Blood test ordering for unexplained complaints in general practice. Results of the VAMPIRE study on diagnosis and prognosis

Koch, H.

Citation for published version (APA):
Chapter 10

General discussion
General discussion

Introduction to the general discussion

This thesis focused on patients with newly presented unexplained complaints and the role of blood test ordering for these patients in general practice. Unexplained complaints were defined as those complaints that remain unexplained for the GP after initial, adequate history taking and physical examination. This definition has been taken from the Dutch College of General Practitioners’ (DCGP) guideline on blood test ordering, which advises to postpone blood test ordering for patients with newly presented unexplained complaints, and to restrict the number of tests ordered after initial postponement. As it happens, blood tests are often ordered for patients with newly presented unexplained complaints. The diagnostic yield of these tests was presumed to be limited since the prevalence of any serious somatic disease in this population is low. However, exact data on the diagnostic yield were lacking. Little was also known about the course of newly presented unexplained complaints. In this thesis, characteristics of patients with newly presented unexplained complaints, the diagnostic yield of immediate and postponed blood test ordering as well as the course of such complaints in general practice and its determinants are described. This final chapter presents the main findings of the study per chapter, in the light of methodological issues. Finally, implications for daily practice and future research are discussed.

Terminology used in literature and in this thesis

The literature about unexplained complaints is not always consistent when it comes to the conceptualization and definition of the subject of study. Terms as varied as ‘medically unexplained symptoms’, ‘functional symptoms or ‘syndromes’ and ‘somatoform disorders’ appear to be used as synonyms. We struggled with terminology as well, because we aimed to focus on the unexplained complaints that patients present with for the first time, rather than the unexplained complaints at the other end of the spectrum, of complaints remaining unexplained over time. There is hardly any literature about this initial phase of episodes with unexplained complaints. The terminology used throughout this thesis is varied as well, and the term ‘unexplained complaints’ may be preceded by e.g. early stage, early phase, new and chronic. We currently prefer the term ‘newly presented unexplained complaint’ for complaints for which no additional diagnostic procedures have been performed as yet and ‘established unexplained complaints’ when both time and additional diagnostic procedures have not helped to establish a diagnosis or explanation.
Unexplained complaints

In the VAgue Medical Problems In REsearch (VAMPIRE) study, general practitioners were asked to include patients presenting one of five types of unexplained complaints: fatigue, abdominal complaints, musculoskeletal complaints, weight changes and itch. In total 513 such patients were included: 325 with unexplained fatigue, 77 with musculoskeletal complaints, 66 with abdominal complaints, 27 with weight changes and 18 with itch. Unsurprisingly, patients with unexplained fatigue were included most often. Fatigue is an even more indefinite complaint than, for example, unexplained abdominal or musculoskeletal complaints. Opinions differ on what to ask, examine and test, since an evidence-based primary care guideline about newly presented fatigue does not exist. Fatigue, though often a non-alarming symptom in patients in general practice, might also be the harbinger of cancer, renal impairment or cardiac failure. Therefore, patients with unexplained fatigue in particular confront GPs with the uncertainties in medical decision-making. This uncertainty is known to be linked to the GPs’ ‘sense of alarm’, and a ‘sense of alarm’ may trigger GPs to take further action to resolve these uneasy feelings, for example by ordering blood tests. This in turn might start a cascade of additional diagnostic procedures, redundant medication or even referral to secondary care, hospitalisation and surgery.

Pre-VAMPIRE test ordering behaviour

Before the start of the trial, we investigated the normal test ordering behaviour of GPs in patients with newly presented unexplained complaints (chapter 4). In our search for determinants influencing the physicians’ decision to order blood tests in patients with newly presented unexplained complaints, we found that fatigue (as compared to abdominal complaints), duration of the complaint for over four weeks, the absence of possible explaining psychosocial factors, and the absence of a syndrome-type of working hypothesis were independently associated with the chances of blood tests being ordered. When evaluating these outcomes in the light of the DCGP guideline on blood test ordering, it is striking that GPs actually seem to heed the guideline’s postponement advice. Duration of the complaints before presentation for over four weeks as a determinant for the decision to order blood tests seems to be in line with postponing blood test ordering for four weeks. Most GPs also take the psychosocial circumstances of their patients into consideration. When there are clues for psychosocial factors influencing the patients’ complaint, GPs are less inclined to order blood tests than if these clues are absent. Finally, when GPs consider a syndrome (more than a mere description of the symptoms presented by the patient (e.g. irritable bowel syndrome or fibromyalgia)) as a working hypothesis, they appear to be less inclined to order blood tests than when an
isolated unexplained symptom dominates their evaluation. A syndrome type of diagnosis might appear to be clearer and therefore less unexplained than a symptom-type diagnosis. The unexplainedness perceived by the GP of a symptom-type diagnosis may lead to diagnostic uncertainty in GPs, which in turn is known to trigger a tendency to act, e.g. by ordering blood tests.

Characteristics at presentation

Our findings in chapter 5 indicate that patients with newly presented unexplained complaints are mainly women in their forties, with secondary or higher-level education. In another Dutch study on unexplained complaints, the majority of patients was female as well, although they were slightly older than our population (the average age was 56). In this study, in addition to the cluster 'pain', most patients also presented symptoms within the cluster 'fatigue'. An explanation of these findings may be that women with a high educational level in this phase of life may struggle with finding a sound balance between private and working life. This struggle may present a heavy load, associated with experiencing and expressing physical complaints for which no somatic explanation can be found. It is known, for example, that among working women, working more hours a week is associated with poorer self-perceived health and a higher probability of medical visits. Also, working women living in family units of more than three persons have a higher risk of poor self-perceived health status and of psychosomatic symptoms. It might also be because of selection bias that predominantly higher educated women were included. This topic is discussed further on in this general discussion.

Quality of life

Surprisingly, our patients turned out to have an overall poor quality of life at presentation (chapter 5). Therefore, we decided, post hoc, to compare their quality of life figures. Compared to a general GP population, patients with newly presented unexplained complaints had poorer quality of life. To some extent, their quality of life was even worse than that of depressed patients. These findings do not match the assumption that newly presented unexplained complaints are supposed to be mild. Though overall poor, especially physical impairments due to the complaints are responsible for the poor quality of life. In our opinion, GPs should therefore pay more attention to the personal experience of impairment of patients with new unexplained complaints. One way to achieve more optimal monitoring of these patients in general practice is to advise all patients with newly presented unexplained complaints to revisit their GPs, e.g. after four weeks, to see whether the complaints persist and remain unexplained. This means not just instructing patients to revisit when their complaints do not resolve within these four weeks, as suggested in the
DCGP guideline. For longer lasting unexplained complaints, the course should be monitored even more carefully. Longer lasting, established unexplained complaints may benefit from cognitive behavioural therapy\(^\text{11,12}\), which can of course only be started when patients are monitored. We did not explore the etiological relationship between unexplained complaints and quality of life in depth. It is possible that patients with a perceived poor quality of life are prone to contract unexplained complaints. The poor quality of life might also be explained by depression. Unfortunately we did not use depression or anxiety scales, since such a poor quality of life was not anticipated at the start of our study. Therefore we do not know whether the complaints of our patients were an expression of the presence of a major depression.

The diagnostic value of blood test ordering

In chapter 5, we showed that patients with newly presented unexplained complaints already have a poor quality of life at the beginning of their complaint episode. From other research it is known that the long-term consequences of unexplained complaints are unfavourable in terms of patients’ quality of life, health care use and absence from work due to illness\(^\text{13-15}\). Therefore, GPs are in need of tools to distinguish patients with newly presented unexplained complaints that are likely to vanish in due time from those with unexplained complaints that turn out to be caused by somatic or psychosocial illnesses or who run the risk of developing established unexplained complaints. One of the diagnostic tools could be ordering blood tests to detect somatic illnesses. However, immediate blood test ordering has its disadvantages because of the high risk of false-positive test results in these patients due to the low prior probability of somatic pathology. In order to overcome these disadvantages, the DCGP guideline recommends a four-week postponement of blood test ordering and a restriction of the number of blood test ordered. GPs are not always inclined to follow this recommendation, because they consider immediate test ordering more attractive than postponement\(^\text{2,16-19}\). After studying the normal test ordering behaviour of GPs in chapter 4, we studied the effects of immediate versus postponed blood test ordering in a randomised trial (chapter 6). Apart from the moment of test ordering, we also explored differences in the outcome when ordering a restricted versus an extensive blood test set. Though only a small number of patients with unexplained fatigue revisited their GPs after four weeks, we did not find strikingly different diagnoses or abnormal blood test results in the postponement group of patients compared to the immediate test-ordering group. We also found that the limited DCGP set works as good as or even better than the extensive fatigue-specific one. Unfortunately, we were not able to present accuracy measures such as sensitivity/specificity or predictive values of the blood tests, as we originally aimed for. This was due to
a methodological problem that became apparent during the study. Most of the diagnoses from the category ‘somatic disorder that can be confirmed by blood test ordering’, the category for which accuracy measures might have been useful, were actually defined by the abnormal test results themselves (anaemia-low Hb; diabetes-high glucose; hypothyroidism-high TSH). Due to this incorporation bias it would not have been scientifically sound to calculate accuracy measures. Still, we conclude that there are no firm arguments against postponement of blood test ordering or against restriction of the postponed test set for patients with unexplained fatigue.

Other effects of postponed blood test ordering

Though from a diagnostic point of view the postponement of blood test ordering turned out to be safe, several other reasons why GPs would want to order blood tests immediately are still valid. These reasons may be the presumed infeasibility of postponement because of patient expectations or presumed dissatisfaction, and anxiety among patients when postponing blood test ordering. Furthermore, the fact that some disadvantages of immediate blood test ordering, such as higher costs, are invisible, can be a reason. Studies from our research group, presented in the thesis of Van Bokhoven, focused on these other aspects of the postponement of blood test ordering. The postponement of blood test ordering was found to be a feasible diagnostic approach in patients with newly presented unexplained complaints and did not lead to testing at a later moment.

The costs of testing were, as expected, statistically significantly lower when blood test ordering was postponed. The differences in other health care related costs between immediate and postponed blood test ordering were not statistically significant. However, the costs of diagnostics and therapy in the twelve-month period after the initial consultation tended to be higher for the immediate test ordering group. Although the overall difference in costs was not very high in absolute terms, the large numbers of patients with unexplained complaints may cause high overall costs of care if GPs frequently order blood tests immediately. Contrary to our hypothesis, postponement did not have an unfavourable influence on patients’ satisfaction with or anxiety after the consultation (chapter 7). Communication aspects of the consultation, however, did influence these aspects. Our studies suggest that a mismatch between GPs’ and patients’ views of the severity of the complaints results in lower satisfaction and higher anxiety among patients.
The course of complaints and of the quality of life of patients with newly presented unexplained complaints

In chapter 8 we showed that most newly presented unexplained complaints remained unexplained during the one-year follow-up period. Of the unexplained complaints that became explained during this period, the majority had a somatic origin, followed by psychosocial explanations. We defined syndrome-type diagnoses (e.g. chronic fatigue syndrome, irritable bowel syndrome) as explained complaints. In addition, the few somatoform disorder diagnoses that were found were categorized as explained. If we had categorized these diagnoses as established unexplained complaints, the percentage of newly presented unexplained complaints becoming established unexplained complaints would have been even higher (64% instead of 57% of the analysed patients).

In this same chapter 8, VAMPIRE showed that unexplained complaints are not as frequently self-limiting as is assumed. Whereas Kroenke and Kenter, for example, found that 90% of unexplained complaints are self-limiting, we found that over 50% of patients still experienced their unexplained complaints after one year\textsuperscript{11,23}. One explanation for this difference might be that the other studies considered complaints to be resolved if patients did not revisit their GPs. Our study showed that although patients frequently did not revisit their GP for their unexplained complaints, when asked after one year, over half of them reported to still or once again perceive complaints. We found that longer duration of the complaint before presentation was associated with lower probability of it resolving. Patients with new UCs tend to have a poor quality of life; we did not find strong determinants that influence this tendency. Especially experiencing musculoskeletal complaints and having had previous episodes of the same complaints were associated with a negative course of the quality of life on the physical dimension. The longitudinal course of the mental summary scale of quality of life was predominantly negatively influenced by experienced fatigue, having had previous episodes of the same complaints and the presence of psychosocial factors.

Because newly presented unexplained complaints often remain unexplained and hinder patients over time and because of the unfavourable prospects (among others high health care facility use and impairment of quality of life) once patients have long-lasting, established unexplained complaints, the evolution from newly presented into established unexplained complaints deserves attention. Our attempts to predict at baseline which patients are at risk of developing established unexplained complaints have not been fruitful, unfortunately. Otherwise, it might have created possibilities for research about whether early treatment, e.g. earlier cognitive behavioural therapy, could prevent evolution from newly presented into established unexplained complaints.
In our opinion, the above-mentioned findings on course and prognosis should stimulate GPs to take unexplained complaints more seriously as from their initial presentation. Paying attention to the determinants we found, may be helpful in early detecting patients with an unfavourable course of complaints and quality of life among the patients with newly presented UCs in general practice.

**Innovative new tests: Carbohydrate deficient transferrin for example**

Frequently, new tests become available and are presented with many promises concerning their value. Quite often these tests, based on research in secondary or tertiary care, are also implemented in primary care. However, if the disease in question has a low prevalence in primary care, the diagnostic yield in terms of predictive values of the test for the primary care situation will probably be less imposing. An example of such a test was the carbohydrate deficiency test (CDT), the accuracy of which in detecting excessive alcohol consumption was reviewed systematically in chapter 9. The reason why we were interested in this laboratory test was that excessive alcohol consumption can be one of the (hidden) causes of complaints that initially are labelled unexplained. We concluded that the validity of CDT as a diagnostic tool in primary care is still questionable. If others can confirm the higher values for sensitivity that some studies report, it might be a useful diagnostic tool in unselected populations. However, more methodologically sound, comparable studies are needed before firm conclusions can be drawn. Our findings also illustrate the fact that one never knows how a test will perform in unselected broad-spectrum populations. Moreover, CDT tests yielded a large part of the false-positive test results in our population. In the years since publication of this chapter, various studies concerning CDT have been published. Bortolotti reviewed some later data on CDT tests. Sensitivities still fluctuated among studies and tend to be low. Found specificities are mostly 60% and up. In this review the literature search was not restricted to unselected populations. The figures showing a higher level of accuracy came from selected populations. More recent literature does therefore not alter our conclusions concerning the accuracy of CDT in detecting excessive alcohol consumption in unselected populations. We did not review the accuracy of CDT in forensic situations, such as at issuing driving licences to persons who have been driving under the influence of alcohol, a setting in which CDT is used in the Netherlands at present. This a setting with a selected population (drunk drivers) and thus our doubts about the validity of CDT for a primary care setting do not conflict with the use of CDT as a diagnostic tool in this specific situation.
Methodological issues

GP and patient recruitment

It was difficult to recruit a sufficient number of participating GPs for the randomised clinical trial in which we evaluated the postponement of blood test ordering. Recruitment of GPs took place in a period in which GPs in the Netherlands stood up for better salaries. In addition, we had to compete with several other health care changes during the recruitment period, in which GPs were investing a lot of time and effort, such as the introduction of practice nurses in general practice and the development of large-scale out-of-hours cooperatives. Many GPs were involved in these developments and stated that they lacked the time to participate in other activities.

The trial aimed for an inclusion of 5000 patients. This number was needed for the diagnostic study that was part of the trial. Due to the low prevalence of serious pathology, many patients needed to be included. The total number of patients who eventually participated in the study was 513. This number restricted us in drawing meaningful conclusions regarding the accuracy measures for the blood tests, but the earlier mentioned incorporation bias had already reduced our possibilities to do so.

GPs from each of the three trial arms included fewer patients than expected on the basis of prevalence research. They experienced several difficulties and ‘Lasagna’s law’ fully applied to this trial. Several risk factors of less successful patient recruitment were present in our trial, such as the need for incident cases and GPs who had to remain alert during consultations and who were the first to inform the patients about the study.

Most reasons quoted by GPs for not including patients were ‘forgot to think about the research’ or ‘too busy’. Some other reasons that were mentioned may explicitly have caused bias. For instance, when a patient was considered eligible for inclusion, GPs sometimes reported that a few extra questions yielded an explanation for the complaints after all. Other GPs reported that they were uncertain whether complaints were ‘unexplained enough’ to warrant eligibility for patient inclusion. It is not easy to determine the direction of the bias due to these factors, nor is it possible to determine the effect of the limited patient inclusion on the generalisability of our conclusions. Nevertheless, patient inclusion took place in daily general practice and GPs decided on inclusion themselves. Our population is thus a reflection of daily general practice and therefore our conclusions are applicable to that setting.

The groups of GPs that were instructed to postpone blood test ordering, groups 2 and 3, possibly had difficulties including patients, as it is known that some GPs feel empty-handed if they have no test or therapy to offer. This may be even more true in the case of our trial, because there was no new test to offer patients as a reward for participating in the trial.
The patients who were included in the trial were not evenly distributed among the three groups. The test postponement group without quality improvement strategy (group 2) included significantly fewer patients than those in the other two groups. The deviation was caused by the fact that several GPs in this arm did not include any patients at all. In terms of background characteristics of the patients, there were no statistically significant differences between the groups. However, although patient inclusion was lower in group 2 than in group 1 (immediate test ordering) this was not the case in the ‘postponement with quality improvement strategy group’ (group 3). Although this might be explained by the fact that the GPs in this group were supported by the quality improvement strategy, the analyses of the strategy actually showed a lack of effect\(^\text{20}\). Participation in the third experimental arm may have caused higher awareness of the trial and therefore improved inclusion, as several GPs mentioned that it was difficult to stay aware of the trial while seeing patients.

We further analysed potential selective inclusion by trying to perform a non-inclusion analysis. This was difficult, because ‘unexplained complaints’ had not been registered as such by the GPs in the patients’ records. We searched the electronic medical files of the participating GPs by means of text words for eligible but not included patients with unexplained complaints. The text words were chosen after exploration of often-used text words in the medical files of included patients. This study did not show significant sex and age differences between included (n=513) and not-included (n=507) patients. However, differences may exist in other characteristics, such as education level.

One way to overcome the skewed distribution might have been randomization at the patient level instead of at the practice level. In the context of the VAMPIRE study, however, this was not an option, because the quality improvement strategy was carried out mainly at the practice level, requiring analysis at the same level. In addition to this methodological argument, there was also a more fundamental argument: it would be impossible for GPs to apply what they had learned in the quality improvement strategy to some patients and to ignore this knowledge with the next. In conclusion, although the distribution of patients over the groups was skewed, we have no indication that the inclusion was selective. Of course, we cannot entirely rule out selection either.

Implications

Implications for daily general practice

The studies presented in this thesis and the other results of the VAMPIRE study presented in Van Bokhoven’s thesis, show that proposing postponement of blood test ordering when patients newly present with unexplained complaints seems safe, is feasible and does not have unfavourable effects on patient
satisfaction and anxiety. Merely postponing blood test ordering, however, is not enough. Postponing blood test ordering should be part of a watchful waiting approach, in which GPs explicitly instruct all patients to revisit after four weeks. This instruction should be explicit, since patients with unexplained complaints are not naturally inclined to do so, among other reasons probably because they often do not feel helped or taken seriously. Therefore, in the first consultation attention should be paid to clarifying the patients’ needs and requests related to test ordering. Patients’ expectations with regard to further diagnostics, treatment and referrals should be explored. These expectations can be a sign of a worrying patient, which needs further exploration. Furthermore, patients’ requests could be a starting point for further explanation of the diagnostic possibilities and limitations of blood tests. Also, in one of the first consultations, GPs should explore how their patients perceive the severity of their complaints, since preventing a mismatch between GP and patient in this regard may favourably influence satisfaction and anxiety. Lastly, during the first or second consultation, attention should be paid to the perceived quality of life at the beginning of the unexplained complaint episode, since this may be quite poor. GPs could explicitly ask their patients for experienced hindrance and impairments in daily life as a result of the complaints.

By explicitly instructing patients to revisit, GPs create possibilities for monitoring the course of the unexplained complaints. In follow-up consultations, the postponement of blood test ordering may be reconsidered, and attention can be paid to the course of complaints and to the quality of life, since this tends to remain poor in patients with newly presented unexplained complaints. Exploring these issues may enable GPs to more efficiently tailor the management of unexplained complaints to individual patients.

**Implications for future research**

Several research topics derive from this thesis and from the VAMPIRE study in general.

First of all, the concept of unexplained complaints should be clarified further. An increased uniformity of the definitions and the development of classification systems for unexplained complaints for research purposes would help to demarcate the research area, and to increase the generalisability and applicability of findings in daily practice.

Furthermore, the incorporation bias problems have to be solved in order to be able to establish the diagnostic value of blood test ordering among patients with newly presented unexplained complaints.

We have presented our findings on the relation between blood test results and diagnoses in patients with unexplained fatigue, but these data may also be analysed for patients with unexplained abdominal complaints and musculoskeletal complaints. This would give more insight into the diagnostic
accuracy of blood test ordering in patients with newly presented unexplained complaints in general, not merely with regard to unexplained fatigue. Another theme is clarifying the consequences of false-positive test results. Although the cascade of superfluous further testing, referrals and treatments that supposedly ensues from these false positive results, is frequently used as an argument for the reduction of test ordering, there is currently no hard evidence of the existence of such a cascade in general practice. We have not succeeded in finding determinants for the evolution of newly presented unexplained complaints into established unexplained complaints. Research into and specification of these, preferably modifiable, determinants may subsequently allow the research and development of early interventions in order to prevent chronicity. Finally, the relationship between a poor quality of life and unexplained complaints deserves to be included in the research of unexplained complaints.
References


