Results of a survey among general practitioners in Germany

What can diagnostic algorithms do for primary care?

Julian Wangler, Michael Jansky

Centre for General Medicine and Geriatrics, Mainz University Medical Centre, Germany

Introduction

The reality of general practice is marked by a wide range of symptoms, complaint pictures and disorders (low-prevalence area) [1]. There are only limited technical and temporal resources for dealing with this potentially broad spectrum of possible diagnoses. It is thus all the more important to make use of a broad range of decision-making strategies to enable a quick orientation and effective diagnosis in everyday practice [2, 3].

Diagnostic algorithms can help in decision-making, for example, in the first diagnosis, in complex or rare clinical pictures and for consistent monitoring. As a way of making the diagnosis path effective, algorithms enable the physician to proceed quickly, simply and clearly [4]. Starting with a specific symptomology, the algorithm proposes pre-structured clarification paths, which customarily recommend a careful anamnesis and differential diagnostics in the form of flow charts. Algorithms offer specific dichotomous steps; they can support medical decisions as well as clarification (laboratory, imaging, waiting and keeping the matter open or transfer, interventions, etc.) [4, 5]. To this extent, dichotomy has diagnostic and therapeutic effects [6]. In the sense of staggered diagnostics, algorithms can help to reduce uncertainty in the GP setting [7–10].

For example, a diagnosis path for clarifying increases of liver test results can include certain blood test results, but also structure the cooperation of primary care and specialists by the GP, who is given specific proposals for the conditions under which waiting with monitoring or a transfer to a specialist or to out-patients are indicated [11]. A significant added value of a diagnostic path can therefore consist in identifying at-risk patients at an early stage and at the same time avoiding excessive use of specialists and special outpatient clinics.

The sources of diagnostic algorithms are extremely varied. On the one hand, they come from evidence-based medicine, for example as integral parts of clinical guidelines. On the other hand, algorithms are published by external actors, including associations close to medicine, foundations or pharmaceutical companies. The objectives and focal points of algorithms vary depending on the area of symptoms and provider [12].

Whereas the use and research of diagnostic algorithms has been established in the hospital inpatients context [13–17], there is – in particular for German-speaking countries – practically no information about the significance that such clarification and action schemes have in the out-patient area. Accordingly, there is hardly any information about the extent to which GPs use algorithms in everyday treatment, the conditions under which they have recourse to them, their expectations of such aids and the experience they have gained up to now.

Diagnostic algorithms can help in decision-making, for example, in the first diagnosis, in complex or rare clinical pictures and for consistent monitoring.

Diagnostic algorithms can be viewed as evidence-based tools that ultimately aim to structure, standardise and streamline care processes. In this respect, they share an important feature with evidence-based guidelines. Studies held in this field show that GPs are often rather hesitant towards guidelines [18–20]. The reason for this is that guidelines are not rarely experienced as a contradiction to individual diagnosis and therapy, and are considered as "patronising and far from reality" due to a perceived "standardising character" [21]. Whereas external evidence from clinical research is seen as an element that is hard to integrate into the diagnostic process and even as a threat to therapeutic self-determination, on the other hand the value of the "lived anamnesis" is emphasised [21–24]. Things are made more difficult by problems that GPs often complain of: too little consideration of GPs in guidelines and recommendations for action; excessive complexi-

The appendices can be found online on www.primary-hospital-care.ch

ty and thus a lack of applicability to practical use. Against the background of these results, one might assume that algorithms – ideally coming from evidence-based medicine – also partly attract criticism or at least reservations from the GP's perspective [25–27].

Research interest

In order to obtain current information on this practice-relevant topic, an explorative study was performed. The research interest can be summarised in the following main questions:

- What attitude do GPs have with regard to diagnostic algorithms? In which fields of application do they consider the use of such decision-making aids to be sensible?
- What must a diagnosis path be able to provide from a GP's point of view so that it can be used for treatment of patients?
- How often are diagnostic algorithms used by GPs in everyday surgery work? What experience has been gained?
- What would have to be improved with diagnostic algorithms so that they can be used more strongly in GPs' treatment in the future?

The present study dealt with all types of algorithms from a general perspective (e.g. paper-based, interactive computer programs, machine learning tools).

Material and methods

Study design and setting

The present study aimed to obtain the broadest possible view of GPs' opinions and experience regarding diagnostic algorithms. A quantitative online survey of GPs was used with an invitation letter sent by post. The study was carried out between April and August 2021 in Baden-Württemberg, Hesse and Rhineland-Palatinate in Germany. It was conducted and processed by the authors, two primary care researchers employed at a Department of General Practice.

Questionnaire and sociodemographic variables

The questionnaire, which was deliberately kept compact (see appendix), was designed on the basis of literature research. Studies were used that dealt with attitudes and experiences of GPs with regard to evidence-based instruments and structured care, amongst others adherence to guidelines and disease management programmes [1, 6, 21, 28, 29]. The studies of Bötler et al. [21] and Vollmar et al. [28] served as a starting point for the derivation of attitude-based item batteries (questions 3, 4, 9). The information from relevant previous studies conducted by the authors, in which various aspects and interconnections of structured treatment in the GP setting were observed, was also included [30, 31].

Preliminary discussions were held with five GPs, who made a decisive contribution to the process of specification and focusing the questionnaire and specifically served the generation of the three item batteries already mentioned.

The questionnaire (needing about 8–10 minutes for filling in) is composed of three focal points of contents: (1) attitudes and expectations with a view to diagnostic algorithms (questions 1 to 4); (2) own use of and experience with diagnostic algorithms (questions 5 to 8); (3) optimisation potentials (questions 9 and 10). Questions 2, 3, 4, 7 and 9 were nominally scaled, questions 1, 5, 6 and 10 verbally scaled. Question 8 was open.

Age, gender, surgery environment, form of surgery and patients per quarter were recorded as sociodemographic features.

Before field use, a pre-test with 25 GPs was held.

Recruiting and sociodemography

All 13,170 GPs active in treatment in Baden-Württemberg (6664), Hesse (3839) and Rhineland-Palatinate (2667) were invited by post to participate in the anonymised survey in April 2021. This was a single letter, in which the physicians to be questioned were, amongst other things, given password-protected access to the online survey. The participants in the survey did not receive any reimbursement of expenditure or incentives.

Ethics

During this study, no sensitive patient data were gathered or clinical tests performed. This was a strictly anonymised survey of a total of 3110 GPs. The Ethics Commission of the State of Rhineland-Palatinate, Germany, informed us that approval by an ethics committee was not necessary.

Written informed consent for participation was obtained from all participants before the start of the study.

Data analysis

We analysed the data using SPSS 23.0 for Windows. Apart from the descriptive analysis, we used a t-test for independent samples to analyse for any significant differences between the two groups, assuming significance at values of $p \leq 0.001$. Only when the P-value is very small (under 0.001) does an inequality become justifiable.

For better visibility of various attitude and expectation clusters (questions 3, 4) of GPs, we made use of factor analysis (varimax rotation), in which variables are put together as factors on the basis of systematic relationships (correlations) to one another [32]. In the run-up, the preconditions for the factor analysis were tested (sampling adequacy according to Kaiser-Meyer-Olkin, significant result in the Bartlett test for sphericity, communality of all the included variables above threshold figure 0.5). As a limit from which an item loads to a factor, the figure 0.4/–0.4 was chosen [30].

Results

Sampling

Of the 3189 questionnaires which were processed, 3110 that were completely filled in were evaluated (return: 24%). The sample has been structured as follows:

- Gender: 52% male, 48% female
- Average age: 53 (median: 54)
- Surgery environment: 44% medium and large cities, 56% rural/small towns
- Form of surgery: 52% individual surgeries, 44% joint surgeries, 4% miscellaneous
- Patients per quarter: 22% <1000, 29% 1000–1500, 28% 1501–2000, 21% >2000

Assessment, potentials of use and expectations

Sixty-eight percent of those surveyed considered diagnostic algorithms to be a useful aid in everyday surgery work as a matter of principle. One quarter (25%) did not see any benefit (7% undecided or no answer).

Positions differed widely onthe situations in which the use of diagnostic algorithms was regarded as effective (multiple replies). For example, 59% were in favour of using algorithms in cases of suspicion; 55% considered them to be usful in specific diagnoses such as rare illnesses. In addition, 41% considered their use sensible in a first diagnosis, 40% in emergency situations, 30% for screening purposes and 18% for checking sequences.

As the answers to the one-item battery show (table S1, published with the online version of this article), the participants perceived the principal benefits of diagnostic algorithms as, in particular, a purposeful diagnostic mode of procedure and improvement of cooperation between the treatment levels. Around half of them had insecurity about the selection and quality of individual algorithms. In addition, applications not consistent with primary care were criticised, as a result of which use of algorithms was not always possible in the time-critical everyday work in surgeries. One in four prefered to rely on his/her own personal methods and not on a diagnostic algorithm.

As factor analysis showed, the participants' attitudes to diagnostic algorithms can be assigned to four clusters of differing size. Perceived application and practicability benefits were in the first two clusters (e.g., the diagnostic and structured procedure); sceptical positions were in the other two groups (e.g., discrepancy between algorithm and intuitive procedure in the clarification steps).

There were some differences between the physicians who principally assessed algorithms positively and those who did not see any benefit worth mentioning. Physicians with a positive attitude stated more frequently that algorithms support onsistent procedures (71% to 26%, p <0.001) and clarification of symptoms (70% to 14%, p <0.001). Physicians with a negative attitude criticised low relevance to application of diagnostic algorithms much more frequently(62% to 35%, p<0.001) and emphasised that they would prefer to rely on their own methods as a GP (55% to 8%, p <0.001).

Noteworthy differences were not seen in the recorded sociodemographic features. However, physicians with a rural surgery stated more frequently than urban physicians that they prefer to rely on their own methods instead of a diagnostic algorithm (33% to 16%, p <0.001).

Expectations and requirements of the GPs for diagnostic algorithms were determined via another item battery (table S2, published with the online version of this article). For the participants the most important thing about diagnostic algorithms is that they are simple and effective to use and enable systematic demarcation of possible clinical pictures on the way to a (suspicion) diagnosis via the identification of clear warning signs. A majority also were in favour of algorithms whose benefits have been proven empirically. A cost-covering alignment of algorithms was also of significance for the GPs.

The factor analysis showed a four-factor model with a comparatively high clarification of variance. The respondents in cluster one attached great importance to instructions on waiting and watching, specific reference values for laboratory diagnostic tests (a clarification procedure that is as precise and application-oriented as possible) and medication effects. Respondents in cluster two prefered a clear evidence base and, thus, a quality check of algorithms. In cluster three, compatibility with primary care played a major role. Respondents in cluster four emphasised, among other things, compliance with the fee schedule.

Compared with urban doctors, rural doctors placed more importance on the fact that a diagnostic algo-

rithm must be simple to apply (91% to 72%, p <0.001). Likewise, rural doctors had a greater requirement for clear compatibility with general practice for such instruments (62% to 47%, p <0.001).

Use of and experience with diagnostic algorithms

Forty-six percent of the participants stated that they used diagnostic algorithms frequently or occasionally in the surgery (35% rarely, 19% never). The participants who used algorithms frequently, occasionally or rarely stated that they had mainly good or very good experience with using them in the surgery up to now (71%); 15% stated that their experience was rather bad (14% no statement).

Specific potential was seen above all in ensuring a standardised, consistent diagnostic procedure and better transparency of diagnosis paths.

An open question concerned areas of symptoms for which diagnostic algorithms have already been used (maximum three areas were to be named). The most frequent ones were chest pains (54%), cardiac disorders (above all coronary heart disease, infarct, 52%), vertigo (48%), thromboses (45%) and abdominal pain (36%).

The participants stated that they originally became aware of the diagnostic algorithms used via postgraduate training (59%), guidelines (57%) and specialist periodicals (49%, multiple replies). Other sources such as works of reference, studies/postgraduate specialist training, recommendations by colleagues or brochures from associations or pharmaceutical companies were a long way behind.

Improvement potentials

Finally, various proposals for optimisation of diagnostic algorithms that are to be used in GPs' practices were presented. The participants were most clearly in favour of ensuring that diagnostic algorithms intended for GPs' use are (also) developed by a neutral and independent agency as far as possible (69%). Fifty-nine percent emphasised that there should be diagnosis paths developed specifically for the situation and perspective of GPs. Fifty-four percent placed importance on the fact that the benefit and the evidence are proven. For a further 45%, the definition of binding excellence and quality criteria for diagnostic algorithms was important. Forty-two percent considered a stronger orientation of algorithms to the fees ordinance to be necessary.

Under the precondition that the improvements mentioned are made, 58% were willing to take diagnostic algorithms into account in treatment of patients more than up to now; 26% would even consider a distinctly wider use (13% no, 3% no information).

Discussion

Main findings and comparison with prior work

A certain degree of diagnostic insecurity is a common phenomenon in decision-making in primary care, particularly under the conditions of shortage of time and resources. This is why strategies for effectively limiting the range of possible diagnoses and recognising threatening sequences of illnesses at an early stage without stressing the patients with unnecessary examinations are necessary. From the point of view of the 3110 German GPs surveyed, diagnostic algorithms can be a useful instrument for achieving this objective. Specific potential was seen above all in ensuring a standardised, consistent diagnostic procedure and better transparency of diagnosis paths.

However, it is noticable that a certain number of the physicians surveyed had reservations and skepticism about the algorithms currently on offer for GPs. With theview that algorithms are hard to assess, a high degree of uncertainty about which algorithm is resilient and trustworthy was conspicuous. From the point of view of the respondents, applications matching practice and adapted to GPs' treatment, as well as recommended for actions taking costs into account, often do not exist. If, in particular, better compatibility with GP practices were guaranteed in diagnostic algorithms, a large number of those questioned would be willing to use them to a greater extent in their own surgery.

Thus, the results correlate to a high degree with national and international research, which, all told, established a critical underlying attitude of independent physicians towards evidence-based instruments [18–27, 33–39]. As authors such as Donner-Banzhoff [1], Heneghan et al. [2], Schneider et al. [7], Blissett et al. [12] or Vollmar et al. [28] emphasise, especially with regard to diagnostic tools, GPs see the potential of the application of guidelines to quickly implement the necessary steps for the appropriate (further) treatment of patients [7–10]. Despite such advantages, they fear that guidelines may restrict their flexibility and freedom of choice in a crucial way [7]. Guidelines, which can also contain diagnostic algorithms, are often regarded with skepticism because they seen as being over-complex and difficult to apply to a GP's situation [19–21]. In addition, GPs often sense restriction of their own freedom of action as a result of such restrictive requirements [18, 22–24, 30]. At the same time, more modern studies show that the acceptance, knowledge and use of guide-

lines and evidence-based structures and as tools, has noticeably grown in recent years [30, 31]. Many GPs nowadays base their work more on standardised, evidence-orientated interventions [36]. In this regard too, this study provided clear indications.

Implications

In the development of new diagnostic algorithms, attention should be paid to the greatest possible consistency with primary care. This includes ensuring simple applicability and clarity, and reduction of complexity of guidelines [32] and can be achieved by increased involvement of GPs themselves. At the same time, there should be assurance that the proposed diagnosis steps are cost-covering. As thre is a perceivable tendency that algorithms are being provided to GPs more frequently and attempts are being made to homogenise action by GPs by various actors in the health service (specialist associations, health insurance schemes, pharmaceutical companies, foundations close to medicine, etc.), attention should be paid to the fact that algorithms do not lead to a restriction of GP's freedom of choice of therapy. Instead, they should maintain individual leeway for action for the GPs. This is the only way in which adequate acceptance by GPs can be found and which would provide added value for GPs' treatment. Because of the unclear offer of diagnostic algorithms, specialist associations could devise algorithms compatible with general practice systematically and examinethem for quality and user-friendliness. This would be a decisive help for GPs in receiving better orientation and developing confidence in their use.

Strengths and limitations

The authors understand the survey as a pattern of opinions giving exploratary access to a subject yet to be extensively examined. The survey had been supported by a previous discussion and preliminary studies from the thematic surroundings and achieved a satisfactory response rate. Nevertheless, it has a series of limitations, which have to be consideredcritically.

For example, the study cannot claim to be representative. This is connected on the one hand with the limited number of cases and the regional recruitment, and on the other hand with the fact that it was an online survey. It can be assumed that this form of data collection could not be used equally well by all GPs, because, for example, of a lack of online affinity or an inadequate internet connection for rural surgeries. In addition, it cannot be ruled out that physicians with an interest in the subject or positive previous experience with guidelines took part to a greater extent, with the result that a selection bias might exist.

We would also point out that it is a question of collecting (self-)appraisals about the application of diagnostic algorithms, which do not automatically correspond with actual conduct. Because of the complexity of the subject, the survey can naturally only cover a small section of opinions. Accordingly, (further) examinations and evaluations of the use of specific diagnostic algorithms by GPs are necessary.

It is important to consider and classify the results of the present study in the context of the German healthcare system. There is a special need for checks, coordination and effectiveness here, since there is no primary doctor system in Germany and patients can therefore go directly to specialists, for example. At the same time, the GP has a strong and independent position in the German healthcare system.

Conclusions

The survey provides indications that GPs principally perceive a considerable benefit with diagnostic algorithms. However, these potentials have not always been used in everyday treatment of patients. As this is a field greatly lacking in order, in which algorithms are provided to GPs from a number of sources, many physicians naturally have difficulty in assessing their quality and reliability. A problem at least as great is that algorithms are often developed in clinical contexts that are not always compatible with the reality of GPs' treatments. In the development of new diagnostic algorithms, which for example are to cooperation between the various levels of treatment, attention should be paid to production of the greatest possible conformity with GPs.

Julian Wangler, PhD
Research Associate
Centre for General Medicine and Geriatrics
Universitätsmedizin Mainz
Am Pulverturm 13
DE-55131 Mainz
julian.wangler[at]unimedizin-mainz.de

Disclosure statement
No financial support and no other potential conflict of interest relevant to this article was reported.

References
Die vollständige Literaturliste finden Sie in der Online-Version des Artikels unter www.primary-hospital-care.ch