Potential Undiagnosed VWD Or Other Mucocutaneous Bleeding Disorder Cases Estimated From Private Medical Insurance Claims

This article was published in the following Dove Press journal: Journal of Blood Medicine

Robert F Sidonio Jr1
Ayesha Zia2
Dana Fallaize3

1Aflac Cancer and Blood Disorders Center, Children’s Healthcare of Atlanta, Emory University, Atlanta, GA, USA;
2Division of Hematology Oncology, University of Texas Southwestern Medical Center, Dallas, TX, USA;
3Department of Life Sciences, Charles River Associates, Boston, MA, USA

Introduction: Von Willebrand disease (VWD) is a common inherited bleeding disorder, but awareness among health care professionals is low. We estimated the number of cases of undiagnosed VWD or other mucocutaneous bleeding disorders among commercially insured patients in the United States with a recent history of bleeding events.

Methods: Patients with a VWD diagnosis who were users of or candidates for von Willebrand factor replacement were identified from the IMS PharMetrics Plus Database (2006–2015). We constructed a unary patient-finding model based on 12 prediagnosis variables that best defined this population, and applied this to undiagnosed patients with recent bleeding events from the same database. Cases of symptomatic undiagnosed VWD or other mucocutaneous bleeding disorders in the commercially insured population were estimated from the “best fit” (positive predictive value [PPV] 83%) and “good fit” (PPV 75%) patients thus identified.

Results: Overall, 507,668 undiagnosed patients with recent bleeding events were identified (86% female, 14% male). Application of the VWD model identified 3318 best-fit and 37,163 good-fit patients; 91% of best-fit patients were females aged <46 years, with heavy menstrual bleeding as the most common claim. Projection to the full commercially insured US population suggested that 35,000–387,000 patients may have symptomatic, undiagnosed VWD or other mucocutaneous bleeding disorders.

Discussion: Computer modeling suggests there may be a significant number of patients with symptomatic, undiagnosed VWD or other mucocutaneous bleeding disorder in the commercially insured population. Enhanced awareness of VWD symptoms and their impact, and of screening and testing procedures, may improve the diagnosis of VWD and reduce disease burden.

Keywords: medical insurance claims, delayed diagnosis, von Willebrand disease

Plain Language Summary

Von Willebrand disease (VWD) is a relatively common disease inherited from your parents that can cause uncontrolled bleeding. However, the level of awareness of this disease in the medical community is low and sometimes patients with the disease are not diagnosed for many years, and therefore do not receive appropriate medical treatment.

Here, we aimed to understand how many people may be living with this disease in the USA without a diagnosis. To do this, we looked at anonymous records from private medical insurance claims. We first identified patients with VWD and looked at the symptoms listed in their records. We then looked for other patients that had insurance claims for symptoms that closely matched those of the patients we knew had VWD, but that did not have a diagnosis of...
VWD, using a computer model. This gave us an estimate of how many patients may actually have VWD (or potentially some other kind of bleeding disease that has similar symptoms, that would also need treatment) but have not been diagnosed.

Our model suggests that between 35,000 and 387,000 people in the USA (or 0.017–0.19% of those with private insurance) may have undiagnosed VWD or other similar bleeding diseases. This tells us there is a need to improve the awareness of, and screening and testing for, these kinds of bleeding diseases in the USA so that those patients may receive appropriate care for their disorders.

**Introduction**

Von Willebrand disease (VWD), the most common inherited bleeding disorder, is caused by defects or deficiency in the von Willebrand factor (VWF) plasma protein. A Canadian study found a prevalence of symptomatic VWD in the primary care setting of at least 0.1%, while a European survey identified a prevalence of treated VWD at 4.5–24 per million people.\(^1\)\(^2\) Although genetic variants associated with VWD are equally common in males and females, symptoms such as heavy menstrual bleeding (HMB) may lead to higher rates of diagnosis in females.\(^3\)

Von Willebrand disease is underdiagnosed among symptomatic patients, owing to a lack of disease awareness\(^5\) and the complexities of clinical and laboratory assessments.\(^4\)\(^6\)\(^7\) This increases the burden of disease on patients and the health care system, as accurate diagnosis and categorization of VWD are necessary to inform the therapeutic strategy for episodic bleed management and, in severe cases, for long-term prophylaxis via VWF replacement.\(^4\)\(^6\)\(^7\) Bleeding symptoms such as epistaxis, HMB, and joint or gastrointestinal (GI) bleeds — together with comorbidities such as iron deficiency anemia — can negatively impact quality of life and increase health care utilization.\(^7\)\(^8\)\(^13\) Patients with VWD are also exposed to a higher risk of future bleeding complications from pregnancy, childbirth, surgical and dental procedures, and minor injuries, which can be managed with appropriate intervention following diagnosis.\(^14\)\(^16\)

We showed recently, through longitudinal analysis of medical insurance claims data from a commercially insured US population, that VWD patients’ need for bleed care (number of medical claims for VWD or bleeding events and prescription claims) was substantially reduced after (vs before) VWD diagnosis.\(^17\) However, misrecognition of VWD was common, with a quarter of patients visiting the same health care specialty for episodic bleed care at least twice before diagnosis. After diagnosis, misrecognized patients continued to have more bleed claims than other diagnosed patients, although the overall frequency of bleed claims was reduced; this further highlights the importance of accurate and timely diagnosis.\(^17\)

The objective of the current analysis was to utilize claims data and unary modeling to estimate the occurrence of undiagnosed, symptomatic VWD (or other mucocutaneous bleeding disorders with similar symptoms) in the commercially insured population. A further objective was to characterize the demographics, symptoms, and care settings of these undiagnosed patients.

**Methods**

**Data Source**

Data are from the IMS Health PharMetrics Plus Database, which contains individual-level, de-identified health care claims information from employers, hospitals, and health plans on more than 150 million commercially insured patients in the United States. Data were extracted for medical claims for VWD (ICD-9 286.4), other bleeding disorders and bleeding events, and for prescription claims for treatments of interest resulting from physician encounters between January 1, 2006 and June 30, 2015. We did not include qualitative platelet disorders coding in this analysis, and primary diagnoses of hemophilia A or treatment with anticoagulants (identified by diagnostic or treatment codes) were excluded.

Because this study used de-identified patient data, ethical approval was not required.

**Development Of The Unary Patient-Finding Predictive Model For VWD**

A 2-class approach is typically used in predictive modeling, with the disease model being built on 2 groups: patients who have the disease and people who do not have the disease.\(^18\)\(^19\) However, in our population, VWD status was definitive only for patients previously diagnosed with the disease, making this approach inappropriate. Therefore, we applied a proprietary approach to a relatively new framework of statistical study, referred to as positive-unlabeled learning,\(^20\) by constructing a model that compares the intrinsic properties of patients with known disease (the reference set) vs a set of patients whose status is unknown (the potential set). By developing a unique, multivariate “fingerprint” of a single (unary) class of patients with VWD in the reference set, we could assign a likelihood of having VWD/bleeding disorders to individuals in the
potential set by assessing their similarity to severely symptomatic VWD patients in the reference set.

In this study, the reference set consisted of severely symptomatic patients with diagnosed VWD (users of, or in the investigators opinion candidates for, VWF replacement therapy), and the potential set consisted of symptomatic patients with minimal criteria that would suggest VWD disease (Table 1). Within the reference set 2 claims with associated VWD codes from separate visits were required to avoid “rule out” diagnoses, and patients were required to be enrolled for at least 2 years to further increase confidence in the diagnosis.21

During the development of the model, multiple clinical variables and patient characteristics were evaluated in order to identify those that contributed the most to the accuracy of the model in identifying an individual with severe VWD symptoms within the reference set. This was achieved by removing each variable in turn, and determining how much the accuracy of the model decreased as a result; the variable with the greatest effect on model accuracy was indexed to 100, and the other variables were then ranked according to the relative decrease in accuracy compared with this index value (mean decrease accuracy). By this process, a set of 12 variables were identified, that best defined the severely symptomatic patients with diagnosed VWD (Table 2).

**Application Of The Predictive Model**

The refined, 12-variable VWD model was applied to the potential set of undiagnosed patients who had what we considered to be the minimal criteria for a symptomatic VWD patient (Table 1). These patients were assigned scores based on their likelihood to have symptomatic VWD/bleeding disorder according to the model.

The final step was to determine the number of candidates having a strong likelihood of VWD/bleeding disorder, balancing the need to have both a sufficiently large number of candidates, and sufficient confidence in the classification of VWD/bleeding disorder. We identified 2 groups of patients from the potential set, based on their similarity to the VWD predictive model: “best fit” patients, in whom the positive predicted value (PPV) of the model was 83%, and “good fit” patients, in whom the PPV was 75%. Thus, we could have between 75% and 83% confidence that the patients in the potential set are similar to those severely symptomatic patients with a diagnosis of VWD, and are therefore likely to have a bleeding disorder that would warrant evaluation and potentially treatment.

**Estimation Of Patients With Potential VWD In The Commercially Insured US Population**

Best-fit patients were characterized by demographic, bleed type, and care setting. The number of symptomatic undiagnosed patients with VWD/bleeding disorders within the wider commercially insured US population was estimated by multiplying the size of the best-fit and good-fit populations by a projection factor of 10.4. This projection factor

---

**Table 1** Eligibility Criteria For Reference Set (Symptomatic, Diagnosed VWD Patients) And Potential Set (Undiagnosed Patients With Minimal Criteria To Define A Symptomatic VWD patient) For Unary Modeling

<table>
<thead>
<tr>
<th>Reference Set</th>
<th>Potential Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWD diagnosis ≥2 VWD medical claims or ≥2 claims for VWF use (excluding hemophilia A patients)</td>
<td>No VWD medical claim in history (or hemophilia A or other bleeding disorder)</td>
</tr>
<tr>
<td>≥24 months’ continuous enrollment in health plan prior to diagnosis</td>
<td>Continuous enrollment in most recent 24 months of data</td>
</tr>
</tbody>
</table>

**Additional criteria for severely symptomatic patient (users of or candidates for VWF replacement)**

- Within most recent 12 months, ≥1 of:
  - History of VWF replacement use after diagnosis
  - ER visit for ≥1 bleeding event in the 12 months before diagnosis
  - ≥1 joint bleed, GI bleed or bleed complicating a procedure in the 12 months before diagnosis
  - ≥2 bleeds within 12 months before diagnosis
  - Not using anticoagulants such as warfarin

**Additional criteria for all patients**

- Within most recent 12 months, ≥1 of:
  - ≥2 bleeding claims
  - ≥1 bleeding claim and a claim for a general bleeding disorder
  - ≥2 procedures to stop bleeding or 1 procedure to stop bleeding and 1 bleeding claim
  - ≥1 bleeding claim and a prescription claim for drugs to treat bleeding
  - Not using anticoagulants such as warfarin

**Abbreviations:** ER, emergency room; GI, gastrointestinal; VWD, von Willebrand disease; VWF, von Willebrand factor.
was based on the number of patients continuously enrolled in PharMetrics Plus between July 2014 and June 2015, the most recent year of the dataset (19,296,554 patients) and the total number of commercially insured patients in the US in 2013 (approximately 201 million patients, ie, 10.4 times the number of patients continuously enrolled in the dataset).22

**Results**

**Patient Populations**

A total of 10,420 patients (72% female, 28% male) met the criteria for the reference set of symptomatic patients diagnosed with VWD, of whom 3153 patients were severely symptomatic (users of, or considered to be candidates for, VWF replacement therapy). A total of 507,668 patients (86% female, 14% male) met the criteria for the potential set (symptomatic, undiagnosed patients). Figure 1A and B shows the sex and age distribution of patients in the reference and potential sets, respectively.

**Predictive Factors For VWD Diagnosis And Potential VWF Replacement Therapy**

The 12 variables that best defined the severely symptomatic subset of reference set patients are shown in Table 2. The variable that had the greatest effect on the accuracy of the model for the detection of potential bleeding disorders was the frequency of procedures to treat bleeding events, and this was indexed to 100. The next strongest predictive variables were younger patient age (mean decrease accuracy 83%), frequency of all bleeds (53%), frequency of GI bleeds (52%), sex (49%), and frequency of emergency room (ER) visits (45%).

**Identification And Characterization Of Undiagnosed, Symptomatic Patients With VWD Or Other Mucocutaneous Bleeding Disorder**

Application of the 12-variable model to the potential set identified 3318 best-fit patients (PPV 83%), and 37,163 good-fit patients (PPV 75%). Most (97%) best-fit patients were female, of whom 94% were less than 46 years of age, and 52% were between 31 and 45 years of age (Figure 1C). The most common bleed type claims among female best-fit patients were HMB (with claims from 49% of female patients), other bleeds (24%), and GI bleeds (21%) (Figure 2).

Female best-fit patients most commonly sought bleed care treatment from a hospitalist/ER physician (38% of patients), obstetrician/gynecologist (33%), or primary care physician (22%); 7% had visits to gastroenterologists, 5% to each of hematologists and pediatricians, and 3% to each of ear nose and throat specialists and surgical specialists. Treating specialist varied by patient age: hospitalist/ER physicians provided care for 43% and 32% of patients aged 16–30 and 31–45 years, respectively, and obstetrician/

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severely Symptomatic Diagnosed Patients</th>
<th>Non-Severely Symptomatic Diagnosed Patients</th>
<th>Mean Decrease Accuracy* (Indexed to 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of procedures to treat bleeding events, b</td>
<td>0.196</td>
<td>0.013</td>
<td>100</td>
</tr>
<tr>
<td>Age category, c</td>
<td>2.468</td>
<td>2.900</td>
<td>83</td>
</tr>
<tr>
<td>Total no. of bleeds, b</td>
<td>22.236</td>
<td>0.772</td>
<td>53</td>
</tr>
<tr>
<td>No. of GI bleeds, b</td>
<td>1.608</td>
<td>0.028</td>
<td>52</td>
</tr>
<tr>
<td>Male sex, %</td>
<td>17.1</td>
<td>30.8</td>
<td>49</td>
</tr>
<tr>
<td>No. of ER visits, b</td>
<td>1.190</td>
<td>0.015</td>
<td>45</td>
</tr>
<tr>
<td>Diagnostic tests for bleeding (all), b</td>
<td>64</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>Gastroenterologist visit for GI bleeds, b</td>
<td>3</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>PCP visit for GI bleed, b</td>
<td>2</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>No. of bleeds complicating a procedure, b</td>
<td>1.472</td>
<td>0.000</td>
<td>8</td>
</tr>
<tr>
<td>No. of joint bleeds, b</td>
<td>0.186</td>
<td>0.000</td>
<td>4</td>
</tr>
<tr>
<td>Hysterectomy procedure, b</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Notes: All variables are significant at a 1% probability level. *Mean decrease accuracy: relative contribution of each variable to the predictive model for VWD, compared with the variable having the greatest impact on the predictive accuracy (frequency of procedures to treat bleeding events), indexed to 100. †Over a 2-year timeframe. #Age categories: 1 (0–15 years), 2 (16–30 years), 3 (31–45 years), 4 (46–60 years), 5 (≥61 years).

Abbreviations: ER, emergency room; GI, gastrointestinal; PCP, primary care physician; VWD, von Willebrand disease.
gynecologists for 31% and 41%. Table 3 shows the percentage of visits by females of different ages to different health care settings.

Among the 101 male best-fit patients, most (66%) were aged 16–30 years (Figure 1C). The vast majority (98%) of male best-fit patients had GI bleed claims (Figure 2).

Estimated Number Of Patients With Undiagnosed Symptomatic VWD Or Other Mucocutaneous Bleeding Disorders

Projections gave an estimate of between approximately 35,000 (3318 best-fit patients×10.4) and 387,000 (37,163

---

**Figure 1** Age and sex distribution of (A) the reference set (symptomatic, diagnosed VWD patients); (B) the potential set (undiagnosed patients with minimal criteria to define a symptomatic VWD patient); and (C) undiagnosed patients who best fit the model for symptomatic VWD.

**Abbreviations:** HMB, heavy menstrual bleeding; VWD, von Willebrand disease.
good-fit patients × 10.4) patients with symptomatic undiagnosed VWD or other mucocutaneous bleeding disorder within the commercially insured US population.

**Discussion**

**Key Findings**

This analysis of medical insurance claims data with unary modeling indicated that up to 1 in 500 commercially insured individuals in the United States may have undiagnosed, symptomatic VWD or other mucocutaneous bleeding disorder. The patients in the potential set who provided the best fit for the VWD model were predominantly women of reproductive age, around half of whom had claims for HMB. This is perhaps unsurprising given that VWD may be disproportionately symptomatic in women of child-bearing age, and that HMB is a common presenting symptom: a systematic review reported a prevalence range of 5%–24% among women presenting with HMB, and the reported prevalence of VWD in adolescents with HMB ranges from 3%–36%, depending on the clinical setting. It is, however, noteworthy that HMB did not emerge as a predictor of potential VWD or other mucocutaneous bleeding disorders in this model. This may have been because HMB was a common feature among women in the reference set, which would decrease the value of HMB as a specific predictor of VWD in this combined model incorporating male and female data. It is possible that gender-specific analyses may reveal criteria (such as HMB) that may more accurately predict undiagnosed bleeding disorders within a single gender only.

**Study Strengths And Limitations**

The PharMetrics Plus database enabled analysis of large numbers of symptomatic diagnosed VWD patients to inform our predictive model, although there was no suitable database to externally validate the model. The database has a geographical bias (with a preponderance of patients from the northeastern United States) and is restricted to the commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commercially insured population; thus, findings may not be generalizable to other geographical or insurance populations. The analyses assume a similar age and gender distribution in the PharMetrics Plus population compared with the overall US commerci
example, is restricted to women of reproductive age). Our model did generate a wide range of estimates for the number of patients who may have undiagnosed symptomatic VWD or other mucocutaneous disorders, nevertheless, we believe even the most conservative estimate supports the need for further research, and to ensure clinicians are aware of this potentially unrecognized group of patients.

Our analysis focused on the prediction of severely symptomatic patients who might be candidates for VWF therapy, and as such may have underestimated the number of undiagnosed patients with milder symptoms (such as those with low VWF levels) who might benefit from diagnosis and specialist care. Indeed, identification of diagnosed patients (for generating the predictive model) relied in part on ICD-9 coding for VWD, which does not provide information on VWD type or distinguish “low VWF” from other conditions. Additionally, the lack of an ICD-9 code may not always mean a patient is undiagnosed, as occasionally a diagnosis code may be omitted from billing claims for patients with medical diagnoses. Miscoding is another potential limitation: for example, some patients may receive a VWD diagnostic code following laboratory evaluation, and retain this even if the diagnosis is not subsequently confirmed. It is also possible that our model may have included patients with other mucocutaneous bleeding disorders such as inherited or acquired platelet function defects or patients on anti-platelet drugs; and patients with chronic liver disease were not excluded.

For these reasons, we acknowledge that it is not possible to definitively state that the patients in the reference set had VWD, rather than another bleeding disorder with similar symptoms. However, we believe the requirement for at least 2 medical VWD or VWF claims should provide a degree of confidence in the VWD diagnosis; we cannot exclude the possibility that the undiagnosed patients all have VWD while sharing similar characteristics to patients with other mucocutaneous bleeding disorders. In the analysis of health care setting, it is possible that some hematologists were represented as hospitalists, owing to limited information within this dataset on physician specialty within hospital settings.

A further potential limitation is that the proportion of male patients identified as having potential bleeding disorders was lower than expected: men comprised only 3% of the best-fit individuals identified in the potential set. This may reflect the higher proportion of women in the reference set, which would be expected given that women are generally more highly represented in VWD diagnoses than men. This in turn may have resulted in characteristics that most confidently identify females with VWD or similar disorders playing a greater predictive role in the model. Similar analyses on single-sex populations might be necessary to estimate the potential number of undiagnosed VWD or other bleeding disorder cases among men.

The strength of this type of analysis is that it relies on real-world evidence of patient-level health care

Table 3 Care Setting Of Undiagnosed Female Patients Who Best Fit The Model For Symptomatic VWD, By Age. Data Represent Percentage Of Patients With Visits To Each Specialist Type In The Most Recent 2 Years Of Their Enrollment In The IMS PharMetrics Plus Database

|                | Hematologist | Ob/Gyn | Pediatrician | PCP | Emergency Room | ENT | GI | Surgical | Dentist |
|----------------|--------------|--------|--------------|-----|----------------|-----|----| Specialist |        |
| Female patients, % (N=3217) | 5 | 33 | 5 | 22 | 38 | 3 | 7 | 3 | 0 |
| 0–15 years (N=248) | 5 | 17 | 42 | 14 | 40 | 12 | 0 | 6 | 0 |
| 16–30 years (N=1101) | 4 | 31 | 4 | 19 | 43 | 4 | 4 | 2 | 0 |
| 31–45 years (N=1669) | 6 | 41 | 0 | 20 | 32 | 1 | 5 | 2 | 0 |
| 46–60 years (N=137) | 1 | 1 | 45 | 41 | 1 | 52 | 4 | 0 |
| ≥61 years (N=62) | 5 | 2 | 0 | 81 | 76 | 0 | 56 | 6 | 0 |

Abbreviations: ENT, ear, nose, throat specialist; GI, gastroenterologist; Ob/Gyn, obstetrician/gynecologist; PCP, primary care physician; VWD, von Willebrand disease.
utilization. We were able to compare the health care utilization history of VWD patients before diagnosis vs a group of potential patients to identify and quantify patients who best fit those patterns of health care utilization. Because this dataset represents such a large percentage of the commercially insured population it allowed for a broad sample to assess as the reference population and a broad sample to search for undiagnosed patients. Further, in this type of study we are limited to observation of patients’ claims histories, without sight of potentially important defining traits of significantly symptomatic VWD, such as the length of menstrual period or the number and type of bleeds with no associated health care utilization. We are also limited to observing the span of patients’ enrollments in the plan, which may not provide their entire symptom histories.

Implications For Education And Research

Diagnosis is important for appropriate management of VWD, and thereby for minimizing the burden of bleeding symptoms and complications on patients and the health care system.\(^3,9,14-16\) Our previous analysis showed that diagnosis of VWD was associated with improved patient outcomes in terms of a lower frequency of bleed claims after (vs before) diagnosis, but that many patients were misrecognized prior to diagnosis.\(^17\)

Our present findings, which suggest (based on computer modeling) that there may be a significant number of patients with symptomatic but undiagnosed VWD or other mucocutaneous bleeding disorders within the commercially insured population, underscore the need for continued health care practitioner education to increase VWD detection and diagnosis and are in line with recent calls for improved VWD community education.\(^25\) Educational efforts to increase awareness of VWD symptoms and evaluation should focus on hospitalists/ER physicians, obstetrician/gynecologists, and primary care physicians — particularly those managing HMB in women. Surveys of obstetricians/gynecologists suggest that while awareness of VWD as a possible cause of HMB has increased, there remains much opportunity for improvement.\(^26,27\) In a previous analysis of the PharMetrics Plus database, we showed that patients who were serially misrecognized before VWD diagnosis were predominantly young and middle-aged women, many of whom had claims for HMB.\(^17\)

Improvements in the recognition and diagnosis of VWD in adolescent girls are also needed.\(^28\) Despite evidence that VWD is among the most commonly encountered bleeding disorders in adolescents with abnormal uterine bleeding\(^29\) and that testing may be cost-effective,\(^30\) laboratory testing is performed infrequently\(^31,32\) and diagnosis may be delayed into adulthood.\(^33\) In 2001, the American College of Obstetricians and Gynecologists (ACOG) indicated that routine screening for coagulation disorders in adolescents with anovulatory bleeding was warranted, and that an underlying coagulopathy such as VWD should be considered in all patients (particularly adolescents) with abnormal uterine bleeding.\(^34\) In 2006, the American Academy of Pediatrics in collaboration with ACOG issued a position statement advising that hematologic disorders (particularly VWD) should be considered in females with HMB, especially at menarche.\(^35\) Prospective studies with objective assessment of menstrual flow and frequency, and standardized hemostatic testing and laboratory definition of VWD, are needed to determine the true frequency of VWD in adolescents.

Organizations such as the Foundation for Women & Girls with Blood Disorders play a vital role in provider education and raising awareness, with the aim of improving diagnosis and treatment of blood disorders in females. Dedicated multidisciplinary young women’s bleeding disorders clinics, with a core team of hemostasis experts and gynecologists, may provide improved care for adolescent and young women with HMB, and increase the rate of detection of previously undiagnosed bleeding disorders including VWD.\(^36\) There is also scope for raising awareness of VWD among patients themselves, who may dismiss their symptoms of HMB as “normal,”\(^37\) and for online campaigns that seek to improve awareness and knowledge of bleeding disorders (www.betteryouknow.org and https://letstalkperiod.ca/).\(^38,39\)

Our findings also highlight a need for increased awareness of VWD as a potential cause of other types of bleeding in both men and women, including GI bleeding, which was the predominant bleeding symptom among males and the second most common in females in this analysis. GI bleeding in VWD is often associated with angiodysplasia and investigators have demonstrated that qualitative defects of VWF alter angiogenesis and can lead to vascular malformations, particularly in the GI tract, leading to difficult-to-manage GI bleeding.\(^40\) It may soon be possible to prevent or slow the progression of angiodysplasia formation in patients with VWD, if identified before adulthood.

Conclusions

Computer modeling using medical insurance claims data suggests there may be a significant number of patients
with symptomatic, undiagnosed VWD or other mucocutaneous bleeding disorders among the US commercially insured population. Such analyses are associated with methodological limitations and require further clinical validation. For instance, these findings may inform further investigation using electronic medical records to identify such patients and refer for clinical examination. It is likely that a large proportion of undiagnosed patients are females of reproductive age with HMB who are not under hematology care and are unaware that they may have an underlying bleeding disorder. Lack of diagnosis prevents the timely initiation of appropriate treatment of VWD, increasing the burden of disease on patients and the health care system. Enhanced awareness among health care practitioners of VWD symptoms and their impact on health, and of screening and testing procedures, would help to increase the diagnosis rate of VWD and reduce the disease burden.

Data Sharing Statement
Data are the proprietary property of Baxalta US Inc., a member of the Takeda group of companies, Lexington, MA, USA and access can be requested by contacting their legal representative.

Acknowledgments
Medical writing support was provided by Mary Berrington, PhD, and Iain Patefield, MS, of Parexel, Hackensack, NJ, USA, and was funded by Baxalta US Inc., a member of the Takeda group of companies, Lexington, MA, USA. Sarah A. Hale, Baxalta US Inc., a member of the Takeda group of companies, Cambridge, MA, USA, and Emily Brouwer, Shire US Inc., a member of the Takeda group of companies, Cambridge, MA, USA, also provided input and reviewed the manuscript for scientific accuracy. Imrran Halari, Charles River Associates, Boston, MA, provided additional statistical analyses for the work, in addition to providing input and reviewing the manuscript drafts.

Author Contributions
All authors contributed to data analysis, drafting and revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work. Although employees of the sponsor were involved in the design, collection, analysis, interpretation, fact checking of information, and coordination and collation of comments, the content of this manuscript, the interpretation of the data, and the decision to submit the manuscript for publication in the Journal of Blood Medicine was made by the authors.

Funding
Research was funded by the sponsor, Baxalta US Inc., a member of the Takeda group of companies, Lexington, MA, USA. Data were obtained from the IMS PharMetrics Plus Health Plan Claims Database, IMS Health Incorporated, all rights reserved. Analyses undertaken by Charles River Associates, Boston, MA, including Dana Fallaize, Mike Roy, Michael Agne and Imrran Halari, and was funded by Baxalta US Inc., a member of the Takeda group of companies, Lexington, MA, USA.

Disclosure
AZ has received an investigator-initiated grant from Shire (a member of the Takeda group of companies). DF was an employee of Charles River Associates, which received consulting fees from Shire, at the time of the study and is currently an employee of UCB, Atlanta, GA, USA. RFS has acted as a paid consultant to Bayer, Shire, Novo Nordisk, Aptevo, Bioverativ, LFB, HEMA Biologics, and CSL Behring, and has received investigator-initiated grants from Bioverativ and Shire. He reports grants, personal fees from Genentech, Grifols, Kedrion, Takeda, and Octapharma, during the conduct of the study; also an investigator initiated grant (as well as personal consulting) from Octapharma, Genentech, Grifols, Kedrion, and Takeda. The authors report no other conflicts of interest in this work.

References
Sidonio et al


18. Elkan C, Noto K. Learning classifiers from only positive and unlabeled data Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 24–27 2008; Las Vegas, New York(NY, USA): Association for Computing Machinery; 2008:213–220.


