercise workloads.[125,126] As such, sweating rates measured on the forearm most accurately reflect the vigor of the exercise.

Age itself is not known to affect the degree of sweating responses to exercise, but the decline in sweat production observed with aging is more closely related to the concomitant decrease in VO_2 max (maximum oxygen consumption).[127] While sweating does not differ substantially between fit and non-fit individuals, the delay to the onset of sweating is shorter in trained individuals.[128]

There is no standard method to measure the amount of sweat produced during exercise. Sweat is comprised of salt (sodium chloride) and water. A filter paper patch test initially developed as a screening tool for cystic fibrosis can be used to collect the salt produced in sweat.[129,130,131] These patches can also be useful in measuring the amount of salt produced in sweat during exercise. Further, because salt becomes permanently trapped in these patches, this measure remains accurate even if the analysis of the patches is delayed for an extended period of time. A modified version of this patch was developed to collect sweat produced during exercise for the purpose of the present study.

x. Limitations of the Existing Literature

There is little doubt that exercise can bring about changes in vWF:Ag and FVIII:C under a variety of conditions. However, there are several shortcomings in the published literature in this regard. First, most of the studies that have addressed this phenomenon are based on small sample sizes. Although no confidence intervals for the means are provided in any of the studies, these would likely reflect imprecise estimates due to the small sample sizes involved. Second, multiple outcomes in addition to vWF:Ag and FVIII:C were measured in most studies, which may have resulted in statistically significant findings by chance due to a high Type I error.[132] Third, although the effect of

exercise on the shortening of the aPTT (a reflection of the increases in FVIII:C) has been examined in numerous studies, there has been virtually no measure of the effect of exercise on the bleeding time (a reflection of the level of vWF:Ag), which is consistently abnormal in individuals with vWD. This shortcoming may be related at least in part to the fact that the PFA-100® instrument, which reliably measures the *in vitro* bleeding time was not routinely available when these studies were conducted. Fourth, the results of these studies, carried out predominantly on male volunteers, may not be completely generalizable to women. The effect of exercise on coagulation has not been studied exclusively in women and the various studies in which women are included (most often comprising a minority)[79,102,106,112] did not have sufficient power to assess gender differences. Fifth, although the most impressive changes in FVIII:C and vWF:Ag were seen in trained athletes after long distance marathons,[96,98,99] this type of exercise intervention has limited applicability to untrained individuals. Sixth, neither exercising to exhaustion, the expensive equipment used to determine VO_2 max, nor the sequential blood tests used to measure lactate during exercise in most studies, is practical for routine use or use in large-scale studies. Finally, the effect of exercise has been studied in only 3 women with vWD.[79,110] Our interest was to study the effect of exercise on vWF:Ag, FVIII:C, CT and aPTT in healthy women, using a standardized exercise protocol that could be done at home and easily integrated into one's daily activities, as a potential therapeutic modality to treat menorrhagia in vWD.

STUDY HYPOTHESIS:

Both exercise and DDAVP can induce significant increases in circulating vWF:Ag and FVIII:C. The therapeutic efficacy of DDAVP in treating menorrhagia in women with vWD has been clearly established. However, the possible beneficial effect of exercise on menorrhagia in women with vWD has not yet been addressed in a formal clinical trial. If the induced levels of these proteins attained with exercise are sufficient to correct the laboratory profile of affected women, then it may be reasonable to hypothesize a positive effect on their menstrual bleeding, analogous to the effect of DDAVP therapy.

Contrary to the usual tendency to remain inactive during heavy menses, we hypothesize that short periods of daily vigorous exercise may significantly reduce the amount of menstrual blood loss in women with Type I vWD. Our hope is that this simple and physiological intervention could be considered as a viable alternative to pharmacological therapy to treat menorrhagia in highly motivated women affected by this disorder.

STUDY OBJECTIVES:

Part I

The main objective of Part I of this study was to assess the impact of a high intensity graded stepping exercise on the level of circulating vWF:Ag, FVIII:C and their respective functional measures, namely, the CT (closure time) and the aPTT (coagulation time). A secondary objective of this study was to evaluate the utility of a filter paper patch to collect sweat during exercise as an index/indicator of compliance with exercise; to assess the relation between the amount of salt (referred to herein after as sodium) collected in the sweat patch during exercise and the change in vWF:Ag; and to assess a similar relation for several other blood markers of exercise intensity.

Part II

The main objective of Part II of this study was to assess the feasibility of a 4-period crossover trial in order to evaluate the effectiveness of an intensive home exercise protocol on the quantity of menstrual blood loss (using a pictorial blood assessment chart (PBAC)) in adult women with established Type I vWD and menorrhagia.

FUNDING:

Phase I

This study was supported by a 'Care until Cure' research grant from the Canadian Hemophilia Society and the Genetic's Institute.

Phase II

This study is being supported by a research grant from Aventis Behring Canada.

RESEARCH TEAM:

Phase I

Georges-Etienne Rivard MD - Principal Investigator
Rochelle Winikoff MD - Project Director, Co-Investigator
Claire Infante-Rivard MD, PhD - Collaborator
Patrick St-Louis PhD - Collaborator

Phase II

Georges-Etienne Rivard MD - Principal Investigator
Rochelle Winikoff MD - Project Director, Co-Investigator
Dianne Francoeur MD - Co-Investigator
Claire Infante-Rivard MD, PhD - Collaborator

ETHICAL CONSIDERATIONS:

The Ethics Committee of Sainte-Justine Hospital approved both studies. All subjects gave their informed consent to participate in this study. Unique codes were used to identify subjects and their test results in our hospital-based computer system. The only nominal information appeared on the data collection sheets used by the project director during the initial interview. These records were kept in a secure place and were only accessible to members of the investigating team for the purpose of this study. Access to the results was offered to subjects upon request.

PHASE I

METHODS:

Study Design

A one-group pre-test-post-test design, without a control group, was used to study the research hypothesis. Laboratory measurements of several coagulation parameters were taken before and after an exercise intervention on a single experimental group consisting of healthy female volunteers.

Study Population

The target population consisted of healthy pre-menopausal adult women (between the ages of 18-40 years old).

Subject Recruitment

A convenience sample of healthy pre-menopausal females comprised the experimental group. Volunteers were recruited from Sainte-Justine Hospital. Publicity for this study was made by word of mouth and from announcements posted in the hospital. A monetary honorarium was offered to participants.

Subject Eligibility

Pre-menopausal women between the ages of 18 to 40 were eligible for the study.

All volunteers answered a medical questionnaire and underwent a physical examination in order to assess their eligibility for the study. The exclusion criteria were:

- irregular menstrual cycles (absence of regular monthly menstrual periods)
- a personal or family history of bleeding or a known bleeding disorder

- established coronary artery disease (CAD) or risk factors for CAD including a family history of premature CAD
- a history of exertional syncope (loss of consciousness with heavy exertion)
- a history of impaired balance
- a significant motor disability
- asthma or other pulmonary conditions
- significant anemia
- pregnancy (known or a positive urine pregnancy test)
- an uncontrolled thyroid disorder (known or an abnormal TSH result)
- chronic steroid use or sympathomimetic drugs (drugs that stimulate the cardiovascular system)
- difficult venous access
- women who did not feel that they could complete the described exercise protocol for any reason

Eligible candidates signed a written consent form.

Study Protocol

All participants selected for this study were required to attend a private exercise session, scheduled at 8:00 a.m. in a designated room with the temperature set at 22^0 C. All participants were asked to refrain from smoking, from ingesting caffeine and foods that are high in fat content, and from intense physical activity for 24 hours prior to their designated exercise session.

Each exercise session began with a 30-minute rest period. A filter paper patch was applied to an alcohol cleansed area of both the forearm and forehead of each participant at the start of the rest period and were removed after 30 minutes had elapsed (control patches, Time 0). Baseline blood samples were drawn immediately after the rest period (Time 0). The exercise protocol was then reviewed with each participant, after which two new filter paper patches were applied. After a 5-minute stretch period, each participant

was required to perform a step exercise, at a rate of 32 steps per minute. A music cassette specially designed for step exercise (The Beat®, 1995) guided this pace. The exercise session was carried out under medical supervision. Each participant wore a cardiac monitor (Polar Beat®; Polar Electro, 1995) with a continuous heart rate display. Participants were coached to achieve their maximum heart rate during exercise and were advised to decelerate if and when they exceeded their maximum heart rate. After completion of the exercise session, a second set of blood samples was drawn (Time 1). During a 30-minute rest period that ensued, both patches were removed and replaced with new ones and as often as necessary, until the patches no longer became wet with sweat (Time 1). A final set of blood samples was drawn after the second 30-minute rest period had ended (Time 2).

Exercise end-points

Each participant was instructed to exercise until either (i) the patch on the inner surface of the forearm became wet, or (ii) the participant became physically exhausted and was unable to continue exercising.

Blood sampling and sodium measurements

The tests that were done on the blood samples included blood type, hemoglobin, hematocrit, platelet count, aPTT, vWF:Ag, FVIII:C, CT with each of epinephrine and ADP (adenosine triphosphate), epinephrine, norepinephrine and lactate. All blood samples were drawn with the help of a technical assistant and were drawn from the antecubital vein (central arm vein) with minimum stasis (tourniquet removed immediately once needle inserted) using a 22" gage needle short tube butterfly catheter. Samples were drawn using a two-syringe technique so as to ensure that there was no spurious local FVIII:C production at the venipuncture site. With the exception of blood type, which was drawn only once at baseline, all other blood tests were done at baseline (Time 0), immediately after completion of the exercise (Time 1) and 30 minutes after completion the exercise (Time 2). The sodium was measured by extracting it from the

filter paper patches placed on the participant's forearm during and after the exercise session. A detailed description of specimen processing and analytical instruments and methods is provided in Table 1.

Data Collection, Processing and Verification

Paper data collection sheets were used during the initial candidate interviews. Age, menstrual phase, contraceptive use, smoking status and fitness level were determined for each candidate at the time of the initial interview. This information was then transferred manually into a SAS compatible database. All entries were verified in duplicate for accuracy. All laboratory test results were initially stored in our hospital-based computer system and then retrieved after the completion of the study and entered into the SAS compatible database. Verifications for illegitimate codes and illogical entries were made on the final database.

Study Variables

The change in vWF:Ag due to exercise was used as the main dependent variable in the regression analysis, although both FVIII:C and vWF:Ag were measured before and after exercise. Using FVIII:C change as the dependent variable in a separate regression analysis was felt to provide redundant information, since these two factors are closely related physiologically, are known to change in parallel with exercise and, in this study, were highly correlated both at baseline ($r = 0.68$ $p < 0.01$) and at both time points after exercise ($r = 0.81$ $p < 0.01$ and $r = 0.85$ $p < 0.01$ for Time 1 and 2, respectively). Further, factors that determine blood levels of vWF:Ag are better established in the literature than those that determine blood levels of FVIII:C. The post-exercise level at Time 2 was used to calculate the change in vWF:Ag from baseline in all regression models. Measuring the change from baseline to Time 2, as opposed to the change from baseline to Time 1, more accurately reflects newly released vWF:Ag from its storage sites due to exercise since the additional concentration effect induced by sweating during exercise, which would be apparent immediately after exercise, is no longer present 30 minutes after the exercise.

The main exercise marker variables (or independent variables) tested in turn, in *separate* regression models, included the following: sodium change, adrenaline change, noradrenaline change and lactate change. The changes in adrenaline, noradrenaline and lactate from baseline (Time 0) to immediately after exercise (Time 1) were used in the regression analysis. The change in sodium in the forearm patch rather than that in the forehead patch was used in the regression model since it became wet much closer to the time of maximal heart rate for all participants studied. The change in sodium was calculated by subtracting the baseline sodium level in the control patches from the total sodium content of all the patches worn during and after the exercise session.

Fitness level and age were evaluated as covariates in regression. Fitness level was determined according to the score calculated from a fitness questionnaire administered by the investigator at the time of the initial interview. A copy of the fitness questionnaire is attached as Appendix A. Age was determined at the initial interview.

The definitions and coding of all study variables are provided in Tables 2 through 5.

STATISTICAL ANALYSIS:

All analyses were carried out on the data set generated during the study using SAS version 8 statistical software package.

Our data set included observations made on 40 individuals. There were no missing values for vWF:Ag. We identified 2 individuals who had missing values for both adrenaline and noradrenaline at all 3 time points, and an additional 9 values of various variables that were missing among the remaining individuals. Given the limited sample size and the fact that we had no *a priori* reason to suspect that the values for these 2 individuals differed systematically from those that were present in the data set (since the tubes for these individuals were known to be lost), the missing values for these 2 individuals, as well as for all other isolated missing observations, were replaced by the median values of

the group of the respective variable and were included in the analysis. In order to ensure the absence of bias by using this approach, all final models were run both with and without observations with missing values. Since the results were not affected by the removal of observations with missing values, these observations were kept in the final models. There were no outliers or illogical entries detected on the initial inspection of the data set.

First, descriptive statistics for all 4 exercise marker variables, namely, sodium change, lactate change, noradrenaline change and adrenaline change (independent variables), and for vWF:Ag change (dependent variable), hemoglobin, hematocrit and platelet count were calculated. A paired-Student's t-test was then carried out for vWF:Ag, FVIII:C, aPTT, CT ADP and CT epinephrine. Separate paired analyses were carried out for Time 0 and Time 1, and for Time 0 and Time 2. A two-sided alpha level of 0.05 was chosen to indicate significance. Linear regression was then used to assess whether any of the 4 exercise markers measured in this study could be useful in predicting the change in vWF:Ag levels due to exercise.

The change in vWF:Ag was used as the dependent variable in regression. While, several authors advocate the use of the post-intervention value as the dependent variable in regression modeling (while adjusting for baseline levels), due to its greater statistical efficiency than modeling change scores,[133] we felt that predicting the absolute level of vWF:Ag achieved after exercise in women who start out from normal baseline values has little clinical applicability. A more clinically relevant outcome is the degree to which these individuals change from their respective baseline values, since predictors of this change may also be relevant to women with vWD who have subnormal levels at baseline. Since these are two fundamentally related approaches, the choice of method is unlikely to influence to a large degree the findings with respect to the exercise marker of interest. Further, statistical efficiency has been reported to be similar between both approaches when the correlation between the pre-intervention and the post-intervention value is greater then 0.6.[134] To verify this assumption, a regression model was run in duplicate

using vWF:Ag measured after exercise while adjusting for baseline levels, and the regression parameters of the exercise marker variables did not change meaningfully.

The choice of using Time 2 to define the change in vWF:Ag outcome variable rather than Time 1 was expected to have little effect on the results, since the vWF:Ag levels at the 2 time points after exercise were highly correlated ($r = 0.99$ $p < 0.01$). Carrying out the regression analysis using the change in vWF:Ag from baseline (Time 0) to immediately after exercise (Time 1) instead of 30 minutes after exercise (Time 2) had a negligible impact on the results.

A univariate analysis was carried out using sodium change, lactate change, noradrenaline change and adrenaline change as the main independent variables in 4 separate models, each with vWF:Ag change as the outcome. The p-values of the regression coefficients were used to determine the significance of the relationship between each independent variable and the outcome using an alpha level of 0.05 to indicate significance. Ninety-five percent confidence intervals (95% CI) were also estimated for all regression coefficients in the univariate analysis. The residual plot of each of these variables was examined and showed no major violation of any of the regression assumptions.

Four different multivariate models were then considered, each using 1 of the 4 markers of exercise intensity as the main independent variable. These variables are: sodium change, lactate change, noradrenaline change and adrenaline change, respectively. In order to evaluate the usefulness of each of these exercise markers as a predictor of the exercise-related change in vWF:Ag, each was studied in a separate model. Since the causal nature of the relationship between the extent of change in the exercise marker and the change in vWF:Ag due to exercise was not the main goal of the regression analysis, it was not considered important to adjust these exercise variables for one another in any given model. Our goal was to identify which of the 4 exercise markers was the best predictor of the change in vWF:Ag resulting from exercise, based on statistical inference and clinical utility. Age (continuous) and fitness level (dichotomous) were then evaluated as covariates in each of the 4 multivariate models.

Potential effect modifiers were tested by adding plausible interaction terms, one at a time, to the base model, which already included the main independent variable of interest (exercise marker variable), as well as age, fitness level and baseline vWF:Ag (only in the case of testing effect modification by baseline vWF:Ag). First order interactions with the main exercise marker variable and each of age, fitness level and baseline vWF:Ag were studied in their respective models. Interaction terms were created by multiplying the main independent variable by each covariate of interest. The significance of the interaction between continuous variables was tested by comparing the significance of the F-statistic derived from the model with the interaction term to that derived from the model without the interaction term, using the appropriate numbers of degrees of freedom (multiple partial F-test). A significance level of 0.05 was used to determine whether the interaction terms should be kept in the multivariate model. There were no significant interaction terms found and, as such, no interaction term was retained in any of the 4 models.

The presence of confounding by age and fitness level was assessed through the change in estimate method.[135] The regression coefficient for the main exercise marker variable in each multivariate linear regression model (MLR) that included fitness level and age as covariates was compared to the regression coefficient for that same variable in an MLR model with each covariate removed in turn. Covariates that changed the regression coefficient of the exercise marker variable by 10% or more were deemed to be confounders of the association between exercise and vWF:Ag change in this study and were included in the final multivariate models on that basis. Age was not found to be a significant confounder of the relationship between the change in any of the exercise markers and the change in vWF:Ag due to exercise. Fitness level was found to be a significant confounder of the relationship between the change in each of lactate, noradrenaline and adrenaline, and the change in vWF:Ag in their respective models. Neither age nor fitness level confounded the relationship between the change in forearm sodium and the change in vWF:Ag. Since fitness level was felt to be an important confounder of the relationship between exercise intensity and the change in vWF:Ag *a priori*, it was also retained in the change in forearm sodium model despite the absence of

demonstrated confounding in this study. Age was not included in any of the final models, seeing as it was neither an important effect modifier nor a confounder of the relationship between exercise and vWF:Ag change.

The final models included each of the exercise marker variables of interest and fitness level. The overall F-test and the adjusted R^2 value are reported for each of the final multivariate models. The regression coefficients, standard errors, p-values and 95% CI for each exercise marker variable (main independent variable) are reported in their respective models. Standardized regression coefficients were calculated for all 4 exercise markers in their respective models in order to allow for a meaningful comparison of their relative utility in predicting the change in vWF:Ag (taking into account differences in measurement units and measurement error) when adjusted for fitness level. This standardized coefficient was calculated by multiplying the regression parameter of the independent variable of interest by the standard deviation of that variable in the data set and then dividing that number by the standard deviation of the vWF:Ag change dependent variable.[136]

Collinearity diagnostics were run on all of the final models by examining variance inflation and Eigen values. There was no evidence of collinearity. Potential outliers and influential observations were identified by observing the distribution of each variable and by using r Studentized and Jacknife residuals, Cook's distance and DF Betas for each observation. The few outliers that were identified were investigated by reviewing the original data for the individuals in question. Because these outliers were not found to be related to illogical and or erroneous data entries, but rather to true extreme values (since they corresponded closely to the extreme values obtained in these same individuals for FVIII:C measures), the outlying data was kept in the final models without modification. In any event, analyses were run both with and without these outliers and did not materially change the overall significance of the model or the estimated regression coefficients in the respective models. Further, visual analysis of the residual plots for each of the 4 models showed that the major assumptions of linear regression were not severely violated.

SAMPLE SIZE DETERMINATION:

A priori, an increase of at least 10% in either vWF:Ag or FVIII:C, either relative or absolute, was considered to represent a clinically meaningful response to exercise. Postulating a two-sided alpha error of 0.05, the study would have 80% power to detect a mean significant increase (defined as a greater than 10% mean increase from baseline) in levels of vWF:Ag and FVIII:C using approximately 40 subjects.[137] The within-subject standard deviation of vWF:Ag and FVIII:C on consecutive determinations is approximately 10% to 15% in our laboratory; as such, 15% was used as a more conservative estimate. A mean difference of 10% was considered to be clinically meaningful and of sufficient magnitude to correct the aberrant laboratory profile of many women affected with Type I von Willebrand disease. It was not considered pertinent to adjust the sample size of this study for noncompliance, since all sessions were supervised until completion.

RESULTS:

Study Accrual

Forty subjects were recruited over a 3-month period. All selected participants presented for their exercise session on the assigned date.

Exercise end-points

The utility of the filter paper patch on each site (forearm and forehead), as a rough indicator of exercise compliance, was assessed among women who successfully completed the standardized exercise protocol. Of the two sites studied, the patch on the forearm corresponded more closely to the achievement of maximum heart rate and was not wet before a minimum of 15 minutes of exercise had elapsed.

All participants achieved the exercise endpoints as outlined in the study protocol; all but 1 participant stopped when the forearm patch became soaked and the last participant stopped due to exhaustion. The mean exercise time was 23 minutes, with a range of 15 to 29 minutes. Forty-five percent of the participants achieved their maximum heart rate, while the remaining 55% of the participants achieved sub-maximal heart rates ranging from 90% to 99% of their maximum heart rates.

Exercise Complications

The exercise protocol was well tolerated by the majority of participants. Two participants experienced transient pre-vagal reactions immediately following the exercise protocol and a third had weakness accompanied by a sustained sinus tachycardia (rapid heart beat) for the first 2 hours after exercise that subsequently resolved. All 3 individuals were in the lowest activity level category at baseline. A fourth participant developed a contact dermatitis at the site of application of her forehead filter paper patch that resolved completely following 3 days of application of topical 1% hydrocortisone and did not recur.

Descriptive Statistics

Baseline Characteristics of the Study Group

Tables 6 and 7 present the baseline clinical characteristics of the study group. All participants, except for 1, were Caucasian. The mean age of participants was 25 years old and 70% of the participants were less than 30 years old (not shown). Approximately ½ of all the participants were physically fit at baseline, were users of caffeine and the contraceptive pill and were of blood type O. Only 10% of the participants were smokers. The distribution across menstrual phases among women participating in this study (luteal 47.5%, follicular non-menstruating 37.5% and follicular menstruating 15%) was proportionate to the known relative duration of each of the phases in a normal menstrual cycle.

Exercise Markers

As can be seen from Table 8, mean changes (+/- SD) in lactate, adrenaline and noradrenaline from baseline (Time 0) to immediately after exercise (Time 1) were 5.57 (+/- 2.31), 86.23 (+/- 130.92) and 1689 (+/- 1290), respectively. The mean change in forearm sodium from baseline (Time 0) compared to immediately after exercise (Time 1) was 3.26 (+/- 1.72). The minimum change in sodium incurred by any participant was 0.8 mmol/l.

Hemoglobin, Hematocrit and Platelets

Both the hemoglobin and the hematocrit levels increased by 8% from baseline (Time 0) to immediately after exercise (Time 1). Levels returned to baseline after 30 minutes of rest (Time 2). There was a 39% increase in the platelet count at Time 1 that returned to baseline levels at Time 2 (not shown).

vWF:Ag Change

As shown in Table 9, the mean change (+/- SD) in vWF:Ag from baseline to Time 1 and Time 2, was 0.31 (+/- 0.26) and 0.23 (+/-0.25), respectively.

Univariate Analysis

Paired Student t-test

Table 10 presents the means (+/- SD) and 95% CI of vWF:Ag, FVIII:C, aPTT, CT ADP (PFA-100®) and CT epi (PFA-100®) before exercise (Time 0), immediately after exercise (Time 1) and 30 minutes after the exercise (Time 2). The mean baseline values of each of these tests were within normal limits (according to the normal values established in our laboratory), thus providing reasonable assurance that individuals with

undiagnosed bleeding disorders were not inadvertently included in this study. As can be seen from the results in Table 10, there was a 0.31 U/ml absolute mean increase in vWF:Ag levels from baseline (Time 0) to immediately after exercise (Time 1) ($p < 0.01$), which corresponds to a mean increase of 37%. Similarly, for FVIII:C, a 0.65 U/ml mean absolute increase was observed from baseline (Time 0) to immediately after exercise (Time 1) ($p < 0.01$), which corresponds to a 56% mean increase. These changes persisted and were still significant 30 minutes after exercise (Time 2). The corresponding absolute mean increase in vWF:Ag measured at 30 minutes after exercise (Time 2) was 0.23 U/ml ($p < 0.01$), representing a mean increase of 27% from baseline. Corresponding values in FVIII:C measured at 30 minutes after exercise (Time 2) were 0.43 U/ml and 39%, respectively ($p = 0.02$). Differences in both vWF:Ag and FVIII:C were significantly higher immediately after exercise (Time 1) than after a 30 minute rest period (Time 2) ($p < 0.01$, $p < 0.01$, respectively). The mean absolute and relative changes in vWF:Ag and FVIII:C are depicted in graphic form in Figures 1 and 2.

Figures 3 and 4 illustrate the frequency distribution of the mean absolute increase in vWF:Ag and FVIII:C, respectively, immediately following exercise (Time 1), as compared to baseline (Time 0). As can be seen in these Figures, 92% of participants achieved a > 10% absolute increase in vWF:Ag levels immediately after exercise (Time 1) as compared to baseline (Time 0), and 65% achieved a > 20% absolute increase. Similarly, 97.5% of all participants achieved a > 10% absolute increase in FVIII:C levels immediately after exercise (Time 1) as compared to baseline (Time 0), and 92% achieved a > 20% increase.

Furthermore, exercise was accompanied by a mean shortening of 1.0 second of the aPTT from baseline to immediately after exercise ($p < 0.01$) that persisted but was no longer significant at 30 minutes after exercise ($p = 0.08$). Closure times (CT) decreased significantly after exercise with a mean absolute decrease of 12.6 seconds and 24.8 seconds with ADP and epinephrine, respectively ($p = 0.01$ and 0.01, respectively). These changes persisted and remained statistically significant 30 minutes after the exercise ($p = 0.05$ and < 0.01, respectively).

Simple Linear Regression

A univariate analysis was performed using simple linear regression. The unadjusted regression coefficients and their accompanying standard errors, p-values and 95% CI for the association between the changes in sodium, lactate, noradrenaline and adrenaline, respectively, and the change in vWF:Ag are shown in Table 11. The squared correlation coefficients are also reported for each univariate model.

As can be seen from Table 12, the regression coefficient for a 1 unit change in sodium (rounded to 2 decimal places) was 0.05 (95% CI 0.00-0.09). The regression coefficient for a 1 unit change in lactate (rounded to 2 decimal places) was 0.03 (95% CI -0.00-0.06). The corresponding changes in vWF:Ag for a 5 unit and a 10 unit change in each of sodium and lactate are also shown. The regression coefficient for a 1000 unit change in noradrenaline was 0.02 (95% CI -0.04-0.08). A 100 unit change in adrenaline corresponded to a mean change in vWF:Ag of 0.02 (95% CI -0.04-0.09) on average. Mean changes in forearm sodium of 3.26 and of 5.57 for lactate corresponded to mean changes in vWF:Ag of 0.15 and 0.17 from a baseline of 0.83 U/ml, respectively. Mean changes in noradrenaline of 1689 and of 86.23 for adrenaline (not shown) corresponded to mean changes in vWF:Ag of 0.03 and 0.02, respectively. As shown by the squared correlation coefficients for these variables, the change in sodium level alone explained 10% of the total variation of the change in vWF:Ag due to exercise ($p = 0.04$), whereas lactate alone explained 8% of the total variation. Individually, noradrenaline and adrenaline changes explained 1% and 2%, respectively, of the variation of the change in vWF:Ag. These results suggest that the change in forearm sodium alone explains a small but significant amount of the change in vWF:Ag due to exercise. The change in lactate explains a comparable amount of the change in vWF:Ag due to exercise, although this estimate is of only borderline significance given its corresponding 95% CI that includes zero.

Modeling vWF:Ag after exercise, while adjusting for baseline levels, resulted in regression coefficients of a similar magnitude for each of these variables, albeit with slighter wider confidence intervals. Results were essentially unchanged when the level of vWF:Ag immediately after exercise (Time 1) was used to calculate the change in the dependent variable.

Multivariate Analysis

A multivariate analysis was performed using linear regression. All estimates of the effect of exercise intensity variables were adjusted for fitness level, which changed the estimate of effect of each exercise intensity variable by at least 10% when added to each of the 4 separate models containing change in lactate, adrenaline and noradrenaline, respectively. Because the baseline fitness level was considered to be an important confounder of the relationship between exercise intensity and the change in vWF:Ag due to exercise on biological grounds, it was retained as a variable in the sodium change model as well, although it was not a confounder based on the change in estimate method. No interaction terms were added to the models, since no effect modification was detected between the exercise intensity variable of each model and the candidate covariates, namely, baseline vWF:Ag, age and fitness level, when tested for this purpose. The regression coefficients, standard errors, 95% CI and p-values for each of the interaction terms tested are shown in Table 13.

The regression coefficients and their accompanying standard errors, p-values and 95% CI for changes in forearm sodium, lactate, noradrenaline and adrenaline, when adjusted for fitness levels are shown in Table 14. As shown in Table 15, the regression coefficients for a 1 unit change in sodium and a 1 unit change in lactate (rounded to 2 decimal places) were 0.05 (95% CI 0.01-0.09) and 0.02 (95% CI -0.01-0.06), respectively. Corresponding regression coefficients for a 5-unit and a 10-unit change in each of sodium and lactate were 0.26/0.52 and 0.11/0.22, respectively, when adjusted for fitness level. The regression coefficient for a 1000 unit change in noradrenaline, when adjusted for fitness level, was 0.02 (95% CI -0.04-0.08), and was 0.12 for a 5000 unit change. The regression

coefficient for a 100 unit change in adrenaline was 0.05 (95% CI 0.01-0.11) and was 0.24 for a 500 unit change when adjusted for fitness level. Mean changes in forearm sodium, lactate, noradrenaline and adrenaline in this group were associated with an average change of 0.17, 0.12, 0.04 and 0.04 U/ml in vWF:Ag, respectively, when adjusted for baseline fitness level (not shown). Mean changes of 4.81, 11.36, 10 416.67 and 520.83 in forearm sodium, lactate, noradrenaline and adrenaline, respectively, would be sufficient to change the vWF:Ag by 1 standard deviation (SD). The squared partial correlation coefficients (shown in Table 14) for these 4 variables indicate that they explain 14.64%, 4.18%, 1.80%, and 6.63%, respectively, of the total variation of the change in vWF:Ag due to exercise that is not otherwise explained by fitness level. These results confirm the findings in the crude analysis, and again suggest that the change in forearm sodium plays a significant role in predicting the change in vWF:Ag due to exercise. The magnitude of the sodium effect was similar to that in the crude analysis. The magnitude of the adjusted regression coefficients for noradrenaline and adrenaline was higher than in the crude analysis, whereas that for lactate was comparatively lower than in the crude analysis. The corresponding 95% CI for these latter parameters are wide and include zero.

The standardized regression coefficients for changes in sodium, lactate, noradrenaline and adrenaline were 0.36, 0.20, 0.13, 0.25, respectively (shown in Table 14). The higher magnitude of the standardized coefficient for sodium change as compared to those for the other exercise marker variables in their respective models (all adjusted for fitness level), suggests that, of the 4 exercise markers, the change in sodium is the strongest predictor of the amount of change in vWF:Ag due to exercise.

The overall significance levels and adjusted squared correlation coefficients for all of the 4 final models are shown in Table 16. Together, sodium and fitness level (Model 1) explain almost 20% of the total variation of the change in vWF:Ag due to exercise, whereas each of the other 3 models explain less than 11% of the total variation of the change in vWF:Ag due to exercise .

The regression parameters for each of the four exercise marker variables studied in their respective models were essentially unchanged when regression was carried out using the level of vWF:Ag measured 30 minutes after exercise (Time 2) as the dependent variable (and adjusting for baseline vWF:Ag). Also, the results in each of the 4 models were not meaningfully different when the level of vWF:Ag measured immediately after exercise (Time 1) was used as the dependent variable, or when it was used to calculate the change in vWF:Ag which was then used as the dependent variable.

DISCUSSION:

Our results indicate that exercise has a strong effect of exercise on various coagulation parameters. We have shown that intense exercise can increase the absolute levels of vWF:Ag and FVIII:C by an average of 0.30 (37%) and 0.60 U/ml (56%), respectively, in healthy females. These responses were sustained even after 30 minutes of rest following exercise, although they were less marked at this time. A corresponding significant increase in the functional activity of these proteins in coagulation was also shown by virtue of the corresponding significant shortening of the CT and the aPTT.

However, these results were surprisingly lower than we had expected; 50% to 200% increments in both vWF:Ag and FVIII:C levels have been reported in most previous studies.[75,76] It is known that the magnitude of coagulation changes is related to both the intensity[100,101] and duration[96,98,99] of the exercise. Wetting of the forearm patch, which was the exercise endpoint used in our study, occurred for all individuals between 19 and 23 minutes of exercise. For at least 50% of participants, this endpoint corresponded to a submaximal heart rate of 90% to 100%, which may be one possible explanation for the lower changes in vWF:Ag and FVIII:C values observed in our study. It is also possible that exercising for a longer period of time at maximum heart rate could have increased the magnitude of the observed changes. However, we feel that compliance with exercise as a therapeutic modality would be seriously compromised if a more exhaustive protocol

were to be used. The protocol used in our study was well tolerated and all women achieved the prescribed end-points. This protocol, if proven effective to treat menorrhagia in women with Type I vWD, could easily be integrated into a woman's daily routine. Further, a stepping exercise can be done indoors, obviates the need to purchase expensive exercise equipment and is suitable for physically untrained individuals. It may also be possible that for reasons that are unclear, men respond differently to maximal exercise than do women. Women were included in only 4 small exercise studies (n = 10-14) that were insufficiently powered to assess gender differences.[79,102,106,112] The effect of intense exercise on menorrhagia in women with Type I vWD should be the focus of a future study.

Notwithstanding the lower than expected magnitude of the change in the level of vWF:Ag with exercise, we feel that our findings are clinically pertinent. The changes that were observed in this study group could theoretically be beneficial to women with Type I vWD in whom baseline levels of vWF:Ag and FVIII:C in the range of 0.30 to < 0.50 U/ml, respectively, are typical.[2] Since the lower normal limit in healthy individuals for both of these proteins is 0.50 U/ml[16], the 10% to 20% increments in vWF:Ag and FVIII:C, respectively, attained with exercise by the majority of subjects participating in this study should be sufficient to correct the aberrant laboratory coagulation profile in a large proportion of women with Type I vWD. Although this study was conducted using a sample of healthy women, we feel that these results are generalizable to women with Type I vWD who are known to have normal storage pools of both vWF:Ag and FVIII:C.[64] A complete correction of the vWF:Ag level and BT (bleeding time) has been reported following vigorous exercise in 3 women with Type I vWD.[79,110] Furthermore, responses similar to those in healthy individuals have been documented among individuals with Type I vWD who are treated with DDAVP[48,51,52,62,63] and adrenaline infusions,[86,89,90] both of which also work by causing a release of these proteins from their storage sites. The impact of a corrected coagulation profile on excessive blood loss during menstruation remains to be seen and is the main objective of Phase II of this study. Additional benefits of routine exercise include decreased menstrual pain[138] and

decreased risk of coronary heart disease,[139] which make regular exercise a particularly appealing treatment modality for women with vWD.

Although the mean increase in vWF:Ag reported in this study was 37%, there was great variability in the responses among study participants (ranging from 6% to 120%). We evaluated the extent to which the change in forearm sodium in sweat as well as 3 other blood markers of exercise intensity could predict the change in vWF:Ag due to exercise. We have shown that, of the 4 exercise marker variables in their respective models, the change in forearm sodium with exercise is a significant predictor of the change in vWF:Ag from baseline to 30 minutes after exercise. From the regression coefficient for a 1 unit change in sodium ($\beta = 0.052$), it can be predicted that an individual of a given fitness level with a baseline vWF:Ag level of 0.83 U/ml (which was the mean in this study group) would attain a level of 0.88 U/ml on average after exercise for every 1 mmol/l change in forearm sodium. The effect on the change in vWF:Ag, based on the range of values observed for the change in forearm sodium in the group, ranges from 0.04 to 0.44, corresponding to final vWF:Ag levels of 0.87 to 1.27 U/ml. Also, an average change in forearm sodium of 4.4 mmol/l, which is sufficient to increase the change in vWF:Ag by 1 SD, lies within the range of change in forearm sodium observed in this group (0.80-8.40 mmol/l). The reproducible findings for the effect of the change in sodium with different dependent variable definitions in regression modeling adds credibility to the importance of change in sodium as a predictor of the change in vWF:Ag with exercise. Only a larger sample size could lead to a more precise estimate of the effect of the change in forearm sodium on the change in vWF:Ag and should be addressed in a future study.

Little can be said about the importance of the other predictors of vWF:Ag in this study, namely, lactate, noradrenaline and adrenaline, since their accompanying 95% CI are wide and all include zero. The precision of the estimates observed in this study was low and is a reflection of our small sample size. Larger studies would be required to address the importance of these prognostic markers as predictors of the change in vWF:Ag.

Lactate, a known by-product of muscular exercise, was expected alone to account for, to a significant degree, the change in vWF:Ag. In our study, the mean change in lactate was 5.6 +/- 2.3. This mean lactate level is intermediate between those achieved by 5 male subjects in one study, when they performed a graded cycling exercise to maximum capacity (10 +/- 1.08) and when they cycled at a steady state low intensity (3.5 +/- 1.53),[100] and was associated with a mean change in vWF:Ag of intermediate magnitude. In our study, vWF:Ag and FVIII:C levels increased by 150% and 200%, respectively, following high intensity exercise but were unchanged after the low intensity exercise.[100] The change in lactate was not found to be a significant predictor of the change in vWF:Ag due to exercise in our study, at a significance level of 0.05. However, the fact that the 95% CI for this regression coefficient lay mostly above zero and its upper confidence limit was 0.06 suggests that it may be prove to be a clinically important predictor of the change in vWF:Ag with exercise in a larger study.

Although one possible explanation for the apparent insignificant effect of lactate is its imprecise estimate due to our small sample size, there are two other possible explanations. One possibility is that while lactate may have a dose-response linear relationship to vWF:Ag within a wide range of exercise intensities, it may lose its predictive capacity within the narrow range of exercise intensities imposed by our study protocol. Another possibility is that the production of lactate during intense exercise has a stepwise (or threshold) relationship rather than a linear dose-response relationship to the change in vWF:Ag. This latter possibility is supported by a study in which the aPTT was shortened significantly when 10 untrained males cycled to their respective lactate thresholds, whereas there was no change following exercise at 50% of their respective lactate thresholds.[119] If a stepwise relationship was indeed found to exist, then it could explain the weak linear relationship between lactate and the change in vWF:Ag that was observed in our study. Interestingly, the crude correlation between the change in forearm sodium and the change in lactate in our study was not significant. The absence of an association between these two variables suggests that perhaps they are not capturing the same element of exercise intensity, which may explain why they are differentially predictive of the change in vWF:Ag with exercise.

There are several reasons as to why we believe that noradrenaline and adrenaline are unlikely to prove to be important and/or useful predictors (despite our initial hypothesis) of the change in vWF:Ag, even in studies with larger sample sizes. First, their corresponding regression coefficients are of low magnitude even when accounting for their different units (i.e. expressing adrenaline and noradrenaline in terms of 100 and 1000 units of change, respectively, rather than 1 unit of change, as in the case of forearm sodium and lactate). Second, they explain but a very small portion of the variation of vWF:Ag change due to exercise, as demonstrated by the low squared partial correlation coefficients of noradrenaline and adrenaline, being 1.80% and 6.63%, respectively. Further, the mean magnitude of change required in either noradrenaline or adrenaline to induce a 1 SD change in vWF:Ag falls beyond their respective ranges of change observed in this study. Also, their respective standardized regression coefficients, which take into account their different units and measurement errors, suggest that these exercise markers are comparatively less important predictors of the change in vWF:Ag change than either sodium or lactate change. Another drawback to using these substances is that they require a blood test, are highly unstable and easily influenced by factors other than physical stress; mental stress is one such factor and may explain the elevated levels of noradrenaline at baseline in our study.[82] Although lactate levels, by comparison, are not influenced by mental stress, lactate shares the disadvantages with noradrenaline and adrenaline of requiring a blood test for measurement and of being highly unstable.

Among the 4 exercise markers, the fact that the change in sodium levels with exercise, as measured in the filter paper patches, was found to be a significant predictor of the change in vWF:Ag in our study, is more specific to physical exercise, and is clinically practical, stable, non-invasive and inexpensive, suggests that efforts should be directed towards refining the predictive capacity of sodium change in this setting, rather than the other exercise markers.

Arguments against the use of standardized regression coefficients have been raised for many reasons. The fact that standardized regression coefficients are difficult to interpret, do not represent true measures of the effect of changing the independent variable and are

sample-specific are among the most frequently cited criticisms.[136] Nonetheless, standardized coefficients can help assess the relative importance of the independent variables when they are measured in different units, since direct comparisons are otherwise difficult[140] and no definitive method for such comparisons exists.[141] Measuring the relative importance of the main independent variables by using standardized regression coefficients in this study was motivated by our interest to determine if measuring sodium in sweat could be a suitable replacement for other blood markers of exercise intensity. Recognizing the controversy surrounding the use of regression coefficients as a valid means to compare the relative importance of variables in regression, other methods to judge the value of sodium as a predictor of the change in vWF:Ag, as compared to the other markers, were also employed in our study. The relative magnitude of change in these variables required to bring about a 1 SD change in vWF:Ag (method proposed by Greenland)[141] and the comparatively higher partial correlation coefficient of sodium relative to the other exercise marker variables led to analogous conclusions. Most importantly, the greater clinical practicality of measuring sodium in sweat as opposed to measuring the other blood exercise markers was an important consideration.

Despite the moderate capacity of forearm sodium to predict the change in vWF:Ag due to exercise, it only explains 20% of the total variation in this change after adjusting for fitness level. It is probable that no one marker of exercise used in this study was able to completely capture the element of the exercise that is most predictive of the change in vWF:Ag. It is also possible that VO2 max, a more global measure of exercise intensity, would be a better predictor of the response to exercise. This measure, however, requires expensive equipment and trained personnel and is not suitable for testing on a large scale or for individual home use. The possibility that other unknown factors are responsible for the differential effect induced by exercise merits further study.

In this study, the relationship between each of the exercise markers and the change in vWF:Ag was not differentially affected by fitness level, age or baseline vWF:Ag. Although this observation may represent a true absence of interaction by these variables,

the small sample size used in our study restricted our ability to properly assess the presence or absence of effect modification. Despite this limitation in our study, the absence of effect modification by these variables has been supported in other studies. In a randomized trial conducted by El-Sayed *et al*, individuals who were subjected to a physical conditioning program did not display increments in vWF:Ag significantly different from individuals who were not subjected to the conditioning program, following a standardized graded exercise to maximum heart rate and/or exhaustion.[106] Further, Furgusson *et al* showed that there was no difference in the magnitude of coagulation changes between sedentary individuals, joggers and marathon runners following vigorous exercise to their maximum potential.[120] The finding that age does not modify the response to exercise is supported by the work of van den Burg *et al* who showed that increments in vWF:Ag after a standardized exercise were of similar magnitude among sedentary men of 20 to 30, 35 to 45 and 50 to 60 years old.[97] Baseline circulating levels of vWF:Ag were not expected to affect the change in vWF:Ag in response to exercise, which is in large part explained by the new release of vWF:Ag from its storage sites.[75] This expectation was based on the fact that individuals with mild vWD and hemophilia, who have decreased baseline levels of vWF:Ag and FVIII:C, respond normally to DDAVP,[48,51,52,62,63] adrenaline[86,87,89,90] and, in a few cases, to exercise.[79,110]

In our study, we assessed and controlled for the presence of confounding due to baseline fitness levels and age, which were identified as potential determinants of the change in vWF:Ag. Although the relationship between changes in forearm sodium and baseline fitness levels has not been explicitly studied, baseline fitness level is known to influence the metabolism of lactate[142,143] and adrenaline and noradrenaline[144,145] during exercise. Since fitness level is likely to be a strong determinant of the change in vWF:Ag due to its direct influence on exercise capacity (intensity and duration) and was also a strong confounder in our data, it was duly adjusted for in our analysis. However, residual confounding by fitness level, as it is defined in this study, remains a possibility.[146] For instance, it is possible that the broad categories of fitness level that were defined in the questionnaire we used were not sensitive enough to distinguish between true baseline differences in cardiorespiratory status. That being said, this less precise definition of

fitness level used in our study probably did not substantially bias our estimates of effect, since it has been shown in at least in one study, that defining fitness levels by more objective measures of cardiorespiratory status does not to affect the extent of coagulation changes induced by exercise.[106]

Age, on the other hand, was not adjusted for in the final models because it is not known to affect the change in the exercise markers, it is unlikely to influence the change in vWF:Ag with exercise to any significant extent[97] and was not an important confounder in our data. Since age range was restricted by our study design and may have controlled for the possibility of age as a confounding variable in our data, this observation may not be valid outside of the age range studied (18 to 40 years old).

Residual confounding due to other known confounders about which we did not have any information may bias the association we have observed. For example, mental stress is known to increase levels of noradrenaline and adrenaline[82] and is a known mechanism for the increase in vWF:Ag.[77] Although the precise relationship between mental stress and the changes in forearm sodium and lactate is unknown, it is unlikely to be very significant since these markers are generally more specific to physical rather than mental stress. Since mental stress is an important predictor of vWF:Ag change,[77] it is likely also an important confounder of all 4 associations between exercise markers and the change in vWF:Ag addressed in this study. The high mental stress of the participants before the exercise session may explain why the mean noradrenaline level after 30 minutes of rest was lower than the initial baseline mean (not shown) and may have independently contributed to the observed increase in vWF:Ag. Not adjusting for mental stress may have biased our estimates away from the null in the case of adrenaline and noradrenaline, and towards the null for lactate and forearm sodium (which are less closely correlated to mental stress).

Another limitation of this study relates to its generalizablilty. First, the effect of exercise on individuals of the black race, on females outside of the age range we studied and on women with high baseline fitness levels is not known and cannot be extrapolated from

our study results. Although exercise as a means of therapy is likely to be of little value in older individuals due to the potential risk of inducing myocardial infarction,[147] its effectiveness in children remains to be studied. Highly fit women were clearly underrepresented in this study (only 1 woman was in the highest fitness level category) and, therefore, it is not possible to generalize our results to this particular group of women.

In addition, the fact that individuals volunteered for this study implies that the study group was comprised of a select group of highly motivated young women to whom the prospect of exercise was appealing; consequently, the study group may not have been representative of all healthy women of childbearing age on a number of unmeasured factors.[148,149] Whether comparable effects of exercise can be demonstrated in a randomly selected group of women is not known.

Another obvious yet important limitation of this study is that it only addresses the efficacy of vigorous exercise on increasing levels of vWF:Ag and FVIII:C in women. This limitation is related to the fact that compliance with the exercise protocol was optimal due to the supervised setting. Participants made a particularly great effort to complete the protocol as intended. Whether women in an unsupervised environment would be equally as compliant is difficult to ascertain, and is an issue that is adressed in Phase II of this study.

A major distinction between a quasi-experiment and a true experiment is that the former typically does not employ a randomization process to assign the intervention or to select individuals for inclusion in the study. The internal validity of the findings in quasi-experiments is typically threatened for this reason.[150] The findings in our descriptive pre-test post-test study may be biased due to the existence of unknown confounders (related, for instance, to anthropometric values, high HDL cholesterol levels or the use of an oral contraceptive pill) that may have significantly affected the observed association between exercise and changes in the coagulation proteins. The lack of a control group in this particular study limits our ability to address this possibility. A random effect due to

clustering of an unknown determinant of the response to exercise among individuals in our single experimental group could have severely biased the observed association in this study. Nonetheless, such a bias, if any, is unlikely to explain the association in its entirety since this association has already been noted in the context of one other, albeit small, randomized trial[106] and is based on sound physiological data. Because all individuals recruited in this study were assigned to the same intervention without randomization, selection bias is probably minimal.[151]

Despite all of the foregoing limitations, the study design we used was a convenient way for us to test our hypothesis. The ease with which this design was implemented (due to the absence of a complicated randomization scheme) and the fact that there was a readily accessible group of young healthy female volunteers willing to participate in the study expedited recruitment and allowed for completion of this study in a timely fashion. Understandably, it is for this these reasons that the repeated measures quasi-experimental design is the most frequently used design in exercise physiology[152] and was used in all but one study (to the best of our knowledge) to assess the effect of exercise on coagulation.

The main drawback of quasi-experiments in general is they do not permit a strong or even reasonable causal inference of the intervention-outcome associations being evaluated.[150] Quasi-experiments are particularly susceptible to several non-experimental processes that can influence, and therefore serve as plausible alternative explanations for, the observed results, which in many cases are not possible for the investigator to exclude. This possibility is especially true in studies such as ours where there is no control group. Nonetheless, there are several features of our study that render it immune from such competing influences, and in so doing, permit reasonable causal inference.

In this study, possible alternative explanations for the results we observed relate to the effects of history, maturation and statistical regression to the mean.[150] The effect of history refers to environmental patterns that change over time (in either direction) and that can directly influence the outcomes that are observed in a study, irrespective of the

specific intervention applied. For example, an improvement in the underlying condition or severity of the illness being studied over time could serve as a viable explanation for the outcomes observed. In our study, the stability of circulating blood levels of vWF:Ag over time (that are genetically controlled) and the very short interlude between repeated tests offer strong arguments against the effect of history playing a paramount role. An analogous argument can be made regarding the effect of maturation (which refers to the systemic fluctuations that can be observed in any dependant variable), since levels of vWF:Ag do not undergo appreciable changes over time. While it is true that there is a cyclical variation in the levels of vWF:Ag during a menstrual cycle,[24] and under the influence of various medications[25] and medical conditions,[17,18] these factors cannot explain the immediate effects of exercise on the change in vWF:Ag. The effect of regression to the mean is to bring score levels of repeated tests towards the usual statistical mean. Particularly vulnerable to this effect are studies that include individuals whose pre-test scores are significantly higher or lower than the usual mean. However, given that all of the women in this study have normal circulating vWF:Ag levels, and therefore have scores that are close to the expected mean, this statistical phenomenon is unlikely to have a significant influence on our study results. In summary, the novel outcome that was measured, its known stability and the very-short pre-test post-test interval between repeated measures served to neutralize these alternative phenomena and therefore, strengthen the causal nature of the association between exercise and the increase in vWF:Ag observed in our study.

Causality of the exercise-vWF:Ag relationship in this study is ostensibly supported by the fact that certain features of our study satisfy criteria put forth by Sir Bradford Austin Hill,[153] although none of these features are sufficient in and of themselves to prove causation. The first of the criteria and arguably a compulsory feature for causality is the temporal relationship observed between exercise and the increment in vWF:Ag after exercise. Further, the very short interval between the intervention (exercise) and the result strongly support the notion that exercise was the cause of the increased levels of vWF:Ag. Additionally, the phenomenon we observed has reproduced what others have shown in the past. The mechanism for this phenomenon has been clearly elucidated and

is analogous to the effect of mental stress and DDAVP on vWF:Ag, and thus is biologically plausible. Although the response we observed is not specific to exercise, there are few factors other than mental stress and DDAVP therapy that are known to bring about such changes in vWF:Ag, particularly in such a short interval. Although mental stress as a contributing factor to the increase in vWF:Ag could not be ruled out in this particular study (especially in the absence of a control group), its contribution is likely to be minimal based on what has been previously reported in the literature.[77]

Despite all the limitations noted above, we believe that the present study gives compelling evidence indicating that performing an intense exercise to one's maximum heart rate in a supervised and structured protocol can significantly raise vWF:Ag and FVIII:C in healthy female volunteers. The accompanying shortening of the aPTT and the CT were in keeping with the degree of increase in FVIII:C and vWF:Ag, respectively. The fact that significant increments in both vWF:Ag and FVIII:C were achieved in the large majority of the participants included in our study, and that these changes were sustained even after a 30-minute rest period, underscores the clinical worthiness of intense exercise as a natural intervention. Moreover, the change in sodium level in sweat associated with exercise is a reasonably good predictor of the variation in change vWF:Ag observed in this study group. Furthermore, the easy implementation, low cost and excellent tolerability of this intervention make it worthy of further study in vWD. Finally, the sweat patch can also be used as a measure of compliance when exercise is being evaluated in women exercising in the home setting.

PHASE II

METHODS:

Study Design

Crossover

The design used for the feasibility study was a 4-period crossover randomized clinical trial. In this study design, there were 2 study arms; Arm A and Arm B. Each arm of the study represented a different treatment order sequence. Individuals in Arm A began with an "exercise" period and those in Arm B began with a "rest" period and each then alternated between "exercise" and "rest" for the remaining 3 consecutive study periods. Thus, individuals in Arm A of the study followed the treatment order sequence of "exercise", "rest", "exercise", "rest", whereas those in Arm B followed a "rest", "exercise", "rest", "exercise" treatment order sequence. The "exercise" or "rest" interventions applied only to days in which the participant was menstruating. A regular menstrual cycle is 28 days, with menstruation every 21 days. Seeing as only women with regular menstrual cycles were eligible to participate in this study, the crossover points in both study arms occur at regular intervals and are inherently time-dependent.

Randomization

A randomization scheme was used to allocate the study arm to each participant. Each consecutive candidate that satisfied the entry criteria was assigned a random number from a random digit permutation table.[154] Individuals with even numbers were assigned to Arm A, whereas individuals with odd numbers were assigned to Arm B. Study arm assignments were kept in sealed envelopes and stored in the serially numbered study kits. After assembly, each study kit remained closed and was not opened until it was received by the participant. Participants were instructed not to open the envelopes until they left the hospital.

Blinding

Blinding the subjects to the crossover point was not feasible given that exercise was the assigned intervention. The study investigator was, however, blinded to the study arm assignment, which was assigned by a clinical research nurse. The decision to blind the investigator in this feasibility study was done in order to reproduce the exact conditions planned for the definitive trial. The investigator will only be unblinded to the treatment assignment in the definitive trial upon the completion or early termination (for any reason) of the study. The investigator will tabulate all menstrual scores while still blinded.

Study Population

Subject Eligibility criteria

Pre-menopausal women between the ages of 18 and 40 with menorrhagia and Type I vWD were eligible for this study. The definition of *definite* Type I vWD that was used in our study was that of the ISTH (International Society of Thrombosis and Hemostasis) published in 1996, and includes *all* of the following 3 components: i) a significant mucocutaneous (mucous membrane) bleeding history, ii) laboratory tests compatible with Type I vWD and iii) either a positive family history of Type I vWD or the documentation of a vWF:Ag mutation. The laboratory criteria for the definition of *definite* Type I vWD (item ii above) requires that i) both vWF:Ag and vWF:RcoF (vWF:Ag activity) are < 2 SD below the population mean and ABO blood type adjusted mean on ≥ 2 determinations, and ii) ristocetin-induced platelet aggregation and von Willebrand multimer analysis, if performed, must not indicate abnormal sensitivity to low concentrations of ristocetin, or abnormal multimer distribution, respectively.[26] All patients registered in our Type I vWD registry had been diagnosed with vWD according to the foregoing criteria.

The presence of menorrhagia in women with Type I vWD was determined based on either a previously documented menstrual PBAC (pictorial blood assessment chart) score of > 100[36] or menstrual bleeding of sufficient severity (in volume and/or duration) to cause 1 or more of the following: i) significant impairment in quality of life (defined as interference with social, family or sexual activities, sleep or mood alterations, and/or inability to enjoy life), ii) work or school absenteeism on at least 2 occasions on account of severe bleeding, iii) the use of more than 8 sanitary pads or tampons per day on 3 or more consecutive days, iv) soiling of clothing or bedding requiring clothing or linen changes on at least 2 occasions, v) iron deficiency anemia, or vi) blood transfusion and/or hospitalization to treat uncontrollable hemorrhage. The features listed above are all associated with menorrhagia in women with Type I vWD.[27,28,29,34,36,41]

Eligible women were excluded if they had irregular menstrual cycles (absence of regular monthly periods), significant anemia, asthma or other pulmonary conditions, a personal or family history of heart disease, impaired balance, motor disability, history of exertional syncope (loss of consciousness with heavy exertion), an uncontrolled thyroid disorder or diabetes. Women who used steroids or sympathomimetic drugs, were pregnant (or had a positive pregnancy test taken 1 week prior to their participation in the study) and those who did not feel that they could comply with the prescribed exercise protocol for any reason were also excluded.

Subject Recruitment

Potential candidates for this study were initially identified from the computerized database of the Women's Bleeding Disorders Clinic of Sainte-Justine Hospital, in Montreal, with the help of a research nurse. Those identified were then contacted by telephone in order to review their eligibility and invite them to participate. Those individuals that were eligible and who agreed to participate in the study were then scheduled for a hospital visit. During the hospital visit, participants i) underwent a medical questionnaire and physical exam, ii) reviewed the details of the study protocol with the investigator, iii) signed a consent form, iv) were randomised to 1 of 2 study arms

(Arm A or Arm B as defined above) and v) were provided with a study kit containing the necessary materials for the study. Each study kit contained a cardiac monitor, a stepping music cassette (The Beat®, 1995), an aerobic step, a menstrual blood flow questionnaire, 4 PBACs, filter paper patches, sterile urine cups (for placement and transport of the collected filter papers) and alcohol swabs.

Compensation

A $200 honorarium was given to each participant upon successful completion of the study.

Study Protocol

A home stepping exercise protocol was prescribed to each study participant. Participants were required to exercise in the morning of each day of their menstrual period (if they were in the "exercise" period of the protocol) while wearing a self adhesive filter paper patch on the inner surface of their forearm. The exercise protocol required participants to perform a step exercise at a pace of 32 steps per minute (as set by the music rhythm) at close to maximum heart rate. Participants were instructed to assess their heart rate from the cardiac monitor display every 15 minutes during the exercise. If, at any time their maximum heart rate was surpassed, they were instructed to decrease the intensity of the exercise and to adjust their pace in order maintain their heart rate at close to their maximum. Approximate maximum heart rate estimates[155] were made by the investigator and were provided to each participant at the time of randomisation. Participants were instructed to continue the exercise until 1 of 2 end-points was achieved; either i) the participant became physically exhausted and was unable to continue, or ii) the forearm patch became visibly wet and a minimum of 20 minutes had elapsed. An absolute minimum of 20 minutes of exercise was required as an extra safeguard to ensure that an adequate exercise effort was made despite potentially higher ambient temperatures in the home that can independently increase sweating. This time corresponded to the shortest

duration of exercise performed by any single participant before the forearm patch became wet in our initial trial.

Participants were asked to complete a PBAC on each day of their menstrual period for all 4 consecutive menstrual periods (i.e. even when they were in the "rest" period of the protocol), as a means to quantify their menstrual blood flow. The PBAC used in this study is adapted from that used by Highams *et al*[36] and is included as Appendix B.

Participants were asked to apply the filter paper patch to their forearm before beginning the exercise and to remove it only once the exercise was completed. Once removed, the patches had to be placed separately into the pre-labelled sterile plastic containers provided in the study kit and then allowed to dry. Once the patches appeared to be dry, participants were asked to seal the containers. During the "rest" periods of the protocol, participants were asked to wear the forearm patch for 30 minutes on each morning of their menstrual period, during the same time that they would have otherwise performed exercise during the "exercise" period of the protocol. The sodium content measured in these patches applied during "rest" periods represents the baseline amount of sodium excreted by the forearm during the participants usual morning activity (non-exercise control patches). After having completed the study, participants were required to return all the patches and the PBACs to the investigator for analysis. Participants who successfully completed the study and returned all of the loaned study equipment were given the $200 honorarium. A clinical research nurse was available 24 hours a day for participant assistance.

Women who participated regularly in vigorous exercise were asked to refrain from doing so during, or within 24 hours prior to, their menstrual periods. Routine exercise, however, was permitted between menstrual periods since long-term physical conditioning is known not to affect the response to an acute exercise.[106,120] The women in the study were instructed to use, whenever possible, the same brand and absorbency of sanitary napkins and tampons for the duration of the study protocol so as to allow for meaningful comparisons between menstrual periods using the PBAC score.[37] Also, women who were

on the birth control pill, who smoked or drank coffee regularly were asked either stop prior to their commencement of the study or to continue throughout the entire study so that these variables would not differentially confound the treatment effect across periods of the protocol. The decision of each participant was documented. As well, participants who used antifibrinolytic agents routinely to control menstrual bleeding were asked to cease such usage prior to their participation in the study. The use of medications during the study was, however, permitted when it was deemed necessary by the investigator to control significant bleeding of any type. In such instances, the use of the particular agent either during a menstrual period or within 3 days of the onset of a menstrual period was clearly recorded. Based on the known short half-life and inhibitory effects of DDAVP,[48,50] Amicar and Cyklocapron,[156] the use of any of these agents more than 3 days before the onset of menstrual bleeding was unlikely to have a substantial effect on menstrual blood flow.

Compliance

The sodium content of the filter paper patch was used as an objective measure of compliance with the exercise protocol. An absolute post-exercise increase in sodium content of greater than or equal to 1.0 mmol/L was chosen to indicate compliance with the exercise protocol. This cutoff level was determined based on the results of our pre-test post-test trial and corresponded to the minimum post-exercise sodium levels observed among all participants that successfully completed the exercise protocol while under supervision.

Study Variables

The main outcome (dependent) variable for the definitive trial will be the PBAC score. The investigator, from the indictors filled in on the PBAC by the participant, will tabulate scores for both the "exercise" and "rest" periods for each participant. The scoring system originally developed by Highams *et al*[36] will be used to tabulate the scores. The PBAC score will be coded as a continuous variable. Compliance will be determined and coded

as a dichotomous variable (either 'yes' or 'no'). Sodium levels below 1.0 mmol/l in a single patch will indicate non-compliance with the corresponding exercise period, whereas scores greater than or equal to 1.0 mmol/l will indicate compliance. Compliant periods will be defined as exercise periods in which at least 90% of the patches worn during that menstrual period indicate sodium levels of greater than or equal to the cutoff level of 1.0 mmol/l.

RESULTS:

Eligibility and Participation

Two hundred and six individuals with vWD were identified from our hospital clinic database. After excluding individuals with non-Type I vWD, males, and individuals outside of the age range being considered in this study, 65 individuals were identified as potentially eligible for enrollment in this study. As can be seen from the enrolment algorithm in Figure 5, 43 individuals were contacted by telephone. An additional 14 individuals were excluded based on information acquired by the telephone contact. The reasons for these exclusions and their relative frequencies are presented in Table 17. The most common reasons for exclusion were absence of menorrhagia or regular menstrual cycles, which together accounted for 72% of all exclusions. Of the 29 eligible individuals, 14 declined and 15 accepted to participate, corresponding to a refusal rate of 48%. Of the 15 individuals who accepted, 13 have already been randomized in our study. The remaining 2 individuals will be randomized shortly. The reasons for refusal to participate and the percentage of individuals who refused for each reason are shown in Table 18. As can be seen from Table 18, the 3 most common reasons for refusal to participate were insufficient time, dislike of exercise and satisfaction with current treatment, which corresponded to frequencies of 35.7%, 21.4% and 14.3%, respectively. A single individual refused to participate because she felt certain that doing exercise during her menstrual periods would not affect her menstrual blood flow, based on her past experience.

Recruitment and Withdrawals

At the time of the preparation of this thesis, 13 individuals had already been randomized in this study. Recruitment took place over a 12-month period. As can be seen from the enrollment algorithm in Figure 5, 6 individuals have already completed the study and 3 others have withdrawn themselves from the study, corresponding to a withdrawal rate of 23%. Two participants withdrew from the study prior to commencing the exercise protocol due to time constraints. Interestingly, these 2 individuals were somewhat reluctant at first to participate for this very reason but ultimately agreed despite their initial reservation. The third participant withdrew after having completed 1 period of the protocol because she was intolerant to the intense exercise. Accordingly, at the present time, 5 individuals remain in the study at various phases of the protocol.

Tolerability and Compliance

The prescribed exercise was well tolerated by all 6 individuals who completed the study. All 6 individuals reported having achieved 1 of the 2 predetermined endpoints during each exercise session. None of the participants experienced difficulty in filling out the PBAC or following the instructions for the exercise protocol. This finding was further corroborated by the absence of telephone calls from participants to our clinical research nurse regarding such issues during the course of the study. There were no missing PBACs for any of the participants and all forms were filled out legibly. One participant, however, indicted the use of several different tampons and/or sanitary napkins during a single menstrual period although this pattern was consistent for all 4 menstrual periods. None of the participants reported having required treatment for bleeding during her menstrual periods, or within the 3 days prior to menstruation, or placed a telephone call to the investigator for that reason. All filter paper patches were returned and the number of patches for each individual corresponded exactly to the number of days of menstruation recorded on the PBAC. Using a cutoff level of greater than or equal to 1.0 mmol/L of sodium in the patches, we will be able to determine the compliance rate with the exercise protocol during exercise periods in the definitive trial. Further, we will be

able to determine the percent agreement between the number of patches that had a sodium level of equal to or greater than 1.0 mmol/l and the number of menstruation days indicated on the PBAC for the exercise periods. Taken together, these determinations will allow us to establish the compliance rate based on subjective and objective criteria. The assessment of compliance, based on the sodium levels in the patches and the number of menstruation days, will be made the research nurse in order to maintain blinding of the investigator, since the sodium quantities contained in the patches are likely to be informative of the type of period ("exercise" or "rest") in question.

DISCUSSION:

There are several issues relating to the recruitment, withdrawal rate and acceptability of our feasibility study that merit discussion. From the eligibility rate of 67% in our study (29 out of 43 individuals contacted by telephone), we can anticipate that 44 of the 65 individuals identified in our clinic database will be eligible for inclusion in our study. This number represents 21% of the total 206 individuals identified from the database. With an overall eligibility rate of 21% and a participation rate of 52%, we can expect to recruit 23 subjects for our study from our institution alone. A few additional patients for the study can be expected to participate from among the new diagnoses made in our outpatient clinic (i.e. approximately 2 per month). It is possible that recruitment for this study will need to be extended to other treatment centers in order to complete our study with sufficient power and within a reasonable time frame. Recruitment for this study was extremely slow, with only 13 individuals recruited in 1 year. The extent of the contribution that we will require from other treatment centers will depend on the estimated sample size for the formal trial. In the event that recruitment from other treatment center becomes necessary, appropriate stratified analyses according to treatment center will be employed.

Although there was only a small number of subjects in this feasibility study, we experienced a relatively high withdrawal rate of 23% (i.e. 3 out of 13 individuals). In only 1 of the 3 cases, the withdrawal was related to the exercise protocol, whereas the other 2 participants withdrew before they commenced the first period of the protocol. This finding was of great concern to us since a high withdrawal rate is especially costly in a crossover trial, where each individual contributes an important proportion of the overall information to the analysis.[157] This observation is especially true when sample size estimates are based on multi-period crossovers, where the contribution of any given individual to the analysis is proportionately greater than in a 2-period design.[158] A high withdrawal rate in a crossover trial results in a significant loss of statistical precision of the estimates generated in the trial, which can be comparatively large relative to that in a parallel group design for an equal number of withdrawals.[157] Failing to keep withdrawals below an acceptable level would defeat one of the principal advantages of the crossover design. The best way to circumvent this problem is to implement precautionary measures to prevent withdrawals whenever possible. The best precautionary measure to minimize this problem in our definitive trial will be to better select individuals for the trial, that is, to ensure that only those individuals who are firmly committed to completing the protocol are randomized. Despite undertaking necessary precautions, a certain percentage of withdrawals is unavoidable. The estimated sample size[137] for the definitive trial will be increased appropriately in order to accommodate withdrawals.[159,160]

As was expected from the successful implementation of an analogous exercise program in our pre-test post-test trial, the exercise program was well tolerated by all individuals who completed this study. Exercise during heavy menstrual periods was deemed acceptable by all of the women who completed the study and none of the participants experienced difficulties adhering to the protocol.

Compliance with our extended 4-period protocol, based on subjective reports and the number of complete PBACs returned by the participants, was 100%. In recognition of the limitation of these methods to determine compliance, a more objective assessment of compliance with exercise, using the sodium content measurements from the filter paper

patches collected from each participant, is planned for the definitive trial. Further, the extent to which subjective and objective measures of compliance are consistent with one another will be determined when the sodium levels are available from the definitive trial. Based on the high compliance rate in our feasibility study and the fact that this exercise was simple to perform without instruction, easily integrated by most participants into their daily routine and did not require the purchase of expensive exercise equipment for home use, we have decided to use the same exercise protocol for the definitive trial.

The rationale for employing a crossover design to study the effect of exercise on menorrhagia in women with Type I vWD merits a discussion of its advantages and disadvantages compared to a group parallel design for the same purpose.

A major advantage of having used a crossover design to study the impact of exercise on menorrhagia is that we were able to use fewer subjects than we would have otherwise needed for a parallel group design. The relative gain in power afforded by a crossover study is related to the fact that a comparison of treatments on the same patient is expected to be more precise than a comparison of treatments on different patients, as a result of the decreased variation in repeated measurements taken on the same individuals, as compared to measurements taken on different individuals.[137,159,160] There is high inter-patient variability in the amount of menstrual blood loss, which would inflate the required sample size needed to show equivalent differences between regimens in a standard randomized group parallel design.[137,157] In contrast, however, the intra-patient variability in the amount of blood loss in 2 sequential menstrual cycles using the PBAC score is considered to be low.[37] As in the case of menstrual flow, where there is high inter-subject variability relative to intra-subject variability, Brown has shown that a crossover design has a high cost efficiency relative to the parallel design for any given ratio of "cost of treatment per patient per period to cost of recruitment per patient".[161] For reasons of cost efficiency and in view of the limited number of potential eligible candidates with Type I vWD available for this study, the crossover design was especially appealing.

Another important advantage of using a crossover study is that it virtually eliminates the influence of fixed patient-related characteristics (both known, such as age and fitness level, and unknown) on the relationship between variables being studied, because each participant serves as her own control.[157] It does not, however, control for patient-related confounders that change over time. An attempt to control for potential confounders such as oral contraceptive use, smoking and caffeine intake was made by insisting on consistent use of or avoidance of these agents throughout the entire 4-month study period. The relatively short 4-month study period, however, makes it unlikely that other unmeasured or unknown time-dependant confounders will be particularly problematic. These two important advantages collectively justified our choice of using the crossover design for our study.

A major theoretical disadvantage to using a crossover design is the risk of introducing order effects, which include carryover and period effects that may contaminate the results.[157,160] A properly designed crossover trial should have an adequate washout period (time in which the potential influence of the intervention wanes before the second intervention begins) in order to prevent carryover effects from one intervention period to the next. Our study protocol meets this requirement since the short half-lives of 8 to 12 hours of both vWF:Ag and FVIII:C released during exercise, ensure that these effects do not carry over until the individual's subsequent menstrual period. When these same individuals cross back from a "rest" period back to an "exercise" period, there is by definition no carryover effect since there is no effect induced by rest. The "rest" period is equivalent to the placebo period used when studying the effect of a single drug in a crossover study. Further, the normal menstrual cycle of 28 days provides a natural washout period of any unknown or unmeasured effects that may have been induced by exercise.

An evaluation of the effect of exercise on menorrhagia was also feasible in this study design, seeing as exercise is not known to cause any irreversible changes to menorrhagia (or its severity) or to cure it definitively. Irreversible changes to menorrhagia induced by exercise would preclude the use of a crossover design.[159] Rather, the transient nature of a

response, if any, allows for comparisons between "exercise" and "rest" periods. Furthermore, there is no evidence in the literature to suggest that patterns of excessive menstrual bleeding fluctuate, remit or regress over time. The one exception to the stable nature of this condition is that patterns of menstrual bleeding tend to change, sometimes definitively, after pregnancy or after use of the contraceptive pill. Since pregnancy was an exclusion criterion in this study, period effects related to changes in menstrual flow unrelated to exercise over the 4-month follow-up period were not anticipated. Likewise, since oral contraceptive use remained consistent throughout the study for each individual, its use would not explain any changes in menstrual flow observed in the study. Treatment by period interactions is unlikely to arise in trials with short treatment regimens under stable treatment conditions.[158] Further, the use of the 28 day menstrual cycle as the "time-dependent" crossover in our trial, rather than a "response dependent" one, is also a useful way to minimize the influence of order effects and allows for unbiased and easy interpretation of the results.[157] For these reasons, studying the effect of exercise on menorrhagia lends itself nicely to the crossover study design.

The power of a crossover study and the validity of its findings may be compromised when the presence of order effects is comparable to or more significant than the presence of treatment effects.[157] In such a situation, strong preference should be given to a parallel design in which these influences are not pertinent. Based on the existing knowledge of the effects of exercise and the strong clinical knowledge about vWD and its manifestations, we were reasonably assured of the absence of these effects in our trial. This factor weighed heavily on our decision to use the crossover design since according to Brown, estimating and adjusting for order effects requires a sample size greater than that which is needed for a parallel group design,[161] and a greater sample size would not be feasible in our study given the limited number of potential candidates with vWD.

Overall, we believe that the crossover design is an efficient way to address our study hypothesis, particularly in view of the inherent stability of vWD, the existing literature validating the absence of carryover effects and the acceptable withdrawal rate. Moreover,

inclusion of individuals who are committed to do the study should ensure that withdrawal rates in the study are not excessive.

The choice between a 4-period crossover trial as opposed to the standard 2-period crossover trial was a difficult one. With a limited pool of potential candidates, our intention was to attempt to maximize the data gathered from each individual by adding additional periods to the protocol, while still maintaining a reasonable total duration of the trial.[158] In addition, we felt that using this strategy in our definitive trial will allow us to limit the recruitment of subjects to our treatment center. However, the trade-off for extending treatment periods is a potentially higher withdrawal rate and the need for more complicated analyses of the trial data.[158] As the number of periods is increased, we suspect that the high rate of subject withdrawal would become an increasing threat. Considering that a withdrawal rate any greater than that which we observed in our feasibility study would undermine the utility of the crossover trial design, we felt that changing to the 2-period design in the definitive trial will help to minimize this risk. In addition, we speculate that the extended 4-period trial may have adversely affected recruitment, which was especially problematic in this feasibility study. It is possible that individuals who are limited by time schedules (as was the case in 36% of the eligible women we canvassed who refused to participate in this study) would be more willing to commit to a 2-month study than to a 4-month one, which should facilitate the recruitment process. Limiting the study to 2 periods will somewhat simplify the statistical analysis of the results, by obviating the need to consider the longitudinal nature (correlations, for example) of patient responses to the intervention and by keeping the number of withdrawals to a minimum, will allow for a more efficient analysis.[158] Notwithstanding the fact that there are certain advantages of the 4-period crossover design and that some multi-period designs can be equally if not more efficient than the 2-period designs, in our opinion, these advantages are largely outweighed by the potential problems of recruitment and withdrawals that we might experience, as well as the added statistical complexity with which we would be faced. Accordingly, serious consideration will be given to using a 2-period crossover design in our definitive trial.

Randomization of the treatment order was considered essential to our study, in order to ensure the complete absence of order-related bias, in case the menstrual score outcome was found to be sensitive to the treatment order. When the treatment order is not randomly assigned, results may be particularly biased.[157] Consider, for example, a hypothetical scenario in which all women begin with an "exercise" period. It is conceivable in this circumstance, that if exercise were to have a true or perceived positive effect on menstrual blood loss, then this period would be perceived as the baseline flow to which the subsequent "rest" period would be compared. In such a situation, menses during "rest" periods may be perceived to be worse than they actually are. Even though the PBAC, a semi-quantitative tool, is being used to measure outcome in this study, it is still possible that participants may complete the PBAC according to their perceived changes in menstrual flow. For instance, women may assign a lesser score for a certain degree of sanitary towel soakedness, or not consider as menstrual days the last days of the menstrual cycle when losses are minimal. In the opposite scenario, if all women begin with a "rest" period, which should in theory represent their baseline heavy menstrual flow, exercise may be perceived as having a benefit even in the absence of a true response. Therefore, we felt that the use of random order treatment assignments between the 2 study groups at least controlled, to some extent, the possibility of order-related bias.

It was not possible to conceal the crossover point in our study from participants since the intervention in the trial, namely exercise, is not amenable to blinding. The fact that exercise was being investigated as a therapy for menorrhagia was also not concealed from participating women. We are aware that knowledge of the crossover point can influence the treatment response or assessment.[157] For instance, anticipated or perceived changes (positive or negative) of the effect of exercise on a participant's menstrual flow may inadvertently influence her menstrual score assignment leading to a misclassification bias of outcome. A tendency to underestimate blood loss during "exercise" periods as compared to "rest" periods is possible, for example, among women who anticipate that exercise will have a positive effect. Methods to ensure the blinding of the investigator to the treatment order sequence were successfully implemented by blinding the investigator to the sodium levels in the sweat patches (which will be carried through until after the

results are analyzed in the definitive trial) and by having the treatment assignments made by the research nurse. This approach will prevent the possibility of misclassification bias introduced by the investigator in the tabulation of the menstrual scores assigned by participants in the definitive trial. Although the menstrual score is a relatively objective measure of outcome, the knowledge of the treatment assignment could influence the classification of ambiguous and/or inconsistent scores by the investigator. If the misclassification of outcome introduced by either the patient or the investigator is non-differential, bias in the menstrual score change will be towards the null hypothesis, or towards no change. If, however, the misclassification of outcome is non-differential, biases away from the truth in either direction are possible. This bias could result in overlooking the beneficial effect of exercise or concluding falsely that one was there when it is not. In either case, wrong conclusions regarding the value of exercise as an intervention for menorrhagia would be reached. Removal of investigator bias in the allocation of subjects to different interventions is achieved by blinding the investigator to the treatment assignment in a parallel group design.[162] However, this feature is less pertinent to the crossover design, where all participants receive identical treatment, albeit in a different randomly assigned order.

Our feasibility study had several other limitations that are worthy to note. The first relates to the possibility of bias that is the result of self-selection. For instance, it is possible that, inadvertently, individuals who already had some prior personal opinion regarding the effect of exercise on their menstrual periods chose to participate in the trial. It is also possible that individuals who exercise routinely may have already noticed some effect, either positive or negative, which may have also influenced their decision to participate in the study. One participant in our feasibility study refused to participate for this reason. The question was not explicitly asked to candidates who consented to participate in the trial, although 1 participant was convinced of the beneficial effects of exercise on her menstrual flow prior to enrollment. If a large proportion of such individuals that experienced prior positive effects of exercise participated in the trial, then the measured effect of exercise on menstrual blood loss may be biased away from the null hypothesis of no exercise effect. Therefore, motivational or behavioral characteristics of the

individuals that choose to participate may also have a positive effect on the outcome being measured.[163] Although self-selection bias remains a possibility, it is unlikely to be substantial in view of the fact that our exercise hypothesis is not well known, and therefore, is unlikely to have influenced a large proportion of participants prior to their enrollment.

The number of participants that withdrew from the study was 2 out of 13, which corresponds to a rate of 23%. If the withdrawals were in some way systematically related to the response to the exercise intervention, selection bias may threaten the validity of our results.[157,164] If, for example, those individuals who found no improvement during the first period chose not to continue, the results may be biased in favor of showing a positive effect when in fact there was none. Although this possibility presents a possible concern for the definitive trial, we anticipate that exercise-related withdrawals will be minimal. Our speculation is based on the fact that since exercise is a physiological intervention, rather than a pharmacological one, withdrawal due to intolerance of the intervention or from side effects is not expected.[157] As well, this expectation is supported by the observation that the exercise session itself was well tolerated by all 40 participants in the pre-test post-test trial, despite the fact that approximately 50% of the women were inactive/sedentary at baseline. In addition, there has been only a single withdrawal so far that is directly related to the exercise protocol of the feasibility study.

The possibility of bias introduced by the selective withdrawals that do occur will be assessed for in our final analysis of the results. An intention-to-treat analysis that includes those women known to have completed at least 1 period of the study protocol and who have completed at least 1 PBAC score, as well as a secondary per-protocol evaluation that includes those women who completed all 4 periods of the study protocol, will be carried out and compared. The intention-to-treat analysis, by including all data collected even from withdrawals, provides a greater accountability of all patients enrolled in the study, by decreasing the influence of withdrawals and non-compliers on the results, and thus allowing for greater generalizabiltiy.[165] Individuals who are randomized, and who have not completed at least 1 period of the trial, will be omitted from the analysis. This

strategy is reasonable and can increase the precision of the estimates derived during the trial. Moreover, this strategy can be done without introducing significant bias, provided, however, that the withdrawals are unrelated to the likelihood of having received the intended treatment (or control).[166]

Another concern raised by our study design is the possibility of contaminating intervention periods with rest periods and visa versa. This problem is a realistic one, since those who experience improvement in their menstrual flow with exercise in the first period may opt to continue exercising during subsequent periods. Although this tendency would violate the study protocol, this possibility is difficult to control. Although sodium measured in the patches should theoretically support or refute this possibility, such would only be the case if the patches were actually worn during exercise. If the patch is worn at any time other than during exercise, the sodium content of the patch will appear to correctly reflect a "rest" period in the protocol, despite the fact that exercise was done. The opposite situation in which women who find no improvement of their menstrual periods with exercise treat the remaining periods as "rest" periods is perhaps easier to control for. Sodium levels lower than 1.0 mm/l on days corresponding to exercise periods of the protocol would indicate that exercise was not done. In both of these situations, biased estimates of the effect of exercise would result. Since the ultimate analysis in our definitive trial will be based on some form of comparison between "exercise" and "rest" periods,[159,160] the possibility of either situation occurring would likely underestimate the effect of exercise (i.e. would bias the estimates toward the null hypothesis of no effect), since there would be little change in the intra-individual comparisons between treatment periods.

The PBAC score has been validated by Highams *et al*[36] We have chosen to use this tool in our study as a semi-quantitative measure of menstrual blood loss. It has also been shown that a score of > 100 has a high sensitivity and specificity to diagnose menorrhagia as compared to the gold standard hematin technique.[36,37] Despite this finding, however, the utility of the PBAC as a screening tool for menorrhagia is limited by the fact that its high sensitivity and specificity have only been established in women who use the same

brand of sanitary napkins or tampons.[36,37] Its ability to discriminate menorrhagia from normal menstrual blood loss among women who use different brands of sanitary napkins has not been studied. However, despite this limitation, both Higham and Janssen have shown that blood loss is highly correlated with the PBAC scores (Pearson correlation coefficient of 0.85).[36,37] This correlation coefficient measures the linear nature of the relation between menstrual blood loss and PBAC scores. Such a strong coefficient suggests that, in general, higher scores represent greater degrees of blood loss, whereas lower scores represent lesser degrees of blood loss. Moreover, this correlation coefficient suggests that the PBAC score can be used to assess the effect of treatment on menstrual blood loss, regardless of the baseline score of the particular individual. Therefore, since each woman will serve as her own control, the actual brand of tampon or pad used, as long as it was not changed for the duration of the study (or if at least consistent changes were made by each participant), should not influence the utility of this scoring system to compare menstrual flow between cycles in the same individual. Janssen *et al* also showed that there is consistency of individual menstrual blood loss and PBAC scores between 2 consecutive menstrual cycles with the median of the menstrual scores on 2 consecutive menstrual cycles of 96 and 85, respectively.[37] In their study, using a cutoff of 185 to define menorrhagia (highest sensitivity and specificity), 85.1% of women showed consistency of assessment of blood loss between 2 consecutive menstrual cycles using the PBAC score as the outcome.[37] This result suggests that differences in menstrual flow assessed by PBAC scores, both with and without exercise, can be attributed to true changes in menstrual blood volume and not to baseline variations in the scores of that particular individual. It is for this reason that we considered the PBAC score to be a valid means to assess the effect of exercise on menstrual blood flow in our study. This outcome tool has been used in order to evaluate the effect of intranasal DDAVP on menorrhagia in several other studies,[68,69,71] including a recently published crossover study.[70]

In our study protocol, exercise is being done in the home setting, where neither patient compliance nor the adequacy of the exercise will be monitored. This design feature will permit us to assess the efficiency of exercise (rather than efficacy) as a therapeutic modality to treat menorrhagia in women with Type I vWD. In general, the effect of non-

compliance results in a bias the trial towards finding no benefit to treatment and leads to an erroneous failure to reject the null hypothesis.[163] Although the compliance rate is of interest to us insofar as the interpretability of the findings in the definitive trial are concerned, even non-compliant periods will be included in the efficiency analysis.

In the definitive trial, we will quantify the degree of compliance by using the sodium content of the filter paper patches. The cutoff that was used to discriminate between compliers and non-compliers with the exercise protocol was determined from data obtained during the pre-test post-test trial. Although all participants in Phase I of this study displayed a post-exercise sodium level of at least 1.0 mmol/l, it is possible that there are individuals for whom this increment will not be achieved during Phase II of the study, even among those who comply with the exercise protocol. This limitation is inevitable seeing as we are using an arbitrary cutoff to determine compliance in Phase II since a threshold criterion for sodium levels to distinguish compliers from non-compliers has not yet been published. Therefore, our compliance estimates will only be considered as rough estimates.

The effect of a supervised exercise session on the laboratory coagulation parameters in women with Type I vWD was not assessed prior to Phase II of this study, although we feel that these findings would be generalizable to women with vWD. Also, blood levels of vWF:Ag are not being measured in Phase II of the study. It is possible that women completing the exercise protocol at home will not achieve blood levels of vWF:Ag as high as those noted in our Phase I trial. However, we do not believe that this possibility would be major drawback since our principal interest will be to assess the efficiency of the prescribed exercise on menstrual blood flow, which we believe to be the most clinically pertinent end-point. The menstrual blood loss outcome used in our study is a global reflection of the balance between the competing coagulant and fibrinolytic changes induced by exercise and, as such, is more clinically relevant than changes in laboratory parameters.

A similar concern was raised in the interpretation of a recently published clinical crossover trial designed to assess the effect of DDAVP therapy on menstrual bleeding in women with vWD.[70] In this trial, the possibility of underdosing was considered since blood levels of the medication were not assessed and administration of the drug was not supervised. Although DDAVP has been shown in previous uncontrolled non-randomized studies to be efficacious to treat this condition,[68,69,71] it was not shown to be effective in this trial.

The results of the definitive trial will not be generalizable to all women with Type I vWD. Highly motivated women, in general, are more likely to participate in studies and are more likely to comply with the study protocol; accordingly, they likely differ significantly from women who do not participate in many other unmeasurable or unquantifiable behavioral and motivational characteristics.[149,163] This limitation in generalizability is acceptable to us, however, in view of the fact that if exercise ultimately proves to be effective, it will only be employed by a very highly select group of women. Further, generalizing the results from our definitive trial to women outside of the age range being considered will not be possible.

OVERALL CONCLUSIONS:

Non-severe forms of vWD affect 1 in 100 persons, making it the single most common inherited bleeding disorder. Menorrhagia is a significant problem among women with Type I von Willebrand disease. However, little focus has been placed on studying natural, low cost and safe treatment alternatives for this condition. The aim of our study was to investigate the effect of exercise on measures of the coagulation system in healthy pre-menopausal women and to assess the feasibility of a 4-period crossover study in order to study the effect of exercise on menorrhagia in women with Type I vWD. The results of our Phase I quasi-experiment showed that circulating levels of vWF:Ag and FVIII:C increase and the bleeding time and clotting times shorten significantly when women perform vigorous exercise. We have also shown that the sodium produced in sweat during exercise is a relatively good predictor of the magnitude of the change in vWF:Ag induced by exercise. It stands to reason, then, that intense exercise should moderate menstrual blood flow in affected women.

The protocol used in our Phase II feasibility study was well tolerated and compliance with the exercise protocol was high. Recruitment, however, was particularly slow. The limited pool of potentially eligible subjects and the lengthy 4-period study design may have adversely affected recruitment. The withdrawal rate was 23%. However, reducing the crossover study to 2 periods and including patients from other treatment centers should accelerate recruitment for our definitive trial. Also, better selection of candidates will help keep the withdrawal rate acceptable. The results of the definitive crossover trial will determine whether exercise is an efficient physiological alternative to standard therapies to treat of menorrhagia in highly motivated women with vWD.

REFERENCES:

[1] Rodeghiero F, Castaman G, Dini E. Epidemiological investigation of the prevalence of von Willebrand's disease. Blood 1987;69:454-9.

[2] Rodeghiero F, Castaman G. Congenital von Willebrand disease type I: definition, phenotypes, clinical and laboratory assessment. Best Practice & Res Clin Haematol 2001;14:321-35.

[3] Ruggeri ZM. Structure and function of von Willebrand factor. Thromb Haemost 1999;82:576-84.

[4] Manusca DJ, Tuley EA, Westfield LA *et al.* Structure of the gene for human von Willebrand factor. J Bio Chem 1989;264:19514-27.

[5] Federici AB, Mannucci PM. Diagnosis and treatment of von Willebrand disease. Haemophilia 1999;5(S2):28-37.

[6] Federici AB, Mannucci PM, Castaman G *et al.* The 20-year (1978-98) natural history of von Willebrand disease in Italy: a multicentre retrospective analysis on diagnosis and therapy in 1234 patients. Haemophilia 2000;6:215 (abstract).

[7] Sadler JE, Mannucci PM, Berntorp *et al* Impact, diagnosis and treatment of von Willebrand disease. Thromb Haemost 2000;84:160-74.

[8] Lee CA. Women and inherited bleeding disorders: menstrual issues. Sem Haematol 1999;36 (S):21-7.

[9] Dilley A, Crudder S. von Willebrand disease in women: the need for recognition and understanding. J Women's Health Gender Based Med 1999;8:443-5.

[10] Sadler JE for the Subcommittee on von Willebrand factor of the Scientific and Standardisation Committee of the International Society on Thrombosis and Haemostasis. A revised classification of von Willebrand disease. Thromb Haemost 1994;71:520-5.

[11] Mammen EF, Comp PC, Gosselin R *et al.* PFA-100 system: a new method for assessment of platelet dysfunction. Sem Thromb Hemostas 1998; 24:195-202.

[12] Fressinaud E, Veyradier A, Truchaud F, Boyer-Neumann C, Trossaert M, Meyer D. Screening for von Willebrand disease with a new analyser using high shear stress: a study of 60 cases. Blood 1998; 91:1325-31.

[13] Cattaneo M, Federici AB, Lecchi A *et al.* Evaluation of the PFA-100 system in the diagnosis and therapeutic monitoring of patients with von Willebrand disease. Thromb Haemost 1999; 82:35-9.

[14] Favaloro EJ, Facey D, Henniken A. Screening for von Willebrand Disease: use of a novel platelet function analyzer (PFA-100) with high sensitivity to disturbances in von Willebrand factor. Am J Hematol 1999;62:165-74.

[15] Kerenyi A, Schlammadinger A, Ajzner E *et al.* Comparison of PFA-100 closure time and template bleeding time of patients with inherited disorders causing defective platelet function. Thromb Res 1999;96:487-92.

[16] Favaloro EJ. Laboratory assessment as a critical component of the appropriate diagnosis and subclassification of von Willebrand's disease. Blood Reviews 1999;13:185-204.

[17] Blomback M, Eneroth P, Andersson O, Anvret M. On laboratory problems in diagnosing mild von Willebrand's disease. Am J Hematol 1992;40:117-20.

[18] Mohlke KL and Ginsburg D. Von Willebrand's disease and quantitative variation in von Willebrand factor. J Lab Clin Med 1997;130:252-61.

[19] Gill JC, Endres-Brooks J, Bauer PJ, Marks WJ, Montgomery RR. The effect of ABO blood group on the diagnosis of von Willebrand's disease. Blood 1987; 60:1691-95.

[20] Blomback M, Eneroth P, Landgren B-M, Lagerstrom M, Anderson O. On the individual and gender variability of hemostatic components. Thromb Haemost 1992;67:70-75.

[21] Conlan MG, Folsom AR, Finch A, *et al.* Associations of FVIII and von Willebrand factor with age, race, sex, and risk factors for atherosclerosis. Thromb Haemost 1993;70:380-85.

[22] Shima M, Fujimura Y, Nishiyama T. ABO blood group genotype and plasma von Willebrand factor in normal individuals. Vox Sang 1995;68:236-40.

[23] Blann AD, Daly RJ, Amiral J. The influence of age, gender, and ABO blood group on soluble endothelial cell markers and adhesion molecules. Br J Haematol 1996;92: 498-500.

[24] Andersson O, Blomback M, Bremme K, Wramsby H. Prediction of changes in levels of hemostatic variables during natural menstrual cycle and ovarian hyperstimulation. Thromb Haemostas 1997;77: 901-4.

[25] Kadir RA, Economides DL, Sabin CA, Owens D, Lee CA. Variations in coagulation factors in women: effect of age, ethnicity, menstrual cycle and combined oral contraceptive. Thromb Haemost 1999;82:1456-61.

[26] 1996 Annual Report of the SSC/ISTH Subcommittee on vWF.

[27] Ragni MV, Bontempo FA, Hassett AC. von Willebrand disease and bleeding in women. Haemophilia 1999;5:313-7.

[28] Kadir RA, Sabin CA, Pollard D, Lee CA, Economides DL. Quality of life during menstruation in patients with inherited bleeding disorders. Haemophilia 1998;4:836-41.

[29] Kouides PA, Phatak PD, Burkart P *et al.* Gynaecological and obstetrical morbidity in women with Type I von Willebrand disease: results of a patient survey. Haemophilia 2000;6:643-8.

[30] Hallberg I, Nilsson L. Menstrual blood loss-a population study. Acta Obstet Gynaecol Scand 1966;45:320-51.

[31] Rees MCP, Chimbira TH, Anderson ABM, Turnbull AC. Menstrual blood loss: Measurement and clinical correlates. Res Clin Forums 1982;4:69-80.

[32] Haynes PJ, Hodgson H, Anderson ABM, Turnbull AC. Measurement of menstrual blood loss in patients complaining of menorrhagia. Br J Obstet Gynaecol 1977;84:763-8.

[33] Barrer AP, Fowler. Blood loss during normal menstruation. Am J Obstet Gynaecol 1936;31:979-86.

[34] Kouides PA. Obstetric and gynaecological aspects of von Willebrand disease. Best Practice & Res Clin Haematol 2001;14:381-99.

[35] Hallberg I, Nilsson L. Determination of menstrual blood loss. Scand J Clin Lab Invest 1964;16:244-8.

[36] Higham JM, O'Brien PMS, Shaw RW. Assessment of menstrual blood loss using a pictorial chart. Br J Obstet Gynaecol 1990;97:734-9.

[37] Janssen CAH, Sholten PC, Heintz APM. A simple visual assessment technique to discriminate between menorrhagia and normal menstrual blood loss. Obstet Gynaecol 1995;85:977-82.

[38] Christiaens GC, Sixma JJ, Haspels AA. Hemostasis in menstrual endometrium: a review. Obstet Gynecol Survie 1982;37:281-303.

[39] Kadir RA, Economides DL, Sabin CA. Frequency of inherited bleeding disorders in women with menorrhagia. Lancet 1998;351:485-9.

[40] Ragni MV, Bontempo FA, Hassett AC. von Willebrand disease and bleeding in women. Haemophilia 1999;5:313:7.

[41] Kadir RA, Economides DL, Sabin CA, Pollard D, Lee CA. Assessments of menstrual blood loss and gynaecological problems in patients with inherited bleeding disorders. Haemophilia 1999;5:40-8.

[42] Gleeson NC, Buggy F, Sheppard BL, Bonnar J. The effect of tranexamic acid on measured menstrual blood loss and endometrial fibrinolytic enzymes in dysfunctional uterine bleeding. Acta Obstet Gynaecol Scand 1994;73:274-7.

[43] Preston JT, Cameron IT, Adams EJ, Smith SK. Comparative study of tranexamic acid and norethistrone in ovulatory menorrhagia. Br J Obstet Gynaecol 1995;102:401-6.

[44] Bonnar J, Sheppard BL. Treatment of menorrhagia during menstruation: randomised controlled trial of ethamsylate, mefanamic acid and tranexamic acid. Br Med J 1996;313:579-82.

[45] Ong YL, Hull DR, Mayne EE. Case report: menorrhagia in von Willebrand disease successfully treated with single daily dose tranexamic acid. Haemophilia 1998;4:63-5.

[46] Onundarson PT. Treatment of menorrhagia in von Willebrand's disease. Haemophilia 1999;5:76.

[47] Rybo G. Tranexamic acid therapy is effective treatment in heavy menstrual bleeding. Clinical update on safety. Ther Adv 1991;14:1-8.

[48] Mannucci PM, Ruggeri ZMé Pareti FI, Capitanio A. 1-deamino-8-D-arginine vasopressin: a new pharmacological approach to the management of hemophilia and von Willebrand disease. Lancet 1977;1:869-72.

[49] Mannucci PM, Aberg M, Nilsson IM, Robertson B. Mechanism of plasminogen activator and factor VIII increase after vasoactive drugs. Br J Haematol 1975;30:81-93.

[50] Mannucci PM. Desmopressin (DDAVP) in the treatment of bleeding disorders: the first factor to DDAVP in healthy subjects and patients with hemophilia A and von Willebrand`s disease. Br J Haematol 1981;47:283-93.

[51] Mannucci PM, Canciani L, Donovan BS. Response of factor VIII/von Willebrand factor to DDAVP in healthy subjects and patients with hemophilia A and von Willebrand's disease. Br J Haematol 1981;47:283-93.

[52] Mannucci PM. Desmopressin (DDVAP) for treatment of disorders of haemostasis. Prod Hemost Thromb 1986;8:19-45.

[53] Mannucci PM. DDAVP in the treatment of bleeding disorders: the first twenty years. Blood 1997;9:2515-21.

[54] Mannucci PM, Pareti FI, Holmberg L, Ruggeri ZM, Nilsson IM. Studies on the prolonged bleeding time in von Willebrand disease. J Lab Clin Med 1976;88:662-71.

[55] Kobrinsky NL, Gerrard JM, Watson CM, Israels ED, Cheang MS, Bishop AJ, Schroeder. Shortening of the bleeding time by 1-deamino-8-D-arginine vasopressin in various bleeding disorders. Lancet 1984; May 26:1145-8.

[56] Sakariasses KS, Cattaneo M, van der Berg A, Ruggeri ZM, Sixma JJ. DDAVP enhances platelet adherence and platelet aggregate growth on human artery subendothelium. Blood 1984;64:229-36.

[57] Balduini CL, Noris P, Belletti S, Spedini P, Gamba G. In vitro and in vivo effects of desmopressin on platelet function. Haematologica 1999;84:891-6.

[58] Fressinaud E, Veyradier A, Sigaud M, Boyer-Neumann C, Le Boterff C, Meyer D. Therapeutic monitoring of von Willebrand disease: interest and limits of a platelet function analyser at high shear stress. Br J Haematol 1999;106:777-83.

[59] Favaloro EJ. Utility of the PFA-100 ® for assessing bleeding disorders and monitoring therapy: a review of analytical variables, benefits and limitations. Haemophilia 2001;7:170-9.

[60] Favoloro EJ, Kershaw G, Bukuya M *et al.* Laboratory diagnosis of von Willebrand disease and monitoring of DDAVP therapy: efficacy of the PFA-100 ® and vWF:CBA as combined diagnostic strategies. Haemophilia 2001;7:180-9.

[61] Cash JD, Gader AMA, da Costa J. The release of plasminogen activator and factor VIII by LVP, AVP, DDAVP, ATIII, and GT in man. Br J Haematol 1974;27:363-4.

[62] Mariani G, Ciavarella N, Mazzucconi MG *et al*: Evaluation of the effectiveness of DDAVP in surgery and bleeding episodes in hemophilia and von Willebrand disease. A study of 43 patients. Clin Lab Haematol 1984;6:229-38.

[63] De la Fuente B, Kasper CK, Rickles FR, Hoyer LW: Response of patients with mild and moderate hemophilia A and von Willebrand disease to treatment with DDAVP. Ann Intern Med 1985;103:6-14.

[64] Mannucci PM, Lombardi R, Bader R *et al.* Heterogeneity of Type I von Willebrand disease: evidence for a subgroup with abnormal von Willebrand factor. Blood 1985;66:796-802.

[65] Rodeghiero F, Castaman G, Mannucci PM. Clinical indications for desmopressin (DDAVP) in congenital and acquired von Willebrand disease. Blood Reviews 1991;5:155-61.

[66] Mannucci PM, Cuttaneo M. DDAVP. A non-transfusional treatment of hemophilia and von Willebrand`s disease. Haemostasis 1992;22:276-80.

[67] Mannucci PM. A non-transfusional form of treatment for congenital and acquired bleeding disorders. Blood 1998;72:1449- 55.

[68] Lethagen S, Tennvall GR. Self-treatment with desmopressin intranasal spray in patients with bleeding disorders: effect on bleeding symptoms and socioeconomic factors. Ann Hematol 1993;66:257-60.

[69] Leissinger C, Becton D, Cornell C, Cox Gill J. High dose DDAVP intranasal spray (Stimate) for the prevention and treatment of bleeding in patients with mild haemophilia A, mild or moderate type I von Willebrand disease and symptomatic carriers of haemophilia A. Haemophilia 2001;7:258-66.

[70] Kadir RA, Lee CA, Sabin CA, Pollard D, Economides DL. DDAVP nasal spray for treatment of menorrhagia in women with inherited bleeding disorders: a randomized placebo-controlled crossover study. Haemophilia 2002;8:787-93.

[71] Dunn Al, Powers JR, Ribeiro MJ, Rickles FR, Abshire TC. Adverse events during use of intranasal desmopressin acetate for haemophilia A and von Willebrand disease: a case report and review of 40 patients. Haemophilia 2000;6:11-14.

[72] Smith TJ, Gill JC, Ambroso DR, Hathaway WE. Hyponatremia and seizures in young children given DDAVP. Am J Hematol 1989;31:199-202.

[73] Bond L, Bevin D. Myocardial infarction in a patient with hemophilia A treated with DDAVP. NEJM 1988;318:121.

[74] Byrnes JJ, Larcada A, Moake JL. Thrombosis following desmopressin for uremic bleeding. Am J Hematol 1988;28:63-5.

[75] Streiff M, Bell WR. Exercise and hemostasis in humans. Sem Hematol 1994;31:155-65.

[76] El Sayed MS. Effects of exercise on blood coagulation, fibrinolysis and platelet aggregation. Sports Med 1996;22:282-98.

[77] Jern C, Eriksson E, Tengborn L, Risberg B, Wadenvik H, Jern S. Changes of plasma coagulation and fibrinolysis in response to mental stress. Thromb Haemost 1989;62:767-71.

[78] Jern C, Wadenvik H, Mark H, Hallgren J, Jern S. Haematological changes during acute mental stress. Br J Haematol 1989;71:153-6.

[79] Stibbe J. Effect of exercise on FVIII-complex: proportional increase of ristocetin cofactor (von Willebrand factor) and FVIII-AHG, but disproportional increase of FVIII-AHF. Thromb Res 1977;10:163-8.

[80] Kopitsky RG, Switzer MEP, Williams RS, McKee PA. The basis for the increase in factor VIII procoagulant activity during exercise. Thromb Haemost 1983;49:53-7.

[81] Dimsdale JE, Moss J. Plasma catecholamines in stress and exercise. JAMA 1980;243:340.

[82] Selye H. The physiology and pathology of exposure to stress. Montreal, Canada: Acta Inc Medical Publ, 1950.

[83] Schneider DA, McLellan TM, Gass GC. Plasma catecholamine and blood lactate responses to incremental arm and leg exercise. Med Sci Sports Exercise 2000;32:608-13.

[84] Prentice CRM, Forbes CD, Smith SM. Rise of factor VIII after exercise and adrenaline infusion, measured by immunological and biological techniques. Thromb Res 1972;1:493-506.

[85] Ingram GIC, Vaughan Jones R. The rise in clotting factor VIII induced by man by adrenaline: effect of α and β-blockers. J Physiol 1966;187:447.

[86] Edgeberg O. Changes in the activity of antihemophilic A factor (F:VIII) and in the bleeding time associated with muscular exercise and adrenaline infusion. Scand J Clin Lab Invest 1963;15:539-49.

[87] Ingram GIC. Increase in antihaemophilic globulin activity following infusion of adrenaline. J Physiol 1961;156:217-24.

[88] Forwell GD, Ingram GIC. The effect of adrenaline infusion on human blood coagulation. J Physiol 1957;135:371.

[89] Rickles FR, Hoyer LW, Rick ME, Ahr DJ. The effects of epinephrine infusion in patients with von Willebrand's disease. J Clin Invest 1976;57:1618-25.

[90] Eyster ME, Ballard JO, Prager D. Comparison of factor VIII levels after adrenaline in classic hemophilia and von Willebrand's disease. Thromb Haemost 1978;39:657-62.

[91] Booth F, Allington MJ, Cederholm-Williams SA. An in vitro model for the study of acute release of von Willebrand factor from human endothelial cells. Br J Haematol 1987;67:71-8.

[92] Vischer UM, Wollheim CB. Epinephrine induces von Willebrand factor release from cultured endothelial cells: involvement of cyclic AMP-dependent signalling in exocytosis. Thromb Haemost 1997;77:195-202.

[93] Cohen RJ, Cohen SC, Epstein SE, Dennis LH. Alterations of fibrinolysis and blood coagulation induced by exercise, and the role of beta-adrenergic-receptor stimulation. Lancet 1968; Dec 14:1264-6.

[94] Small M, Tweddel AC, Rankin AC *et al* Blood coagulation and platelet function following exercise: effects of beta-adrenoreceptor blockade. Haemostasis 1984;14:262-8.

[95] van den Buerg PJM, Hospers JEH, Mosterd WL, Bouma BN, Huisveld IA. Aging, physical conditioning and exercise induced changes in hemostatic factors and reaction products. J Appl Physiol 2000;88:1558-1564.

[96] Rock G, Tittley P, Pipe A. Coagulation factor changes following endurance exercise. Clin J Sport Med 1997;7:94-9.

[97] van den Burg PJM, Hospers JEH, van Vliet M, Mosterd WL, Bouma M, Huisfeld IA. Changes in haemostatic factors and activation products after exercise in healthy subjects with different ages. Thromb Haemost 1995;74:1457-64.

[98] Arai M, Yorifuji H, Ikematsu S *et al.* Influences of strenuous exercise (triathlon) on blood coagulation and fibrinolytic system. Thromb Res 1990;57:465-71.

[99] Hansen J-B, Wilsgard L, Olsen JO, Osterud B. Formation and persistence of procoagulant and fibrinolytic activities in circulation after strenuous physical exercise. Thromb Haemost 1990;64:385-89.

[100] Andrew M, Carter C, O'Brodovitch H, Heigenhauser G. Increases in factor VIII complex and fibrinolytic activity are dependent on exercise intensity. J Appl Physiol 1986;60:1917-22.

[101] Wheeler ME, Davis GL, Gillespie WJ, Bern MM. Physiological changes in hemostasis associated with acute exercise. J Appl Physiol 1986;60:986-90.

[102] Brown JE, Baugh RF, Hougie C. Effect of exercise on the factor VIII complex: a correlation of the von Willebrand antigen and factor VIII coagulant antigen increase. Thromb Res 1979;15:61-7.

[103] Ferguson EW, Barr CF, Bernier LL. Fibrinogenolysis and fibrinolysis with strenuous exercise. J Appl Physiol 1979;47:1157-61.

[104] Lin X, El-Sayed MS, Waterhouse J, Reilly T. Activation and disturbance of blood haemostasis following strenuous physical exercise. Int J Sports Med 1999;20:149-53.

[105] Van den Burg PJ, Hospers JE, van Vliet M, Mosterd WL, Bouma BN, Huisveld IA. Effect of endurance training and seasonal fluctuation on coagulation and fibrinolysis in young sedentary men. J Appl Physiol 1997;82:613-20.

[106] El-Sayed MS, Lin X, Rattu AJM. Blood coagulation and fibrinolysis at rest and in response to maximal exercise before and after a physical conditioning programme. Blood Coagul Fibrinolysis 1995;6:747-52.

[107] Van den Burg PJM, Hospers JEH, van Vliet M, Mosterd WL, Huisveld IA. Unbalanced haemostatic changes following strenuous physical exercise. Eur Heart J 1995;16:1995-2001.

[108] Prisco D, Paniccia R, Guarnaccia V *et al.* Thrombin generation after physical exercise. Thromb Res 1993;69:159-64.

[109] Davis GL, Abildgaard CF, Bernauer EM, Britton M. Fibrinolytic and hemostatic changes during and after maximal exercise in males. J Appl Physiol 1976;40:287-92.

[110] Egeberg O. The effect of exercise on the blood clotting system. J Clin Lab Invest 1963;15:8-13.

[111] Iatridis SG, Ferguson JH. Effect of physical exercise on blood clotting and fibrinolysis. J Appl Physiol 1963;18:337-44.

[112] Rizza CR. Effect of exercise on the level of antihaemophilic globulin in human blood. J Physiol 1961;156:128-35.

[113] Mandalaki T, Dessypris A, Louizou C, Panayotopoulou C, Dimitriadou C. Marathom run III: effects on coagulation, fibrinolysis, platelet aggregation and serum cortisol levels. A 3 –year study. Thromb Haemost 1980;43:49-52.

[114] Herren T, Bärtsch P, Haeberli A, Straub PW. Increased thrombin-antithrombin III complexes after 1 h of physical exercise. J Appl Physiol 1992;73:2499-504.

[115] Dufaux B, Order U, Liesen H. Effect of a short maximal physical exercise on coagulation, fibrinolysis, and complement system. Int J Sports Med 1991;12:38-42 (suppl)

[116] Bärtsch P, Haeberli A, Straub PW. Blood coagulation after long distance running: antithrombin III prevents fibrin formation. Thromb Haemost 1990;63:430-4.

[117] Hyers TM, Martin BJ, Pratt DS, Dreisin RB, Franks JJ. Enhanced thrombin and plasmin activity with exercise in man. J Appl Physiol 1980;48:821-5.

[118] Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehab Med 1970;2:92-8.

[119] Handa K, Terao Y, Mori T et al. Different coagulability and fibrinolytic activity during exercise depending on exercise intensities. Thromb Res 1992;66:613-6.

[120] Fergusson EW, Bernier LL, Banta GR, Yu-Yahiro J, Schoomaker EB. Effects of exercise and conditioning on clotting and fibrinolytic activity in men. J Appl Physiol 1987;62:1416-21.

[121] Kvernmo HD, Østerud B. The effect of physical conditioning suggests adaptation in procoagulant and fibrinolytic potential. Thromb Res 1997;87:559-69.

[122] Foster KG, Ginsburg J, Weiner JS. Adrenaline and sweating in man. J Physiol 1967;191:131-2.

[123] Allen JA, Roddie JC. The role of circulating catecholamines in sweat production in man. J Physiol 1972;227:801-14.

[124] Aoki K, Shiojiri T, Shibasaki M, Takano S, Kondo N, Iwata A. The effect of diurnal variation on the regional differences in sweating and skin blood flow during exercise. Eur J Appl Physiol Occup Physiol 1995;71:276-80.

[125] Takano S, Kondo N, Shibasaki M, Aoki K, Inoue Y, Iwata A. The influence of work loads on regional differences in sweating rates. Jpn J Physiol 1996;46:183-86.

[126] Kondo N, Takano S, Aoki K, Shibasaki M, Tominaga H, Inoue Y. Regional differences in the effect of exercise intensity on thermoregulatory sweating and cutaneous vasodilation. Acta Physiol Scand 1998;164:71-78.

[127] Tankersley CG, Smolander J, Kenney WL, Fortney SM. Sweating and skin blood flow during exercise: effect of age and maximal oxygen uptake. J Appl Physiol 1991;71:236-42.

[128] Yamazaki F, Fujii N, Sone R, Ikegami H. Responses of sweating and body temperature to sinusoidal exercise in physically trained men. J Appl Physiol 1996;80:4911-5.

[129] Yeung WH, Palmer J, Schidlow D, Bye MR, Huang NN. Evaluation of a paper-patch test for sweat chloride determination. Clin Ped 1984;23:603-7.

[130] Warwick WJ, Huang NN, Waring WW *et al.* Evaluation of a cystic fibrosis screening system incorporating a miniature sweat stimulator and disposable chloride sensor. Clin Chem 1986;32;850-3 .

[131] Warwick WJ, Hansen LG, Werness ME. Quantification of chloride in sweat with the cystic fibrosis indicator system. Clin Chem 1990;36;96-8.

[132] Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. Biometrics 1987;43:487-98.

[133] Frisson L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. Stat Med, 1992;11:1685-1704.

[134] Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. BMC Med Res Method 2001;1:6.

[135] Greenland S. Modeling and variable selection in epidemiologic analysis. Am J Public Health 1989;79:340-349.

[136] Bring J. How to standardize regression coefficients. Am Stats 1994;48:209.

[137] Julious SA, Campbell MJ, Altman DG. Estimating sample sizes for continuous, binary, and ordinal outcomes in paired comparisons: practical hints. J Biopharm Stats 1999;9:241-51.

[138] Golomb LM, Solidum AA, Warren MP. Primary dysmenorrhea and physical activity. Med Sci Sports Exercise 1998;30:906-9.

[139] Powell KE, Thompson PD, Caspersen CJ, Kendrick JS. Physical activity and the incidence of coronary heart disease. Annu Rev Public Health 1987;8:253-287.

[140] Walsh A. *Statistics for the Social Sciences.* Harper and Row, New York, 1990.

[141] Greenland S, Schlesselman JJ, Criqui MH. The fallacy of employing standardized regression coefficients and correlations as measures of effect. Am J Epidemiol 1986;123:203-8.

[142] Fukuba Y, Walsh ML, Morton RH *et al.* Effect of endurance training on blood lactate clearance after maximal exercise. J Sports Sci 1999;17:239-48.

[143] Stallknecht B, Vissing J, Galbo H. Lactate production and clearance in exercise. Effects of training. A mini-review. Scand J Med Sci Sports 1998;8:127-31.

[144] Greive JS, Hickner RC, Shah SD, Cryer PE, Holloszy JO. Norepinephrine response to exercise at the same relative intensity before and after endurance exercise training. J Appl Physiol 1999;86:531-5.

[145] Strobel G, Friedmann B, Siebold R, Bartsch P. Effects of severe exercise on plasma catecholamines in differently trained athletes. Med Sci Sports Exercise 1999;51:560-5.

[146] Greenland S, Morgenstern H. Confounding in health research. Annu Rev Public Health 2001;22:189-212.

[147] Willich SN, Lewis M, Lowel H *et al.* Physical exertion as a trigger of acute myocardial infarction. NEJM 1993;151:459-67.

[148] Britton A, McKee M, Black N *et al.* Threats to applicability of randomised trials: exclusions and selective participation. J Health Serv Res Policy 1999;4:112-21.

[149] Amori G, Lenox RH. Do volunteer subjects bias clinical trials? J Clin Psychopharmacol 1989;9:321-7.

[150] Cook TD, Campbell. *Quasi-Experimentation. Design and Analysis Issues for Field Settings.* Houghton Mifflin Company, Boston, 1979.

[151] Berger VW, Exner DV. Detecting selection bias in randomized clinical trials. Control Clin Trials 1999;20:319-27.

[152] Winter EM, Estion RG, Lamb KL. Statistical analyses in the physiology of exercise and kinanthropometry. J Sports Sci, 2001;19:761-75.

[153] Hill AB. The environment and disease: association or causality? Proc R Soc Med 1965;58:295-300.

[154] Cochran WG, Cox GM. *Experimental Designs.* John Wiley and Sons, New York, 1957.

[155] Astrand P-O, Rodahl K. *Textbook of work physiology. Physiological basis of exercise. 3rd ed.* McGraw-Hill book company, New York,1986:188-91.

[156] Verstraete M. Haemostatic drugs. In: *A critical Appraisal.* Matinus Nijhoff , The Hague, Netherlands,1997:1-153.

[157] Louis TA, Lavori PW, Bailar JC, Polansky M. Crossover and self-controlled designs in clinical research. NEJM 1984;310:24-31.

[158] Matthews JNS. Multi-period crossover trials. Stat Meth Med Resear 1994;3:383-405.

[159] Huitson A. A review of cross-over trials. Statistician1982;31:71-80.

[160] Hills M, Armitage P. The two-period cross-over clinical trial. Br J Pharmacol 1979;8:7-20.

[161] Brown, Jr. BW. The crossover experiment for clinical trials. Biometrics 1980;36:69-79.

[162] Berger VW, Exner DV. Detecting selection bias in randomized clinical trials. Control Clin Trials 1999;20:319-27.

[163] Kramer MS, Shapiro SH. Scientific challenges in the application of randomized trials. JAMA 1984;252:2739-45.

[164] Overall JE, Shobaki G, Shivakumarc C, Steele J. Adjusting sample size for anticipated dropouts in clinical trials. Psychopharmacol Bull 1998;34:25-33.

[165] Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319:670-4.

[166] Fergusson D, Aaron SD, Guyatt G, Hebert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ 2002;325:652-4.

TABLES

Table 1- Analytical Methods

VARIABLE MEASURED	SPECIMEN HANDLING	ANALYSIS TECHNIQUE	INSTRUMENT	ADDITIONAL STEPS	UNITS
Sodium	i. filter paper patches allowed to dry on room air ii. each patch was eluted into a fixed volume of distilled de-ionized water iii. the concentration of NaCl in the eluate was measured	electrochemical determination	CMT 10 chloride titrator (Radiometer, Denmark)	Concentrations multiplied by fixed volume of eluate to derive total amount	mmol/L
Chloride		Flame Emission Spectrophotometry	IL10 flame photometer (Instrument Laboratories, USA)		
FVIII:C	i. immediate centrifugation at 4°C ii. snap frozen on dry ice iii. stored at –80°C for later analysis in batches	aPTT-based one-stage clotting assay using FVIII deficient plasma (Diagnostica Stago's Reagent)	STA compact instrument (Asnière, France)	Supplementary dilutions for values exceeding 1.0 U/ml and repeated (carried out to ensure readings from the most accurate portion of their respective standard curves)	U/ml (normal range: 0.5-1.5 U/ml)
vWF:Ag		Automated immunoturbidometric assay (LIATEST)			
aPTT	analyzed immediately at 37°C	aPTT-based one-stage clotting assay	STA compact instrument (Asnière, France)		Seconds (normal range 29-43)
BT		in vitro with both ADP and epinephrine platelet agonists	Platelet Function Analyzer (PFA-100) test cartridge system		Seconds (normal ranges: ADP- 50-105 Epi - 60-150)
Blood group		standard forward and backward blood bank procedure			
Catecholamines (adrenaline and noradrenaline)	stored immediately on dry ice	High Performance Liquid Chromatography (HPLC)	Perkin Elmer LC system		pmol/l
Lactate		enzymatic end-point procedure	Beckman Synchron System (Beckman Instruments, CA, USA)		mmol/l

Table 2- Dependent Variables

Variable	Definition	Scale (Units)
vWF:Ag change* (Time 1)	Absolute level measured at Time 1 minus the value measured at baseline (Time 0)	Continuous (U/ml)
vWF:Ag change* (Time 2)	Absolute level measured at Time 2 minus the value measured at baseline (Time 0)	Continuous (U/ml)

*used in regression

Table 3 – Main Independent Variables (continuous markers of exercise)

Variable	Definition	Scale (units)
Forearm sodium change*	Total absolute level as measured after exercise (Time 1 plus Time 2) minus absolute value measured at baseline (Time point 0)	Continuous (no units)
Lactate change*	Absolute level as measured after exercise (Time 1) minus absolute value measured at baseline (Time 0)	Continuous (mmol/l)
Noradrenaline change*	Absolute level as measured after exercise (Time 1) minus absolute value measured at baseline (Time 0)	Continuous (pmol/l)
Adrenaline change*	Absolute level as measured after exercise (Time 1) minus absolute value measured at baseline (Time 0)	Continuous (pmol/l)

*used in regression

Table 4 – Potential Confounding Variables

Variable	Definition	Scale (units)
Age*	Chronological age	Continuous (years)
Fitness Level*	Determined from a continuous score calculated from a validated fitness questionnaire[1]. Possible score values range from 0 to >20. Category of fitness level then determined according to absolute score.	Dichotomous unfit <11=0 fit >11=1

*used in regression

[1] adapted from R.L. and J. Henderson, *Fitness Running*, Human Kinetics, Champaign IL, 1994

Table 5 – Determinants of Baseline vWF:Ag

Variable	Definition	Scale (units)
Race	Caucasian or other	Dichotomous Caucasian=0 Other=1
Blood group	Determined by forward and backward blood banking procedures	Categorical O= reference A AB B
Menstrual phase	Categories of menstrual phase determined by day of menstrual cycle by history at the time that they perform the exercise protocol Individuals in the follicular phase are then subdivided into follicular menstruating phase and non-menstruating according to the presence or absence of their menstrual period during the exercise study	Categorical luteal (days 14-28)=reference Follicular non-menstruating (days 6-14 and no menses) Follicular menstruating (days 1-5 and menses)
Smoking status	Smoker defined as current and regular use of tobacco (>1 pack/week) Past-smokers and current infrequent smokers are categorized as non-smokers	Dichotomous non-smoker=0 smoker=1
Contraceptive use	Users are defined by current use or use within 1 month preceding their exercise session Previous users are categorized as non-users	Dichotomous non-user=0 user=1

Table 6 – Distribution of Baseline Categorical Variables

Variable	Frequency Distribution (%)
Race	
Caucasian	97.8
Other	2.2
Menstrual phase	
Luteal	47.5
Follicular non-menstruating	37.5
Follicular menstruating	15.0
Blood group	
O	52.5
A	30.0
B	15.0
AB	2.5
Smoking status	
Y	10.0
N	90.0
Oral contraceptive use	
Y	55.0
N	45.0
Activity Level	
Unfit	52.5
Fit	45.0
Very Fit	2.5

Table 7 – Baseline Continuous Variables

Variable (units)	Median	Mean (+/- SD)	Range
Age	25.5	27.0+/- 5.4	18-39

Table 8 - Independent Variables

Variable*	Median	Mean (+/- SD)	Range
Forearm sodium	3.30	3.26+/- 1.72	0.80-8.40
Lactate	5.75	5.57+/- 2.31	-0.31-10.73
Adrenaline	80.00	86.23+/- 130.92	-215.00-401.00
Noradrenaline	1935	1689+/- 1290	-827-5369

* change in variable indicated from baseline (Time 0) to immediately after exercise (Time1)

Table 9 - Dependent Variables

Variable *	Median	Mean (+/- SD)	Range
vWF:Ag 1	0.25	0.31+/- 0.26	-0.12-1.11
vWF:Ag 2	0.16	0.23+/- 0.25	-0.12-0.98

* vWF:Ag 1 = the difference in vWF:Ag from baseline (Time 0) to Time 1, vWF:Ag 2 = the difference in vWF:Ag from baseline (Time 0) to Time 2

Table 10 – Univariate Analysis-Paired Student t-test

Parameter	Time 0 mean +/- SD (95% CI)	Time 1 mean +/- SD (95% CI)	Time 2 mean +/- SD (95% CI)
vWF:Ag (U/mL)	0.83 +/- 0.19 (0.77 0.89)	1.14 +/- 0.32 * (1.04 1.24)	1.06 +/- 0.30 * (0.97 1.15)
FVIII:C (U/mL)	1.12 +/- 0.20 (1.06 1.18)	1.77 +/- 0.60 * (1.58 1.96)	1.55 +/- 0.59 * (1.37 1.73)
APTT (seconds)	36.0 +/- 3.06 (35.0 36.9)	35.1 +/- 2.94 • (34.1 36.0)	34.9 +/- 3.37 • (33.9 35.9)
CT ADP (seconds)	85.2 +/- 14.12 (80.8 89.6)	72.4 +/- 16.66 * (67.2 77.6)	80.4 +/- 15.31 • (75.6 85.1)
CT Epi (seconds)	129.6 +/- 34.37 (118.9 140.2)	105.1 +/- 28.58 * (96.2 114.0)	116.1 +/- 36.31 • (104.8 127.3)

* indicates significance compared to baseline using a paired-student t-test of $p < 0.0001$

• indicates significance compared to baseline using a paired-student t-test of $p < 0.05$

Table 11 – Univariate Analysis- Linear Regression *dependent variable = vWF:Ag change (Time 2)*

Variable†	Coefficient‡	SE	P-value	95% CI	R^2
Forearm sodium*	0.05	0.02	0.04	0.00-0.09	0.10
Lactate*	0.03	0.02	0.08	0.00-0.06	0.08
Noradrenaline	0.02**	0.03	0.55	-0.04-0.08	0.01
Adrenaline	0.03***	0.03	0.41	-0.04-0.09	0.02

† change in variable indicated from baseline (Time 0) to immediately after exercise (Time1)
‡ rounded to 2 decimal places
*regression coefficient for a 1 unit change
** regression coefficient for a 1000 unit change in noradrenaline
*** regression coefficient for a 100 unit change in adrenaline

Table 12-Univariate Analysis- Linear Regression *dependent variable = vWF:Ag change (Time 2)*

Variable	Units change in independent variable	Units change in dependent variable (vWF:Ag) ‡	Final vWF:Ag level (U/ml) from mean baseline of 0.83 U/ml ‡
Forearm sodium	1	0.05	0.88
	5	0.22	1.05
	10	0.45	1.28
Lactate	1	0.03	0.86
	5	0.15	0.98
	10	0.30	1.18
Noradrenaline	1000	0.02	0.85
	5000	0.09	0.92
Adrenaline	100	0.03	0.86
	500	0.13	0.96

‡ rounded to 2 decimal places

Table 13 – Multivariate Analysis – Interactions *dependent variable = vWF:Ag change (Time 2)*

Variable†	Coefficient	SE	95% CI	P-value
Sodium*Age	0.004	0.003	-0.002-0.011	0.21
Sodium*Fitness	0.017	0.044	-0.007-0.108	0.70
Sodium*vWF:Ag*	-0.142	0.109	-0.363-0.080	0.20
Lactate*Age	-0.003	0.003	-1.010-0.321	0.32
Lactate*Fitness	-0.071	0.037	-0.147-0.004	0.06
Lactate*vWF:Ag	0.047	0.082	-0.120-0.215	0.57
Noradrenaline*Age	-0.0000033	0.0000056	-0.0000146-0.0000080	0.56
Noradrenaline*Fitness	-0.000092	0.000059	-0.000212-0.000029	0.13
Noradrenaline*vWF:Ag	-0.00014	0.00018	-0.00050-0.00022	0.43
Adrenaline*Age	-0.000041	0.000065	-0.00017-0.000091	0.53
Adrenaline*Fitness	-0.00021	0.00062	-0.00146-0.00105	0.74
Adrenaline*vWF:Ag	0.0039	0.0020	-0.0038-0.0046	0.85

† Interaction term tested in base model containing respective main independent variable and adjusted for fitness and age

* vWF:Ag = vWF:Ag at baseline (Time 0)

Table 14 - Multivariate Analysis – Final Models *dependent variable = vWF:Ag change (Time 2)*

Model	Variable†	Coefficient‡	SE	P-value	95% CI	Partial R^2	Stand Coeff¥
1	Forearm sodium*	0.05	0.02	0.02	0.01-0.09	0.146	0.36
2	Lactate*	0.02	0.02	0.21	-0.01-0.06	0.042	0.20
3	Noradrenaline**	0.02	0.03	0.42	-0.04-0.08	0.018	0.13
4	Adrenaline***	0.05	0.03	0.11	-0.01-0.11	0.066	0.25

† Main independent variable of base model adjusted for fitness
‡ rounded to 2 decimal places
¥ standardized regression coefficient
* regression coefficient for a 1 unit change
** regression coefficient for a 1000 unit change in noradrenaline
*** regression coefficient for a 100 unit change in adrenaline

Table 15- Multivariate Analysis – Final Models *dependent variable = vWF:Ag change (Time 2)*

Variable	Units change in independent variable	Units change in dependent variable (vWF:Ag) ‡	Final vWF:Ag level (U/ml) from mean baseline of 0.83 U/ml ‡
Forearm sodium	1	0.05	0.88
	5	0.26	1.09
	10	0.52	1.35
Lactate	1	0.02	0.85
	5	0.11	0.94
	10	0.22	1.05
Noradrenaline	1000	0.02	0.85
	5000	0.12	0.95
Adrenaline	100	0.05	0.88
	500	0.24	1.07

‡ rounded to 2 decimal places

Table 16 – Multivariate Analysis – Final Models *dependent variable = vWF:Ag change (Time 2)*

Model	Variable*	Covariate	F-value	P-value	Adjusted R^2
1	Forearm sodium	Fitness	5.39	0.01	0.18
2	Lactate	Fitness	2.78	0.07	0.08
3	Noradrenaline	Fitness	2.27	0.12	0.06
4	Adrenaline	Fitness	3.34	0.05	0.11

* Main independent variable of base model adjusted for fitness

Table 17: Frequency (in descending order) according to reason of exclusion

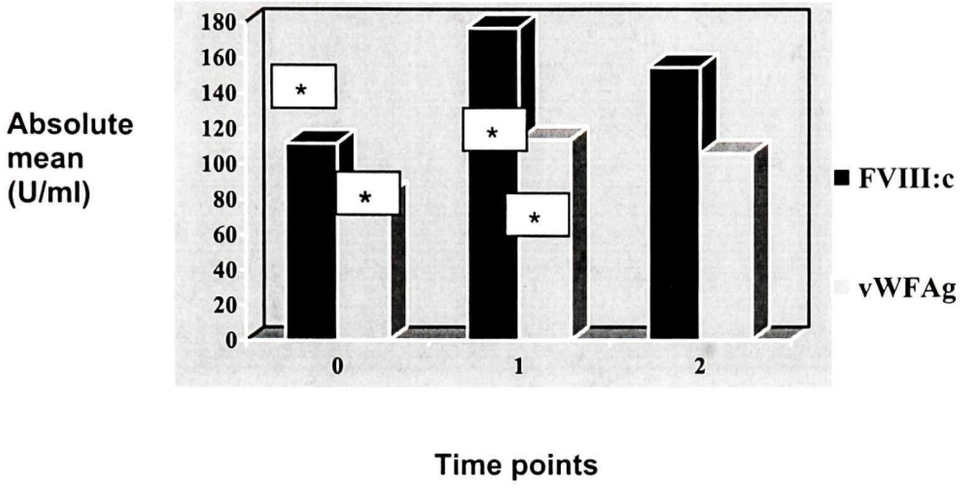
Number	Reason	Number	Frequency (%)
1	No menorrhagia	6	42.9
2	Irregular periods	4	28.6
3	On medication	2	14.3
4	Pregnant	1	7.1
5	Exercise intolerance	1	7.1
Total		14	100

Table 18: Frequency (in descending order) according to reason of refusal

Number	Reason	Number	Frequency (%)
1	Insufficient time	5	35.7
2	Dislike exercise	3	21.4
3	Satisfied with current treatment	2	14.3
4	Unwilling to modify current exercise program	1	7.1
5	Trying to become pregnant	1	7.1
6	Unconvinced that exercise helps	1	7.1
7	None	1	7.1
Total		14	100

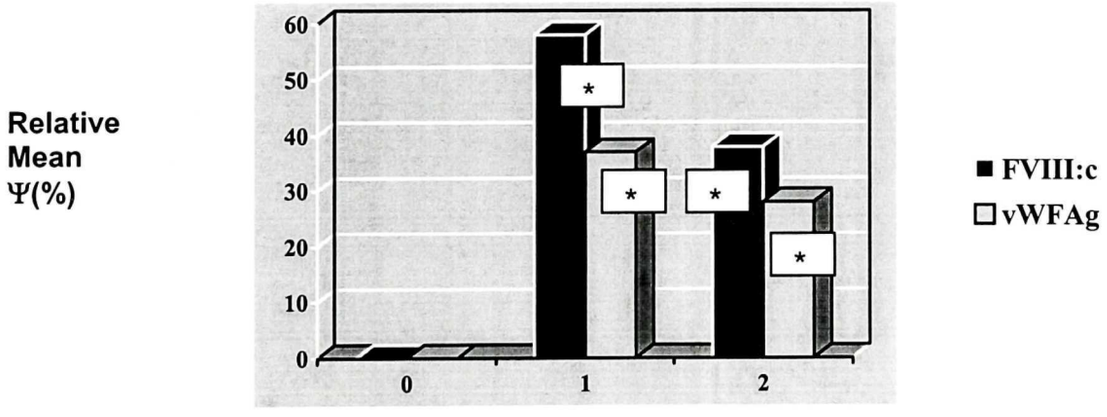
FIGURES

Figure 1: Mean absolute values of vWF:Ag (U/ml) and FVIII:C (U/ml) at baseline (Time 0), immediately after exercise (Time 1) and 30 minutes after exercise (Time 2).



* indicates significance at a $p < 0.05$ using a paired Student t-test

Figure 2: Mean increases (in per cent) in vWF:Ag and FVIII:C immediately after exercise (Time 1), and 30 minutes after exercise (Time 2) adjusted for baseline (Time 0).



* indicates significance at a $p < 0.05$ using a paired Student t-test

Ψ mean % increase was calculated by subtracting the mean pre-exercise values from the mean post-exercise values, dividing by the pre-exercise values and multiplying by 100.

Figure 3: Frequency Distribution of absolute increase in vWF:Ag (U/ml) immediately after exercise (Time 1) compared to baseline (Time 0).

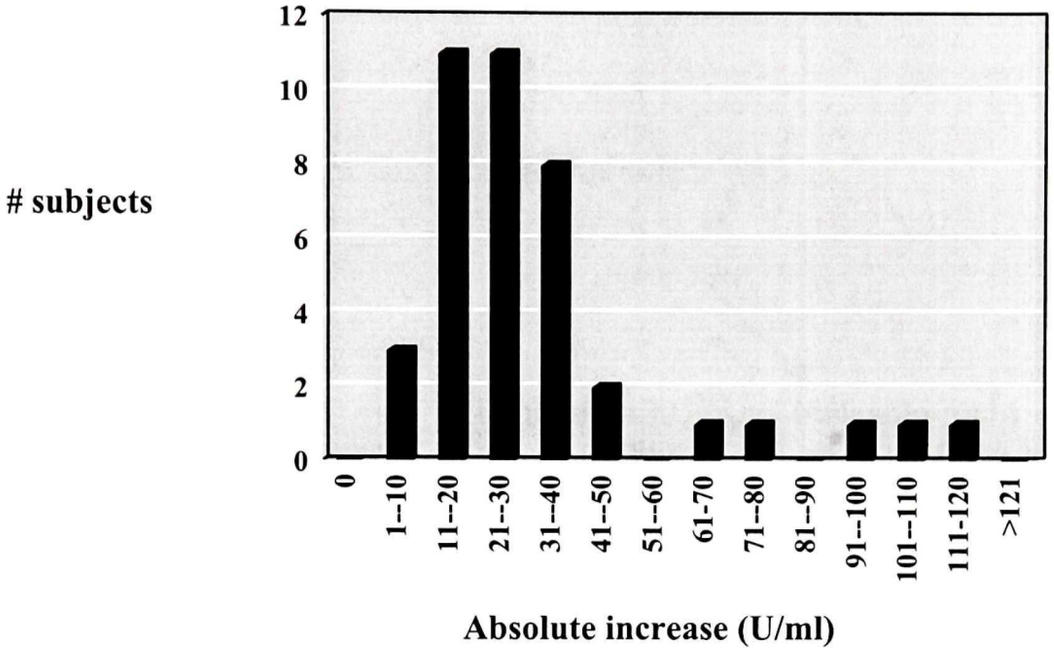


Figure 4: Frequency distribution of absolute increase in FVIII:C immediately after exercise (Time 1) compared to baseline (Time 0)

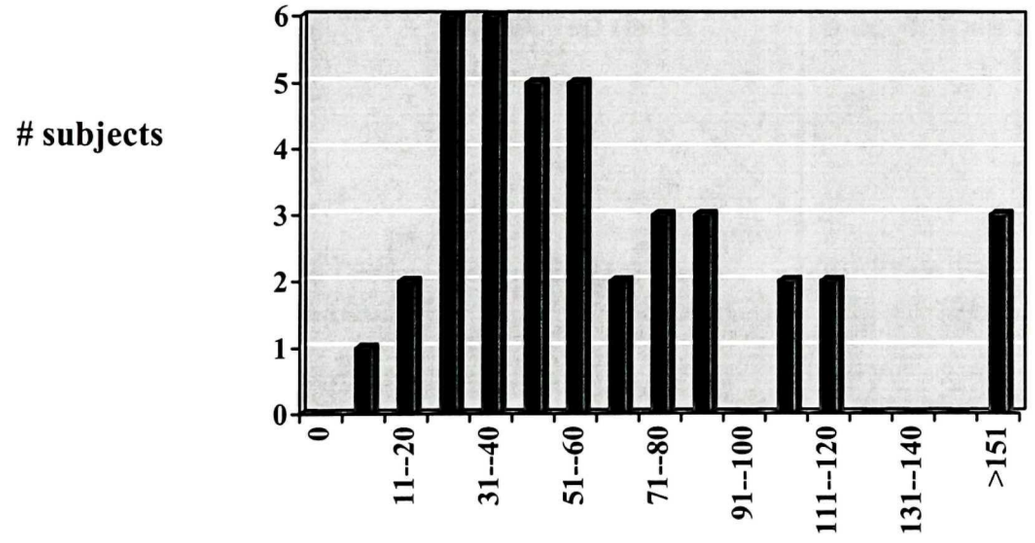
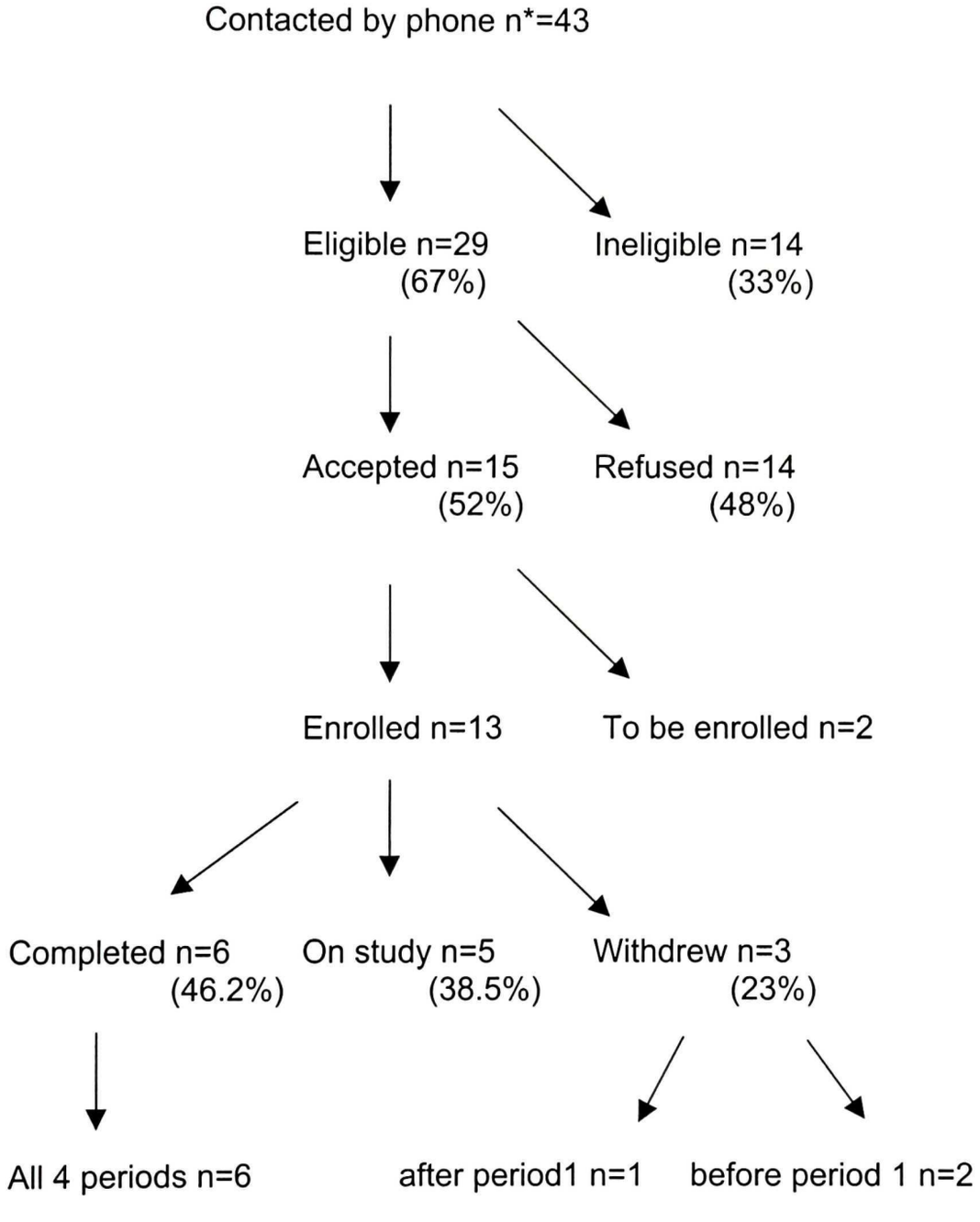


Figure 5: Enrollment Algorithm



* n= number of subjects

APPENDIX: Fitness Questionnaire

From: R.L. and J. Henderson, *Fitness Running*, Human Kinetics, Champaign IL, 1994

Please choose the statement that best describes you or your condition in the categories that follow:

Cardiovascular health

No history of heat or circulatory problems	____(3)
Past ailments have been successfully treated	____(2)
Such problems exists but no treatment required	____(1)
Under medical care for cardiovascular illness	____(0)

Injuries

No current injury problems	____(3)
Some pain in activity but not limited by it	____(2)
Level of activity is limited by the injury	____(1)
Unable to do much strenuous exercise	____(0)

Illnesses

No current illness	____(3)
Level of activity not limited by the illness	____(2)
Level of activity limited by the illness	____(1)
Unable to do much strenuous training	____(0)

Age

Age <19	____(3)
Ages 20-29	____(2)
Ages 30-39	____(1)
Ages 40 and older	____(0)

Weight

Within 5 pounds of ideal weight	____(3)
6-10 pounds above or below ideal	____(2)
11-19 pounds above or below ideal weight	____(1)
>20 pounds above or below ideal	____(0)

Resting Pulse Rate

Below 60 beats per minute	____(3)
60-69 beats per minute	____(2)
70-79 beats per minute	____(1)
>80 beats per minute	____(0)

Smoking

Never a smoker	____(3)
Once a smoker but quit	____(2)
An occasional, light smoker now	____(1)
A regular, heavy smoker now	____(0)

Most Recent Aerobic Workout

Participate in aerobic exercise classes >3 times a week	____(3)
Participate in aerobic exercise classes 2 to 3 times a week	____(2)
Participate in aerobic exercise classes 1 to 2 times a week	____(1)
No recent aerobic classes	____(0)

Aerobic Exercise Background

Participated in aerobic step workouts 1 to 2 years ago	____(3)
Participated seriously on a regular basis within the past year	____(2)
Participate in aerobic steps workouts 1 to 2 times a week	____(1)
No recent aerobic step training	____(0)

Related Activities

Regularly practice similar aerobic activities	____(3)
Regularly practice less vigorous activities	____(2)
Regularly practice nonaerobic sports (weight-lifting, tennis, golf etc..)	____(1)
Not regularly active in any physical activities	____(0)

TOTAL SCORE_____

Find Your Fitness Level

Score	**Fitness level**
20 or more	high
10-19	average
less than 10	below average