Betwixt and between

Gol, Janna

DOI:
10.33612/diss.195066742

Chapter 2

PROFSS: a screening tool for early identification of functional somatic symptoms

Gol JM, Burger H, Janssens KA, Slaets JP, Gans RO, Rosmalen JGM
Journal of Psychosomatic Research
2014; 77: 504-509

Abstract

Objective: To develop and validate a brief screening tool for predicting functional somatic symptoms (FSS) based on clinical and non-clinical information from the general practitioner referral letter, and to assess its inter-rater reliability.

Methods: The derivation sample consisted of 357 consecutive patients referred to an internal outpatient clinic by their general practitioner. Referral letters were scored for candidate predictors for the main outcome measure, which was a final diagnosis of FSS made by the internist. Logistic regression identified the following independent predictors: type of symptoms, somatic and psychiatric comorbidity, absence of abnormal physical findings by the general practitioner, previous specialist consultation, and the use of illness terminology. Temporal validation was performed in a cohort of 94 consecutive patients in whom predictors were scored by two independent raters.

Results: In both the derivation and validation sample, the discriminatory power of the model was good with areas under the receiver operating characteristic curves of 0.84 (95%confidence interval: 0.80-0.88) after bootstrapping and 0.82 (95%confidence interval: 0.73-0.91), respectively. Calibration of the models was excellent in both samples and the interobserver agreement in the validation sample was very good (intraclass coefficient: 0.82 (95%confidence interval: 0.75-0.88)). Based on this model, we constructed the brief screening tool PROFSS (Predicted Risk Of Functional Somatic Symptoms). PROFSS identified patient groups with risks of FSS ranging from 17% (95%CI: 10-26%) to 92% (95%CI:86-96%).

Conclusion: the presence of FSS can be predicted with the brief screening tool PROFSS, based on a limited set of items present in the general practitioner referral letter.

Introduction

Functional somatic symptoms, i.e. symptoms not conclusively explained by organic pathology, are a common and burdensome problem in medical outpatient's clinics. The prevalence is estimated between 25% and 50 % [1] and there is some evidence that this figure is rising [2]. Functional somatic symptoms are hazardous to patients, doctors and the health care system as a whole as they have shown to be associated with extensive but unproductive searches for a biomedical cause [3].

Nowadays specific integrated somatic and psychological treatment facilities are highly recommended for these patients [4,5]. It would be a major step forward if these facilities were offered during the first outpatient clinic contact as timely treatment could prevent symptom fixation, reduce the risk of iatrogenic physical and psychological harm, and reduce the high health care use and costs of this group of patients [6]. However, systematic methods to allocate patients to these specific facilities are lacking.

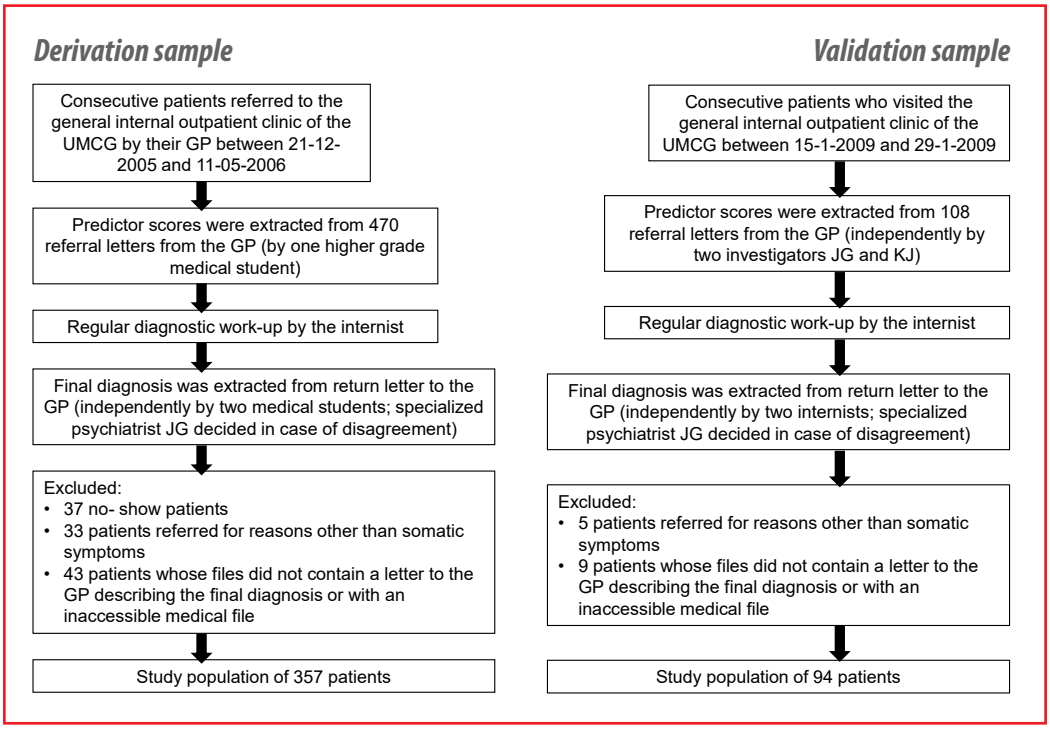
In the Netherlands, a referral letter from a general practitioner (GP) or medical specialist is required for access to a medical outpatient clinic. We analyzed data extracted from these referral letters for the development and validation of a prediction model to identify patients at the greatest risk of functional somatic symptoms [7]. The model was subsequently transformed into the brief screening tool PROFSS (Predicted Risk Of Functional Somatic Symptoms).

Study population and Methods

Study population

Figure 1 describes the two samples in which this study was performed. For derivation of the model, we performed a prospective cohort study including all consecutive patients referred to the general internal outpatient clinic of the University Medical Centre Groningen by their GP between 21-12-2005 and 11-05-2006. After excluding 37 no- show patients, 33 patients referred for reasons other than somatic symptoms, 43 patients whose files did not contain a letter to the GP describing the final diagnosis or with an inaccessible medical file, the study population consisted of 357 patients. For temporal validation of the model, we collected data on the first 108 consecutive patients in 2009 who were referred by their GP and visited the same general internal outpatient clinic of the University Medical Centre Groningen between 15-1-2009 and 29-1-2009. The final study population consisted of 94 patients; excluded were 5 patients referred for reasons other than somatic symptoms, and 9 patients whose files did not contain a letter to the GP describing the final diagnosis or with an inaccessible medical file.

Figure 1 Study samples (NB: this file is also uploaded separately for better quality)



Candidate predictors

All candidate predictors were extracted from the referral letters; no other data sources were used. For derivation of the model, the referral letters from the GPs were rated by four higher grade medical students on potential predictors of functional somatic symptoms before the patients were seen at the internal outpatient clinics. No letters were rated by more than one student. As candidate predictors, we started by selecting clinical and non-clinical indicators suggestive of functional somatic symptoms as described in the literature. They included female sex and younger age (less than 50) [1], somatic history defined as number of previous and co-existing somatic symptoms and diseases other than the referral symptoms in the previous 10 years (none, 1 or 2, 3 or more) and psychiatric comorbidity defined as history of a psychiatric diagnosis and / or treatment for psychiatric problems [8]. In addition, we analyzed type of the referral complaint as a potential predictor, categorized according to clusters found in previous studies [9,10]. We recorded the following symptom groups: (1). gastrointestinal, including general (nausea, diarrhea, feeling bloated etc) and pathognomic (rectal bleeding, jaundice etc), (2). musculoskeletal (joint pain, localized weakness, pain in extremities etc), (3). general (fatigue, dizziness, headache etc), [4]. other tracti (frequent urination, skin blotches, cardiovascular symptoms, lump in throat etc), 5 multi-system (combination of at least two of the following symptoms: general gastro-intestinal, musculoskeletal, and general symptoms), or 6. symptoms already diagnosed by GP. Furthermore, we scored predictors that we hypothesized to be associated with functional somatic symptoms based on our own clinical experience: absence of abnormal physical findings by the GP including laboratory test results, and chronicity of the complaint, defined as duration of at least 6 months. Previous health care consumption for the referral complaint(s) was scored as one or more consultations of the GP or consultation of a medical specialist

We also included a potential predictor linked to the GP. We hypothesized that a GP would be inclined to use disease (medical discourse) terminology in case he suspects biomedical pathology. If the symptoms are suspected to be functional, referral would mainly satisfy the needs of the patient, which is reflected in the use of illness (lay discourse) terminology[11]. For example, we scored high blood pressure and tiredness as illness terms, while hypertension and fibromyalgia were scored as disease terms. In case of doubt we used a standard Dutch dictionary.

In the validation study, referral letters were rated in an identical way by two investigators (JG

and KJ) independent of each other. Categorization of variables that were not dichotomously recorded, i.e. age, number of other symptoms and health care consumption, was done prior to the analyses based on their distributional properties.

Diagnostic work-up and outcome assessment

After referral, patients underwent the regular diagnostic work-up by the internist ending with a final diagnosis. The internists involved in the work-up were unaware of the hypotheses and were blind to the predictor ratings.

For derivation of the model, two higher grade medical students scored the final diagnosis on the basis of the return letter in the electronic file of the patient to the GP as a general medical diagnosis or no general medical diagnosis, i.e. functional somatic symptoms. A third medical student independently scored all diagnoses again. In all 128 (36%) cases of disagreement, a psychiatrist specialized in somatoform disorders made the final decision. The latter diagnosis served as the outcome variable. With regard to the criteria, the final diagnosis no functional somatic symptoms was for those patients who were diagnosed with a general medical diagnosis or whose complaints were clearly explained in terms of somatic matters judged by two independent raters. All other patients, including those with functional somatic syndromes such as chronic fatigue syndrome and irritable bowel syndrome, received a final diagnosis of functional somatic symptoms. Scoring of the final diagnosis was done by different students than those who scored the GP's referral letters and who were unaware of the hypothesis. In the validation study, two internists scored independently the final diagnosis in the same way as in the derivation study, blind to the predictor ratings. In all 19 (20%) cases of disagreement, a psychiatrist specialized in somatoform disorders made the final decision.

Statistical analysis

Complete case analysis was performed in view of the small proportion (1.7 %) of patients with missing values.

Derivation of the prediction model

We first tested whether patients with and without functional somatic symptoms differed on the predictor variables using Chi2 or Fisher's exact tests. However, as it has been shown that pre-selection of predictors based on statistical significance in univariable analyses may result in unstable prediction models, we included all potential predictors as independent

variables in a multivariable logistic regression model [12]. In this model, a final diagnosis of functional somatic symptoms served as the binary outcome variable. Dummy variables were made for each predictor variable with more than two categories. For the symptom groups, category 6 (diagnosis made by GP) served as the reference group. For the ordinal variable indicating the number of other symptoms variable the reference category was the highest, i.e. 'three or more', in order to obtain risk increasing predictors only. For the health care consumption variable, single consultation of GP was the reference.

Using a backward stepwise algorithm, the model was reduced by eliminating variables with p-values greater than 0.25, based on the likelihood ratio test. Values higher than 0.05 for alpha are commonly used in prediction research to limit bias in the estimation of predictor coefficients [13, 14].

The predictive accuracy of the resulting model was assessed in terms of discrimination and calibration. Discrimination is the potential of the model predictions to separate those with functional somatic symptoms from those without and an overall measure of this ability was obtained by calculating the Area Under the Receiver Operating Characteristic (ROC)-curve (AUC). Calibration refers to the agreement between categories of predicted risk and observed frequencies of functional somatic symptoms. It was assessed graphically by plotting the observed frequencies by deciles of predicted risk. Calibration was further evaluated by calculating the Hosmer Lemeshow (H-L) test which is statistically significant in case of poor goodness of model fit, i.e. poor calibration.

Both discrimination and calibration were assessed after correcting the model for over fitting or overoptimism using a bootstrapping procedure [15]. In this procedure 100 random bootstrap samples were drawn with replacement and the logistic model was fitted in each sample. From these repeated analyses a shrinkage factor was calculated to correct the logistic regression coefficients and the AUC for overoptimism.

Validation of the prediction model

Temporal validation of the derived model was performed in terms of calibration and discrimination. Calibration was evaluated by comparing the predicted risks averaged over the two observers from the model to the actual observed frequencies of functional somatic symptoms in the validation cohort. This was done graphically and by performing the H-L test. Discriminatory power of the model in the validation cohort was quantified as the AUC. We decided prior to the analyses that the temporal validation of the model was inadequate when the H-L test was statistically significant, or when the AUC decreased by more than 10% [14]. Using data from the validation cohort, we additionally tested the reliability of

the predictor ratings in terms of interobserver agreement. This was done by feeding the ratings of the two observers into the final model and calculating the intraclasscorrelation coefficient between the two predicted risks.

The PROFSS screening tool

To obtain a practical score, we calculated for each predictor a number of points. This was done by dividing each beta coefficient from the logistic model by the absolute smallest beta coefficient obtained from the multivariable regression and rounding it these numbers to the nearest integer. Next, a total risk score for each patient was calculated by summing the number of points for each predictor present. The relation between the total risk score and the model predictions was depicted in a graph. For practical application, we defined four categories of score values corresponding to clinically sensible categories of predicted risks, and related these to the observed frequencies of functional somatic symptoms. These calculations were performed using the combined derivation and validation cohorts to acquire maximum reliability.

Analyses were conducted in PASW Statistics 18.0 and TIBCO Spotfire S+ 8.1.

Results

Study population

Characteristics of the study population according to a final diagnosis of functional somatic symptoms in the derivation and validation cohort are presented in table 1. In the derivation cohort, patients with functional somatic symptoms are more likely than patients with no to be of a younger age, to show general and multi-system symptoms, to have no abnormal physical findings on examination by the GP, to have a history of psychiatric problems, and to be referred with a letter written in illness terminology symptoms. These associations were similar in the validation cohort, although at a lower statistical significance level.

Table 1 *Characteristics of the study population according to a final diagnosis of functional somatic symptoms (FSS)*

	Derivation cohort (N=357)			Validation cohort (N=94)		
Predictor (%)	FSS N=214 (59.9%)	No FSS N=143 (40.1%)	P-value	FSS N=60 (63,4%)	No FSS N=34 (36.5%)	P-value
Female gender	65.9	62.9	0.57	66.1	50.0	0.13
Age < 50 years	65.9	46.2	<0.01	65.0	44.1	0.05
Referral symptom group						
- Gastrointestinal	25.2	20.3	0.28	16.7	11.8	0.76*
- Musculoskeletal	6.1	7.0	0.73	3.3	2.9	1.00 *
- General	24.3	3.5	<0.01	36.7	2.9	<0.01
- Other tracti	18.2	19.6	0.75	18.3.	23.5	0.55
- Multi-system	16.4	1.4	<0.01	20.0	8.8	0.16
- Diagnosed by GP	9.8	48.3	<0.01	5.0	50.0	<0.01
Somatic comorbidity			0.05			0.19
- None	22.4	33.6		36.7	20.6	
- One or two	33.6	35.7		35.0	35.3	
- Three or more	32.7	42.0		28.3	44.1	
Longstanding symptoms (> 6 months)	45.8	42.7	0.56	-$	-$	-$
No abnormal findings on examination by GP	56.1	16.8	<0.01	60.1	23.5	0.01
History of psychiatric diagnosis or treatment	21.9	4.9	<0.01	28.3	17.6	0.25
Health care consumption			0.07			0.19
- Single consultation GP	21.5	23.8		61.7	70.6	
- Multiple consultations GP	38.8	48. 3		26.5	23.3	
- Consultation(s) medical specialist	39.7	28.0		15.0	2.9	
Referral letter written in Illness terminology	70.1	21.0	<0.01	69.5	44.1	0.02

P-values are from Chi2, or Fisher exact tests when indicated not included in the scoring procedure$*

Derivation of the prediction model

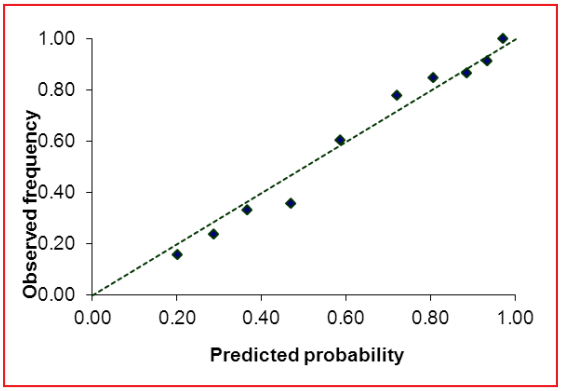
Table 2 shows the reduced multivariable model containing independent predictors with p-values less than 0.25 only. The model was corrected for overoptimism by applying a shrinkage factor from the bootstrapping procedure of size 0.87. The most important predictors were symptom characteristics, especially multi-system symptoms, followed by general symptoms. Other important predictors were no abnormal findings upon examination by the GP, and a referral letter written in illness terminology. The final logistic regression model showed good discriminatory power with an AUC of 0.84 (95%CI: 0.80 - 0.88) after correction for overoptimism. The close agreement between predicted risk and observed frequency of functional somatic symptoms shows that the calibration of the model was excellent (Figure 2) and the H-L test was statistically not significant (P=0.75) indicating no deviance of goodness of fit.

Table 2 *Reduced final multivariable model to predict functional somatic symptoms*

Predictor	Odds ratio (95% confidence interval)	Regression coefficient	Number of points
Intercept (constant)		-1.659	0
Female gender	1.45 (0.81-2.58)	0.371	1
Referral symptom group*			
- Gastrointestinal	2.10 (1.03 – 4.26)	0.7419	2
- General	7.02 (2.34 – 21.03)	1.948	5
- Other tracti	2.06 (1.10 - 4.17)	0.722	2
- Multi system	13.06 (2.76 – 61.81)	2.570	6
Somatic comorbidity			
- None	2.19 (1.19 – 4.04.)	0.786	2
No abnormal physical findings on examination by GP	2.53(1.35 – 4.73)	0.927	2
History of psychiatric diagnosis or treatment	2.28 (0.85 - 6.10)	0.825	2
Referral letter written in illness terminology	3.03 (1.64 – 5.59)	1.108	3

**The group diagnosed by the GP served as reference category.*

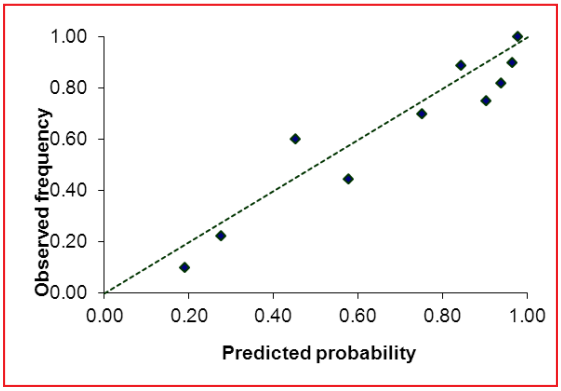
Figure 2 *Calibration plot of the final logistic model in the derivation population, depicting the observed frequencies of functional somatic symptoms by deciles of predicted risk according to the PROFSS tool.*



Validation of the prediction model

The temporal validity of the model was adequate with a less than 10% reduction in the AUC to 0.82 (95%CI: 0.73 - 0.91) and good calibration as shown in Figure 3 (P-value from H-L test:0.46). The reliability of the predictor ratings in terms of their interobserver agreement was good with an intraclass coefficient of 0.82 (95%CI: 0.75 - 0.88).

Figure 3 *Calibration plot of the final logistic model in the validation population, depicting the observed frequencies of functional somatic symptoms by deciles of predicted risk according to the PROFSS tool.*



The PROFSS screening tool

For each predictor in the reduced multivariable regression model, the number of points based on its regression coefficient is presented in Table 2. A total risk score can be calculated summing the number of points for each predictor present and relating the number to the predicted probability using Figure 4.

Figure 4 *Relationship between the total risk score as determined by the PROFSS tool, and predicted probability of functional somatic symptoms (FSS) in the combined derivation and validation cohort.*

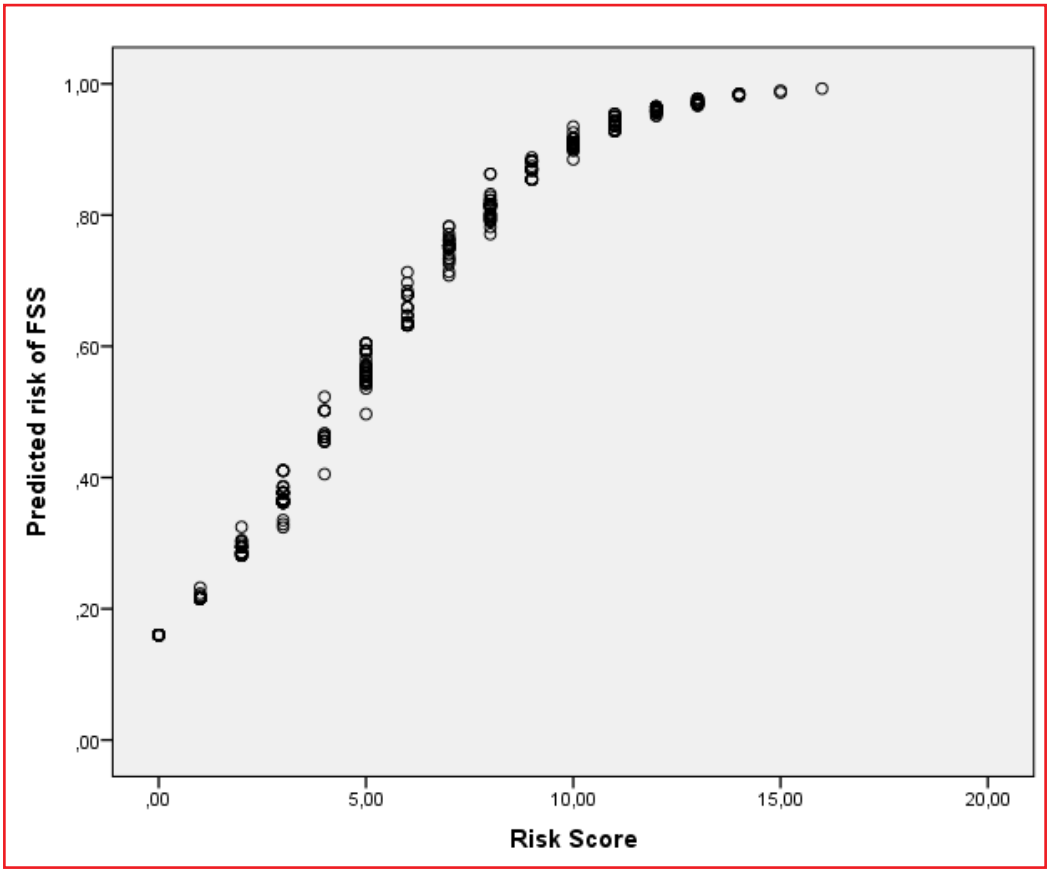


Table 3 shows the predicted probabilities and observed frequencies of functional somatic symptoms in the combined derivation and validation population according to categories of total score. In the lowest category, the risk is lower than a quarter. A total PROFSS score of 10 points or more is associated with a probability of 92 % for functional somatic symptoms.

Table 3 *Risk categories for functional somatic symptoms in the combined derivation and validation cohorts.*

Risk category (PROFSS score range)	Mean predicted probabilities, % (min; max)	Observed frequencies N / Total
Very low (<3)	23 (16; 32)	17/98 (17%; 95%CI: 10 – 26%)
Low (3-5)	47 (32; 60)	49/110 (45%; 95%CI: 35 – 54%)
Moderate (6-9)	77 (63; 89)	93/118 (79%; 95%CI: 70 – 86%)
High (>9)	94 (88; 99)	115/125 (92%; 95%CI: 86 – 96%)
Total	63 (16; 99)	274/451 (61%; 95%CI: 56-65%)

Confidence intervals are exact 95% binomial confidence intervals (95%CI). PROFSS: Predicted Risk Of Functional Symptoms .

Alternatively, the exact individual predicted probability of functional somatic symptoms can be calculated using the following formula which is based on the coefficients from table 2: 1 / (1 + exp (-(-1.659 +0.371*female gender + 0.741*gastro-intestinal + 1.948*general + 0.722*other tracti+ 2.570*multi-system + 0.786*no somatic comorbidity + 0.927*no abnormal findings + 0.825*psychiatric history + 1.108*illness terminology))). Using a cut-off score corresponding to the highest risk category, i.e. 10 or over, the sensitivity of this dichotomized score was 42% (95%CI: 36 – 48%), and the specificity was 94% (95%CI: 90 – 97).

Discussion

This study shows that the PROFSS screening tool accurately predicts the presence or absence of functional somatic symptoms based on a limited set of items that can be derived from the GP referral letter.

Strength of the study is its naturalistic design: it made use of already existing data that are routinely available in hospital settings. A further strength is the large sample size of the derivation study. Given the rule of the thumb of at least ten outcomes per predictor considered, our derivation study had more than enough statistical power. A limitation of our study is the relatively small validation sample, which resulted in similar results but with a lower statistical significance level. Further, the validation of our model was temporal only and many characteristics of the patients as well as their doctors were likely similar between derivation and validation cohort. However, the validation was independent of the original data and development process and yielded excellent results, despite the fact that it was done three years after derivation. Nevertheless, external validation in other general internal outpatient samples is required before the PROFSS screening tool can be widely used. Second, the disagreement on the end diagnosis between raters was relatively large. This is partly a methodological artefact, since the second rating of the end diagnoses in the derivation cohort was done after the first rating, and more information might have been available for the second rater. The ratings in the validation cohort were performed simultaneously, and the disagreement there is indeed lower. It also reflects the doubts of the internal medicine specialist who did the diagnostic work-up, and thereby the inherent difficulty to distinguish functional somatic symptoms from general medical conditions.

To the best of our knowledge, we are the first to use items from GP referral letters to develop a screening tool for functional somatic symptoms. The performance of the PROFSS screening tool is comparable in terms of positive predictive value to that of self-administered screening instruments to detect somatoform symptoms and somatoform disorders such as the Patient Health Questionnaire-15 (PHQ-15) [16].

Contrary to what we expected, our data indicated that counting the number of previous and co/existing somatic symptoms and diagnoses mentioned in referral letters is not a good indicator of risk of current functional somatic symptoms. In fact, previous and co-existing somatic symptoms and diagnoses mentioned in the referral letter appeared to be indicative

for a general medical diagnosis. Most likely, we measured with this item the overall physical weakness of the referred patients instead of presence of numerous functional somatic symptoms. It could be speculated that the GP is more inclined to give an accurate recording of the patient somatic history in the case of assumed organic pathology, thinking this would contribute to the diagnostic process of the medical specialist. In accordance with the literature, patients referred for symptoms in multiple organ systems have an increased chance of an end diagnosis of functional somatic symptoms.

Remarkable is that the "unknown" GP characteristic, the terminology in which the letter is written, i.e. contextual information, is a comparable predictor qua strength to the already in previous studies reported patient' characteristics such as psychiatric comorbidity and abnormal physical findings. This is important for policy makers in light of the recent developments with regard to standardized GP referral letters. Recent research on general practitioner' referral letters focuses on improving the content of the letters by recommending the use of standardized letters [17, 18]. Our results suggest that valuable contextual information will be lost by standardizing GP referral letters, since standardization restrict the freedom for the GP to make his or her own selection and presentation of information. The question arises how to implement this tool in clinical practice. The tool might be useful in countries with a GP system comparable to the Netherlands, since it depends on the presence of referral letters. Because the referral letter is the only source needed to calculate the score, the PROFSS screening tool may allow efficient allocation of patients to integrated treatment facilities from the first patient contact onwards. Also in the absence of these facilities the score could be informative, since it could alert the internist of the high likelihood of functional somatic symptoms in referred patients. With this prior knowledge, the internist will be able to adopt his approach, and opt for a bio-psycho-social approach of the patients' symptoms from the first contact onwards. This would make psychological interventions more acceptable after somatic causes have been excluded. The PROFSS tool is certainly not developed to reject admitting the patient with a high score, because general practitioners may need a second opinion to confirm their diagnosis of a functional condition.

In our study, medical students with little clinical experience obtained the same scores as experienced clinicians. Possibly, also medically administrative personnel would be able to use this simple selection tool after a short instruction. However, the implementation of this tool in clinical practice needs further investigation.
In addition to being practically applicable, the PROFSS tool demonstrated fairly good

performance. At the highest cut-off value that we proposed, the PROFSS tool had a high specificity, meaning that few patients with a general medical diagnosis will be falsely classified as having a high likelihood of functional somatic symptoms. With 42%, the sensitivity of our tool at the highest cut-off was relatively low implicating that 58 % of patients with functional somatic symptoms will be classified as low risk and consequently will receive standard work-up. This constrains the efficiency gain of the PROFSS screening tool.

In conclusion, functional somatic symptoms can be predicted using the PROFSS screening tool based on a limited set of items present in the GP referral letter. The PROFSS screening tool enables to allocate patients to appropriate treatment facilities from the first patient contact onwards. Future studies could investigate whether slightly modified versions of the PROFSS screening tool are useful in other general outpatient clinics.