Universiteit
Leiden
The Netherlands

In between looking and seeing: recognition, referral and assessment of children and adolescents' mental health problems at the interface of primary care and secondary mental healthcare

Aydin, S.

Citation

Aydin, S. (2023, February 23). *In between looking and seeing: recognition, referral and assessment of children and adolescents' mental health problems at the interface of primary care and secondary mental healthcare*. Retrieved from https://hdl.handle.net/1887/3564429

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3564429

Note: To cite this publication please use the final published version (if applicable).

In between

looking and seeing

Recognition, referral and assessment of children and adolescents' mental health problems at the interface of primary care and secondary mental healthcare

Semiha Aydın

In between looking and seeing

Recognition, referral and assessment of children and adolescents' mental health problems at the interface of primary care and secondary mental healthcare

Semiha Aydın

This thesis was printed on 100% recycled paper.

ISBN: 978-94-6469-177-1. Copyright 2022, Semiha Aydın, the Netherlands. No part of this thesis may be reproduced in any form or by any means electronically, mechanically, by print, or otherwise without permission of the author.

Cover by Ayşe Kübra Küpeli Akyazı

Chapter design and lay-out by Semiha Aydın and Proefschriftmaken.nl

Printed by Proefschriftmaken.nl

In between looking and seeing

Recognition, referral and assessment of children and adolescents' mental health problems at the interface of primary care and secondary mental healthcare

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op donderdag 23 februari 2023

klokke 16:15 uur

door

Semiha Aydın

geboren te Rotterdam

Promotores

Prof.dr. P.M. Westenberg

Prof.dr. M.E. Numans

Co-promotor

Dr. M.R. Crone

Leden promotiecommissie

Prof.dr. P.F. Wouters (Decaan/voorzitter)

Prof.dr. A.W.M. Evers (Secretaris)

Prof.dr. V. Hendriks

Prof.dr. E.M.W.J. Utens (Universiteit van Amsterdam)

Prof.dr. M.H. Nauta (Rijksuniversiteit Groningen)

The research described in this thesis was supported with expertise, data en guidance by LUMC-Curium, centre for child and adolescent psychiatry – with a research focus on detection and treatment of complex psychiatric problems in youth, by the former research profile area Health, Prevention and the Human Life cycle of Leiden University, and the Leiden University Medical Centre societal outreach programme Population Health.

“Kindness is more important than wisdom,

and the recognition of this is the beginning of wisdom.”

Theodore Isaac Rubin

TABLE OF CONTENTS

“As for the future, your task is not to foresee it, but to enable it.”

Antoine de Saint Exupery

Chapter 1

General introduction

Box 1

Imagine a consultation with the general practitioner, Dr. Jan. A 13 years old girl, Hanne, visits with her mother for recurrent unexplained abdominal pain. Her mother mentions that Hanne also does not manage to complete full days at school. The physician focuses thoroughly on the abdominal symptoms and continues the physical examination.

When asked for, Hanne's mother explicates that the school attendance problems happened a few times in the last two months. She explains that she divorced Hanne's father over a year ago and wonders whether Hanne might be troubled by this big change at home. "However", she adds, "the problems at school are of a more recent date and seem not so much related". Hanne's adjustment to the divorce of her parents took some weeks after she moved to her mothers' house, but Hanne is now best friends with two other girls at school. This reassures the general practitioner.

The general practitioner decides not to ask further as there is no defined request for help with psychosocial problems and Hanne seems not to be off-balance. Hanne gets a prescription for a fiber solution to treat the presumed bowel problems, the mother expresses thanks for the consultation, and they leave.

Ideally, when minors' emotions persistently hinder wellbeing, an appropriate treatment or management plan should be made and followed through. In practice, however, many steps and challenges exist between the two ends of mental health needs and obtaining adequate support. A major and first challenge is recognizing minors' experiences or expressions that might point towards the existence of a mental health-need and evaluating their symptoms and strengths.[1 2] Not only do patients and their primary surrounding fail to recognize psychosocial needs, but also professionals fail to recognize and assess them sufficiently.[3-5] For example, in the above-described consultation, the general practitioner decided not to enquire further as the mental health related problems were not the primary focus of that consultation and the few explicated problems did not sound much worrisome to the professional (Box 1). The general practitioner underrecognized the likelihood that this might be an early case of school refusal, which is a major risk factor for problems to build up and persist.[6 7] If the general practitioner had perceived risks, they could enquire further and balance attention to both the likelihood that watchful waiting might be sufficient as well as

whether signposting to resources for prevention and health promotion might be adequate. Therewith the general practitioner could reach its potential in navigating minors towards 'the right service in the right place, at the right time delivered by the right person'– a meaningful Chinese proverb often used by authors from the field of healthcare.

In practice, it is not uncommon that a professional has no resources to recognize, assess, and react to the first signs of mental health problems.[8] This is also true for anxiety disorders which are highlighted in decades of research as the most common mental health problem amongst minors and adults.[9] We now more than ever have insight into the correlates and consequences of mental health disorders that are not outwardly visible but have major long-term consequences. Nonetheless, the treatment gap remains and efforts investigating practical steps between recognition and treatment are relatively scarce in the literature.

In this dissertation we present a series of studies on I) whether general practitioners even think about a probable anxiety disorder when presented with descriptions of psychosocial problems (chapter 2), II) what general practitioners write on their referral letters to child and adolescent mental healthcare (chapter 3), and III) what the predictive value of a feasible step-wise assessment method could be at the interface between primary healthcare and secondary mental healthcare (chapter 4). While values for common mental health disorders (e.g. depressive disorders and attention-deficit hyperactivity disorders) will be presented in each chapter, extra focus will be put on anxiety disorders as an example of a widespread yet overlooked and underserved mental health problem.

The results of the studies could inform clinicians on the status quo considering recognition and referral of minors with mental health problems, which are major healthcare topics in several countries including the Netherlands. Moreover, the findings could serve policy and curriculum makers, thereby improving evidence-based practices in child and adolescent mental healthcare.

Epidemiology of mental health in youth

Most of young people develop healthily and show resilience to stressful experiences. Notwithstanding, mental health problems are not uncommon. Each year about one in four minors meet the criteria of mental health disorders as they are hindered in their daily life by

emotional and behavioural problems.[10 11] The prevalence rates are not much different, if not higher amongst adults.[12] In fact, both retrospective and longitudinal studies suggest that a substantial amount of adult cases are formed by caseness from younger ages on.[9 13-15] It is found that about three-quarters of persons with a mental health disorder at some point in their life met the criteria of a psychiatric disorder already before the age of 24, and over half the lifetime cases had mental health problems before the age of 14.[9 13] Indeed, the age of onset distribution of most mental health problems shows an onset in childhood or adolescence (Figure 1[16]). In addition to that, it is not far-fetched to conclude that child mental health has to be highlighted when aiming for healthy functioning of individuals and society.

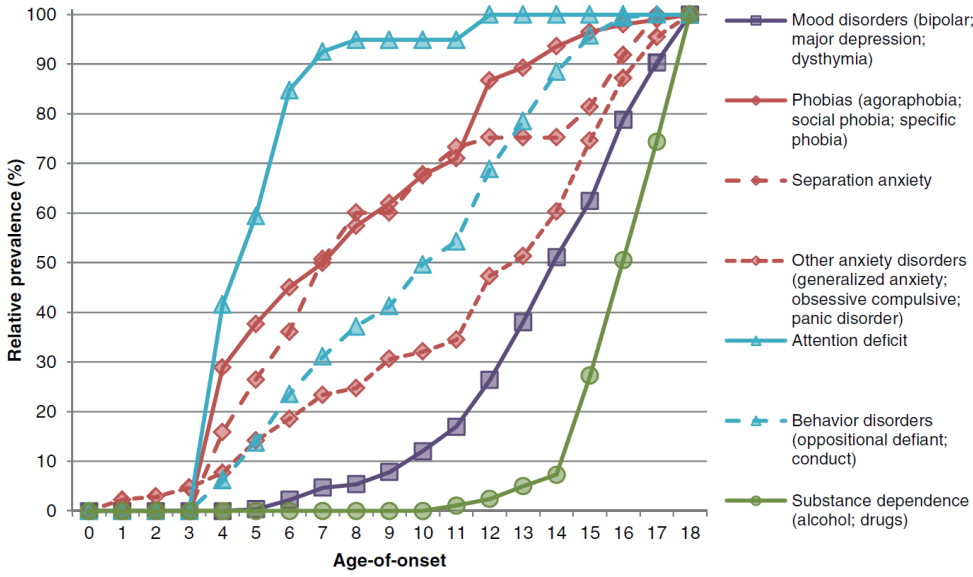


Figure 1. Standardized cumulative prevalence curves for mental health disorders, from Ormel et al. (2015). The plots track per age the relative prevalence of each index disorder.[16]

Mental health disorders account for a significant amount of disease burden worldwide; in literature this is also expressed as disability adjusted life years (DALYs). When grouped altogether, they form the fifth leading cause of DALYs over the life span.[17] Specific categories of mental health problems, such as depressive disorders and self-harm, account for three out of the ten leading causes of DALYs. In female adolescents in the age group 10 to 19 years

anxiety disorders on their selves are even amongst the top five leading causes of disease burden. The cost and cascade of problems triggered by mental health problems often build up over time as a result of their early age of onset. Therefore, as the World Health Organization formulates, “investments in minors might bring a triple dividend of benefits: now, for their future adult lives, and the next generation”.

Effective investments might help prevent a negative snowball effect in minors’ developmental trajectory. Access to mental health treatment is one such example of a valuable investment to prevent their developmental potential from being thwarted.[18 19] Nonetheless, there is a major gap between the occurrence and treatment of child mental health problems, also called a treatment gap. On average over two thirds of young people with a mental health-need do not access adequate support.[20 21] The numbers vary depending on the type of mental health problem with for instance an estimated 3.4% of people with social anxiety disorder making treatment contact within a year of onset, 33.6% of people with a panic disorder, and 37.4% for major depressive disorder.[21] Median delays in access to treatment are estimated as respectively 16 years, 10 years and 8 years after onset.[21]

Challenges in recognition and access to services

To facilitate timely access to services, minors’ mental health needs must be recognized. However, in reality, this is a rather challenging aim. Many steps exist between the two ends of a need for care and obtaining appropriate support and intervention. In their filter model Goldberg and Huxley[22] describe the pathway to mental healthcare by means of five levels of care and four major filters in between these levels (Figure 2[22 23]). They underline that primary care is the first and foremost point where decisions are made and at which most of the sub-filters are.[24] Examples of such filters are the illness behaviour of the patient, and other characteristics of both the doctor and patient.[22 24] A major challenge in mental health needs of children and adolescents is recognizing their experiences and expressions that might indicate a mental health problem and how to evaluate their symptoms and strengths. As children and adolescents are still developing, it is difficult to differentiate signs of problematic development from typical and temporary developmental challenges. Also, as they depend on their primary surrounding, another challenge is formed by problems that are

difficult to perceive outwardly by parents.[25-27] Hence others have expanded the filter-model and included filters such as the social context,[28] problem recognition by parents or important others in minors' lives,[29] or more specific stages of problem recognition, such as perceiving that there might be an issue, what the issue is, and subsequently the realization that there is a need for mental healthcare.

Mental health treatment seeking is not only a multistep process[27] as a result of the persons involved but also due to complexities in the health system. Even though the filter-model of Goldberg and Huxley implicates that after referral patients reach psychiatry, in the current era with various stakeholders and institutions, there are many more pathways – with each also various bottlenecks and hence relationships with the earlier filters. In the next paragraph, we describe the current landscape in child adolescent mental healthcare in the Netherlands, and tap into how the variety of pathways might help benefit from the pros of specialization, as well as increase risks in decision-making through increased ambiguity in responsibilities and opportunities.

The care landscape

The variety of pathways toward mental healthcare is a result of decades of reforms set up to meet increasing healthcare demands both in terms of quality and quantity. In developed countries, in the last decades, the reforms were dominated by the replacement of specialised services with low threshold services, also known as the dichotomization of primary and secondary care services. Primary care services were created in the 1950s, after recognition of their potential to improve monitoring and management of health-related problems in a low-threshold, local and low-cost framework. Since then, the healthcare system has undergone many reforms, and decades later, primary care is still at the center of reforms. Recently, in the Netherlands new reforms took place in child healthcare

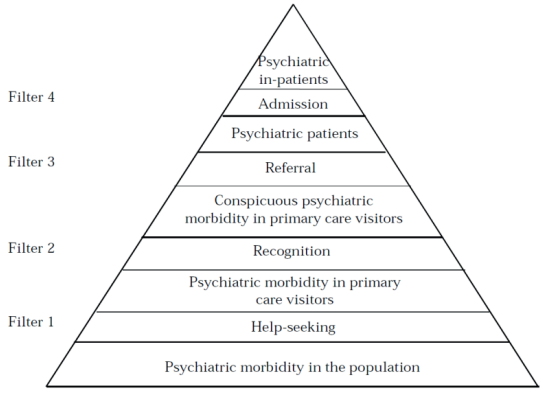


Figure 2. Goldberg and Houxley's Filter-model (1982) describing the process of help seeking for mental health problems, from Zwaanswijk (2005).[22 23]

legislation by the Youth Healthcare Act.[30 31] Before 2015 the care system for youth was financed and organized in different levels and the system with various stakeholders did not facilitate the provision of tailored and harmonious care. Starting from 2015, all authority, financing and responsibility were given to the municipalities to diminish the number of stakeholders and fragmentation in care. So-called inter-professional local youth teams were formed to intensify preventive care and function as a bridge between primary and specialised care services. With this decentralisation it was expected that better-tailored interventions could reach minors earlier and prevent their problems from worsening.[32] Nonetheless, seven years later, at the time of writing this introduction, the delivery of coherent, coordinated, timely and sustainable care for youth with various needs in the social and health domains is still a major topicality. Evaluations on how the youth act turned out report that working methods of local youth teams differ between municipalities and that integrated working approaches might lack even within municipalities.[33] Reports conclude that the matter might be to facilitate what works on the case level, independent of how the health system is organized.[34]

Assessment

In a care-model with various stages and specialisations, by definition, a chain of providers will be involved, and thus communication will form a challenge. Minors' strengths and challenges should be made explicit and assessed for adequate and cohesive care. This is also

highlighted in studies on integrated care. Such studies underline that next to clear clinical pathways, and collaboration between professionals, a broad assessment of problems and possibilities is needed to facilitate support that is tailored to the needs of children and families across life domains.[35-37] A broad assessment approach that is done timely might also contribute to the achievement of several aims that are also formulated with the new Youth Healthcare Act, including ameliorating the quality of care, reducing waiting time, and facilitating access to care.[33 34] But what is assessment? Assessment as part of evidence-based practice refers to "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (Box 2). This definition underlines the importance of both clinicians' day-to-day evaluations as well as drawing from results of empirical research. Both are viewed as an indispensable part of high-quality care that starts with putting from both clinicians' expertise and more formal guidelines in order to explicate patients' "characteristics, culture, and preferences". In absolute terms one might even argue that without assessment no focused treatment can be initiated, as at least a form of hypothesis generating and testing has to precede before and during each treatment intervention.

Box 2. Definitions of evidence based practice and assessment

What is **evidence-based practice** in **medicine**, Sacket et al (1996)	Evidence-based practice, defined as "(…) the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research. (…)".[38]
Definition and purposes of **assessment in psychiatry**, Mash & Hunsley (2005)	What is assessment: Assessment captures the assessment process (hypothesis generating and testing, decision-making) and the assessment methods (tests, observation, measures to gather data). What are the purposes of assessment: the six interrelated purposes of assessment are screening, diagnosis and case formulation, prognosis and prediction, treatment design and planning, treatment monitoring and treatment evaluation.[39]
Evidence-based assessment as defined by the **American Psychological Association**	Evidence-based assessment aims to "promote effective psychological practice and enhance public health by applying empirically supported principles of psychological assessment, case formulation, therapeutic relationship, and treatment".[40]

Assessment captures both the methods and the process to aid hypothesis generating (Box 2). In the clinical field, however, often with assessment, the use of structured assessment methods such as questionnaires is meant, partly as a result of the origin of structured assessment methods for scientific purposes. Structured assessment methods such as rating scales and questionnaires were initially used for scientific and epidemiological purposes as a reliable proxy for what cannot be directly observed in the diagnostic process – i.e. the information gathering to estimate the likelihood of various diagnostic possibilities. The persons experienced subjective inner feelings are than explicated, operationalised, through rating of items that form a proxy for these experiences including feelings, cognitions and sensations.[41 42] Assessment stimulates informants thinking about symptoms and signs that are typically less perceived, as in this way informants are prompted to evaluate daily behaviour and experiences that are less remarkable.[43 44] Later on, structured assessment instruments were included in the clinical field for individual patients to diminish the risk of tunnel vision and unaddressed problems at the side of the clinicians. In child and adolescent mental healthcare, the inclusion of multiple informants into the assessment procedure is emphasized given the still ongoing cognitive development and reliance of minors on their parents, differences in children's behaviour in various contexts including school, at home, and amongst friends, and differences in how various informants perceive children's behaviour.

Focusing on general practitioners

In the Netherlands, minors are invited for a structured mental health screening and health check twice during primary school and twice during high school. Between these preventive moments, the signaling role is up to the direct caregivers of youth and their teachers.[45] When parents and caregivers signal a need, there are several pathways to follow and professionals to visit. Reports on general practice consultations from around Europe show that over two thirds of youth visit a general practitioner at least once a year.[46 47] Recent studies on referral pathways from the Netherlands suggest that over a third of the youth support trajectories indicated between 2016 and 2019 were still indicated by general practitioners, despite the legal shift towards the newly created local youth teams within the Youth Healthcare Act.[48 49] Clearly, the general practitioner is first point of contact for many families in the case of health related problems including mental health.[34 50 51] Considering

the life time perspective of general practice, general practitioners are in a unique position to recognize and manage mental health problems. Last but not least, primary healthcare is approachable in its nature and could thereby have a major role in timely identification, and help disrupt the relation between socioeconomic status and the abovementioned alarming treatment rates. Therefore, in this dissertation, emphasis is placed on general practitioners as an example of an essential partner in facilitating minors reach for mental healthcare.

Focusing on anxiety disorders

Anxiety disorders form a typical example of a widespread yet overlooked mental health problem that starts early in life, and when not intervened, frequently persists or build-ups into adult life. Hence, the case of anxiety disorders portrays a clear image of the importance of tackling mental health problems in youth. Anxiety disorders have a median age of onset of 11 years and show similarities with typical development. Signs and symptoms of problematic anxiety often fly under the radar.[13] Those who do access treatment do mostly so after as much as 9 years after onset.[52] Median delays in receiving treatment range between 9 to 30 years depending on type of anxiety disorder whereas these median delays are estimated to be below 10 years for mood disorders and substance abuse disorders.[21] We know from the literature that these are the cases with the highest life hindrance and are only the tip of the iceberg.[20 52] It is estimated that more than half of the persons with anxiety disorders never receive professional support for their anxiety disorder.[53] Studies focusing on youth conclude similar, if not more alarming, treatment coverage rates.[54-56]

The benefits of a decrease in the treatment gap of anxiety disorders have been modeled by Chisholm and his colleagues.[57 58] In their conservative simulation study, they report an economic benefit-cost ratio of 5.3 (range 2.6 to 10.9) by a treatment gap reduction of about 16 to 25%. This suggests that a reduction of the number of untreated people by a quarter might result in a five-fold monetized short-term benefit relative to the cost of treatment. Nonetheless, studies amongst children in paediatric primary care reveal that only 31% of youth with clinical level anxiety disorders have received any form of intervention whereas, for instance, 75% of children with attention-deficit hyperactivity disorder and 40% of those with depressive disorder received a form of intervention.[54] In general, studies suggest that the so called internalizing disorders that are less expressed outwardly are recognized less or

perceived as not requiring care,[23] and within the group of internalizing disorders, anxiety disorder are recognized less than depressive disorders and receive treatment less frequently.

Besides the short-term economic burden, underdetection makes problems build up. Anxiety disorders are proposed as a risk factor for the onset of comorbidities, as persons with anxiety disorders are often the first to develop a series of problems including depression, substance use, academic difficulties, dropout and difficulties in familial functioning or at the workplace. Often it is only after the development of other issues that these persons seek or obtain help, still with chances that the treatment is not focused on the anxiety disorder.[59]

Objectives and research questions

With this dissertation we aimed to add to the knowledge gap in the referral-intake process for child and adolescent psychosocial problems, specifically anxiety disorders (Figure 3). First, we studied whether general practitioners even think about anxiety disorders when they are presented with cases at risk of anxiety disorders (chapter 2). Also we examined which treatment decisions general practitioners preferred when various mental health disorders are suspected. Next, we quantified what general practitioners write on referral letters about various mental health problems once the decision to refer is made (chapter 3). Here, we also explored which more general reasons, such as problems at school or family environment, were indicated. In the third study, we investigated the added value of a sequential assessment approach with potential for broader use at the interface between primary and secondary mental healthcare (chapter 4).

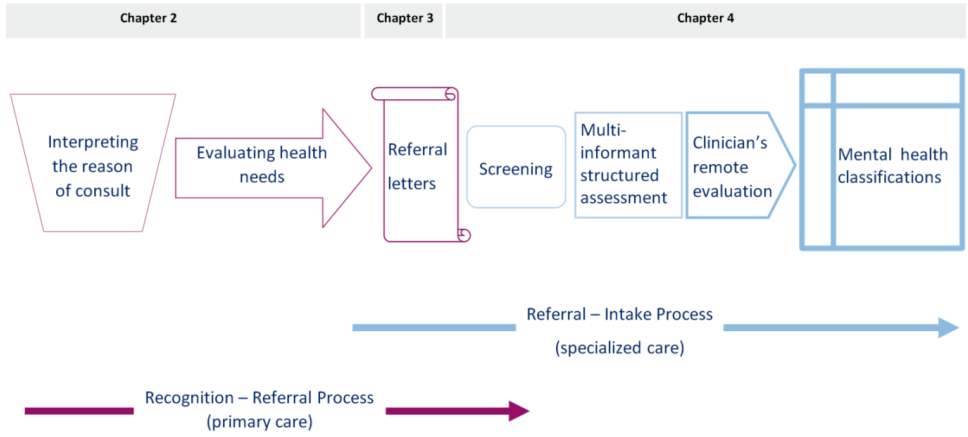


Figure 3. The scope of the studies presented in this dissertation

Study 1: Recognition of anxiety disorders in children: a cross-sectional vignette-based survey among general practitioners

Many studies have found that people with anxiety disorders in particular are infrequently recognized as such. Historically this has been related to many factors, including patient factors such as avoidance, practice- and professional based factors such as mental healthcare skills and knowledge. However, where in the diagnostic process professionals' underdetection starts is less known. We questioned whether the often shown underdetection starts already in the first diagnostic considerations and interpretations professionals make. Therefore, to gain insight into the detection and referral of child anxiety symptoms we presented general practitioners with hypothetical case descriptions, i.e., vignettes, reflecting mixed symptoms in which child anxiety may be manifested. The main question here was whether they attribute the symptoms to anxiety in their first interpretation of presented problems? We asked how they would diagnose and manage problems as described in the vignettes. To gain insight in their conscious management tendencies, general practitioners were also asked what they prefer to do in terms of referral when they assume the existence of the common mental health problems in children.

Study 2: Informative value of referral letters from general practice for child and adolescent mental healthcare

Referral letters are a central part of patients journey through healthcare institutions. Nonetheless there were little to no empirical studies on the informative value of referral letters. As a first step to address the lack of scientific knowledge in this area, we reviewed referral letters of youth that accessed specialised mental healthcare. Based on the available literature, discrepancies were expected, particularly regarding the prediction of anxiety disorders. Contrary to anxiety disorders, developmental and externalizing disorders were expected to show better symmetry between initial reason of referral and the final diagnosis. Thus, we expected that anxiety would not have been mentioned sufficiently often in referral letters of cases clinically identified with an anxiety disorders.

Study 3: The diagnostic process from primary care to child and adolescent mental healthcare services: the incremental value of information conveyed through referral letters, screening questionnaires and structured multi-informant assessment

The use of structured assessment methods is advised to improve early detection, objectify burden and sound decision-making considering the diagnosis and therapeutic progress. However, in many practices, structured assessment is not implemented sufficiently. Often clinicians report that these questionnaires and assessment instruments might limit patient satisfaction and the development of a healthy therapist-client relationship. Studies amongst patients however do not support these concerns.[8] What remains a relevant question for practices is whether the various instruments add to the clinical process. Although many instruments each with different purposes are developed, and well investigated as a standalone measure, studies on their incremental value have not been reported in the literature. As in clinical practice each following information overlaps with the previous, insight is needed in the unique predictive value of the instruments in the sequence. Therefore, in the third study, we present the added value of commonly available sources of information with potential for integrated use in and between primary and secondary mental healthcare: referral letters (RLs), the strength and difficulties questionnaire (SDQ) as a screening, and the development and well-being assessment (DAWBA) as a more elaborate assessment tool.

Reference list General introduction

1. Richardson, L. P., McCarty, C. A., Radovic, A., and Suleiman, A. B., *Research in the Integration of Behavioral Health for Adolescents and Young Adults in Primary Care Settings: A Systematic Review.* J Adolesc Health, 2017. **60**(3): p. 261-269.
2. Lempp, T., Heinzel-Gutenbrunner, M., and Bachmann, C., *Child and adolescent psychiatry: which knowledge and skills do primary care physicians need to have? A survey in general practitioners and paediatricians.* Eur Child Adolesc Psychiatry, 2016. **25**(4): p. 443-51.
3. Kazdin, A. E., *Evidence-based assessment for children and adolescents: issues in measurement development and clinical application.* J Clin Child Adolesc Psychol, 2005. **34**(3): p. 548-58.
4. Reardon, T., Harvey, K., Baranowska, M., O'Brien, D., Smith, L., and Creswell, C., *What do parents perceive are the barriers and facilitators to accessing psychological treatment for mental health problems in children and adolescents? A systematic review of qualitative and quantitative studies.* Eur Child Adolesc Psychiatry, 2017. **26**(6): p. 623-647.
5. Haller, D. M., Sanci, L. A., Sawyer, S. M., and Patton, G. C., *The identification of young people's emotional distress: a study in primary care.* Br J Gen Pract, 2009. **59**(560): p. e61-70.
6. Bernstein, G. A., Bernat, D. H., Davis, A. A., and Layne, A. E., *Symptom presentation and classroom functioning in a nonclinical sample of children with social phobia.* Depress Anxiety, 2008. **25**(9): p. 752-60.
7. Coles, M. E., Turk, C. L., Jindra, L., and Heimberg, R. G., *The path from initial inquiry to initiation of treatment for social anxiety disorder in an anxiety disorders specialty clinic.* Journal of Anxiety Disorders, 2004. **18**(3): p. 371-383.
8. Yonek, J., Lee, C. M., Harrison, A., Mangurian, C., and Tolou-Shams, M., *Key Components of Effective Pediatric Integrated Mental Health Care Models: A Systematic Review*. JAMA Pediatr, 2020.
9. Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E., *Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication.* Arch Gen Psychiatry, 2005. **62**(6): p. 593-602.
10. Currie, C., Roberts, C., Morgan, A., Smith, R., Settertobulte, W., Samdal, O., et al., *Young People's Health in Context Health Behaviour in School-aged Chidren (HBSC) Study: International Report from the 2001/2002 Survey.* . 2004.
11. Merikangas, K. R., Nakamura, E. F., and Kessler, R. C., *Epidemiology of mental disorders in children and adolescents.* Dialogues Clin Neurosci, 2009. **11**: p. 7-20.
12. Wittchen, H. U. and Jacobi, F., *Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies.* Eur Neuropsychopharmacol, 2005. **15**(4): p. 357-76.
13. Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., and Ustun, T. B., *Age of onset of mental disorders: a review of recent literature.* Curr Opin Psychiatry, 2007. **20**(4): p. 359-64.
14. Ramsawh, H. J., Raffa, S. D., Edelen, M. O., Rende, R., and Keller, M. B., *Anxiety in middle adulthood: effects of age and time on the 14-year course of panic disorder, social phobia and generalized anxiety disorder.* Psychol Med, 2009. **39**(4): p. 615-24.
15. Woodward, L. L. and Fergusson, D. M., *Life Course Outcomes of Young People with Anxiety disorders in Adolescence.* Journal of the American Acadaemy of Child and Adolescent Psychiatry 2001. **40**(9): p. 8.

16. Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R., et al., *Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders.* Psychol Med, 2015. **45**(2): p. 345-60.
17. Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., et al., *Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010.* The Lancet, 2012. **380**(9859): p. 2197-2223.
18. Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., et al., *Child and adolescent mental health worldwide: evidence for action.* The Lancet, 2011. **378**(9801): p. 1515-1525.
19. Zechmeister, I., Kilian, R., McDaid, D., and group, M., *Is it worth investing in mental health promotion and prevention of mental illness? A systematic review of the evidence from economic evaluations.* BMC Public Health, 2008. **8**: p. 20.
20. Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., et al., *Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A).* J Am Acad Child Adolesc Psychiatry, 2011. **50**(1): p. 32-45.
21. Wang, P. S., Berglund, P., Olfson, F., Pincus, H. A., Wells, K. B., and Kessler, R. C., *Failure and delay in initial treatmnet contact after first onset of mental disorders in the national comorbidity survey replication.* Archives of General Psychiatry, 2005. **62**: p. 603-613.
22. Goldberg, D. and Huxley, P., *Mental Illness in the Community: The Pathway to Psychiaric Care*. 1980, London/New York: Tavistock.
23. Zwaanswijk, M., *Pathways to Care: Help-seeking for child and adolescent mental health problems*, in *Academic dissertation*. 2005: Utrecht: NIVEL.
24. Huxley, P., *Mental illness in the community: The Goldberg-Huxley model of the pathway to psychiatric care.* Nordic Journal of Psychiatry, 2009. **50**(sup37): p. 47-53.
25. Zwaanswijk, M., van Dijk, C. E., and Verheij, R. A., *Child and adolescent mental health care in Dutch general practice: time trend analyses.* BMC Fam Pract, 2011. **12**: p. 133.
26. Zwaanswijk, M., Van der Ende, J., Verhaak, P. F., Bensing, J. M., and Verhulst, F. C., *The different stages and actors involved in the process leading to the use of adolescent mental health services.* Clin Child Psychol Psychiatry, 2007. **12**(4): p. 567-82.
27. Logan, D. E. and King, C. A., *Parental facilitation of adolescent mental health service utilization: A conceptual and empirical review.* Clinical Psychology: Science and Practice, 2001. **8**(3): p. 319-333.
28. Costello, E. J., Pescosolido, B. A., Angold, A., & Burns, B. J., *A family network-based model of access to child mental health services.* Research in community and mental health, 1998. **9**: p. 165-190.
29. Verhulst, F. and Koot, H., *Child Psychiatric Epidemiology: Concepts, Methods, and Findings*. 1992.
30. Ministerie van Volksgezondheid Welzijn en Sport & Ministerie van Veiligheid en Justitie, *Jeugdwet [Youth Act]*. 2013.
31. Jeugdwet. 2014; Available from: https://wetten.overheid.nl/BWBR0034925/2018-01-01.
32. Raad van State. *Jeugd-wet.* W13.13.0096/III 2013 16 February 2022]; Available from: https://www.raadvanstate.nl/@62589/w1-3-13-0096-iii/.

33. Friele, R. D., Hageraats, R., Fermin, A., Bouwman, R., Zwaan J. van der *De jeugd-GGZ na de Jeugdwet: een onderzoek naar knelpunten en kansen*. 2019.
34. Netherlands Youth Institute, *Reform of the Dutch system for child and youth care. 4 years later.* 2019: Utrecht.
35. Fickel, J. J., Parker, L. E., Yano, E. M., and Kirchner, J. E., *Primary care - mental health collaboration: an example of assessing usual practice and potential barriers.* J Interprof Care, 2007. **21**(2): p. 207-16.
36. Nooteboom, L. A., van den Driesschen, S. I., Kuiper, C. H. Z., Vermeiren, R., and Mulder, E. A., *An integrated approach to meet the needs of high-vulnerable families: a qualitative study on integrated care from a professional perspective.* Child Adolesc Psychiatry Ment Health, 2020. **14**: p. 18.
37. Nooteboom, L. A., Mulder, E. A., Kuiper, C. H. Z., Colins, O. F., and Vermeiren, R., *Towards Integrated Youth Care: A Systematic Review of Facilitators and Barriers for Professionals.* Adm Policy Ment Health, 2021. **48**(1): p. 88-105.
38. Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., and Richardson, W. S., *Evidence based medicine: what it is and what it isn't.* BMJ, 1996. **312**(7023): p. 71-2.
39. Mash, E. J. and Hunsley, J., *Evidence-based assessment of child and adolescent disorders: issues and challenges.* J Clin Child Adolesc Psychol, 2005. **34**(3): p. 362-79.
40. Practice, A. P. A. P. T. F. o. E.-B., *Evidence-based practice in psychology.* Am Psychol, 2006. **61**(4): p. 271-85.
41. Kosten, T. A. and Rounsaville, B. J., *Sensitivity of psychiatric diagnosis based on the best estimate procedure.* Am J Psychiatry, 1992. **149**(9): p. 1225-7.
42. Nordgaard, J., Sass, L. A., and Parnas, J., *The psychiatric interview: validity, structure, and subjectivity.* Eur Arch Psychiatry Clin Neurosci, 2013. **263**(4): p. 353-64.
43. Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., et al., *Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment.* Clinical Psychology: Science and Practice, 2017. **24**(4): p. 331-363.
44. Doss, A. J., *Evidence-based diagnosis: incorporating diagnostic instruments into clinical practice.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(9): p. 947-52.
45. Zwaanswijk, M., van der Ende, J., Verhaak, P. F., Bensing, J. M., and Verhulst, F. C., *Help-seeking for child psychopathology: pathways to informal and professional services in the Netherlands.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(12): p. 1292-300.
46. Kramer, T. and Garralda, M. E., *Child and adolescent mental health problems in primary care.* Advances in Psychiatric Treatment, 2018. **6**(4): p. 287-294.
47. Jorm, A. F., Wright, A., and Morgan, A. J., *Where to seek help for a mental disorder? National survey of the beliefs of Australian youth and their parents.* Med J Aust, 2007. **187**(10): p. 556-60.
48. Statistiek, C. B. v. d., *Jeugdhulp 2019*. 2020: Den Haag/Heerlen/Bonaire.
49. Statistiek, C. B. v. d., *Jeugdhulptrajecten in natura; verwijzer, regio (gemeente), 2015-2020*. 2021.
50. Sampson, R., Cooper, J., Barbour, R., Polson, R., and Wilson, P., *Patients' perspectives on the medical primary-secondary care interface: systematic review and synthesis of qualitative research.* BMJ Open, 2015. **5**(10): p. e008708.

51. Sayal, K., Tischler, V., Coope, C., Robotham, S., Ashworth, M., Day, C., et al., *Parental help-seeking in primary care for child and adolescent mental health concerns: qualitative study.* Br J Psychiatry, 2010. **197**(6): p. 476-81.
52. Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., et al., *Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries.* Depress Anxiety, 2018. **35**(3): p. 195-208.
53. Essau, C. A., *Frequency and patterns of mental health services utilization among adolescents with anxiety and depressive disorders.* Depress Anxiety, 2005. **22**(3): p. 130-7.
54. Chavira, D. A., Stein, M. B., Bailey, K., and Stein, M. T., *Child anxiety in primary care: prevalent but untreated.* Depress Anxiety, 2004. **20**(4): p. 155-64.
55. Ramsawh, H. J., Chavira, D. A., and Stein, M. B., *Burden of anxiety disorders in pediatric medical settings: prevalence, phenomenology, and a research agenda.* Arch Pediatr Adolesc Med, 2010. **164**(10): p. 965-72.
56. Ford, T., Hamilton, H., Meltzer, H., and Goodman, R., *Child Mental Health is Everybody's Business: The Prevalence of Contact with Public Sector Services by Type of Disorder Among British School Children in a Three-Year Period.* Child and Adolescent Mental Health, 2007. **12**(1): p. 13-20.
57. Andrews, G., *It Would be Cost-Effective to Treat More People with Mental Disorders.* Australian & New Zealand Journal of Psychiatry, 2006. **40**: p. 613-615.
58. Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., et al., *Scaling-up treatment of depression and anxiety: a global return on investment analysis.* The Lancet Psychiatry, 2016. **3**(5): p. 415-424.
59. Bittner, A., Egger, H. L., Erkanli, A., Jane Costello, E., Foley, D. L., and Angold, A., *What do childhood anxiety disorders predict?* J Child Psychol Psychiatry, 2007. **48**(12): p. 1174-83.

When light meets our eyes

we only see colours we are susceptible to

Chapter 2

Recognition of anxiety disorders in children: a cross-sectional vignette-based survey among general practitioners

Published as: Aydin S., Crone M.R., Siebelink B.M., Vermeiren R.R.J.M., Numans M.E., Westenberg P.M. Recognition of anxiety disorders in children: a cross-sectional vignette based survey among general practitioners. BMJ Open, 2020, doi:10.1136/bmjopen-2019-035799

Unnoticed anxiety disorders (ADs) have an important impact on human development and well-being. Only one in ten children with ADs, including those with subthreshold levels of severity, will be free of any mental health problem later in life.[1-7] With estimated prevalence rates of up to one in four, ADs are the most common mental health issue over the lifespan.[8-10] Fortunately, treatment helps to mitigate risks and adversities.[11-14] Based on their gate-keeper role, approachability, and their continuous care relationship with families, general practitioners (GPs) are in a unique position to recognise ADs characterized by an early age of onset, a chronic or relapsing-remitting course, somatic manifestations and comorbidities.[15-20] Although around 70% of children and adolescents consult their GP more than twice a year,[21-24] studies in primary healthcare indicate that over two thirds of children with ADs go unnoticed.[25-28]

Recognition is especially problematic in early onset, less severe and less explicit disorders.[28-31] Interestingly those factors are inherent to anxiety, including the covert nature of its core symptoms, its gradual development, fluctuating course, and the wide variety of accompanying symptoms that are not prototypic for anxiety.[3] ADs are tied to problems as temper tantrums, control neediness, withdrawn behaviour, difficulties in connecting with others, poor concentration and physical complaints.[32] Each of these indicators are seemingly unrelated and may not equally impact a child's situation, increasing the chances that professionals fail to appreciate them as a relevant or important sign. These symptoms also mimic and co-occur with other mental health disorders and contextual problems. Parents and teachers show differences in recognition of various symptoms and tend to underdetect or misattribute some more than the others.[33-36]

Given the limited time available to understand and act on the breadth of disclosed problems, a GP's first interpretation of the presented problems might be pivotal in successfully navigating a consultation and recognising anxiety in children. The decision-making literature highlights the importance of the first diagnostic opinion and suggests that amendments after this point are often insufficient ("anchoring bias").[37] However, studies that relate recognition to the variety of symptoms are scarce, and to our knowledge no study has investigated in how far physicians even consider anxiety as the probable factor underlying the reason of consult. To address this, we presented GPs with *mixed-anxiety vignettes*. That is, vignettes that capture the variety and indefinite nature of early symptom presentation in general

practice and do not solely describe specific mental health disorders as demarcated in psychiatric nosologies. Notwithstanding that we created the vignettes to suggest a probable AD, we hypothesised that GPs would not proportionately cite anxiety as the underlying cause of the presented symptoms. In addition to the question how GPs would interpret and refer the described cases, we also asked how they generally tend to manage cases when they suspect specific mental health problems in practice. Based on literature suggesting that ADs might be evaluated as relatively benign,[38] we hypothesised that GPs would be less likely to report that they refer children with suspected anxiety to mental healthcare (MHC) than children with suspected developmental, behavioural or mood problems.

Methods

Setting and Design

The study was conducted during the yearly 2-day continuing medical education conference "Boerhaave Progress and Practice"[39] for primary healthcare physicians (Leiden, The Netherlands, 14 December 2018). Each year a range of somatic and mental health topics are covered by specialists within 25-minute time slots. During one of those slots, we presented the survey supported by PowerPoint slides that additionally carried embedded subtitled audio fragments of the vignettes (60-64 seconds each). A total of 275 GPs were registered. Attendees were seated with a device that digitally recorded their anonymous responses. As such, data collection was not subject to Dutch law governing medical research involving human subjects, nor to European law on general data protection and privacy.

Procedure and measurement

Attendees were introduced with I) that they would hear five case descriptions of children with psychosocial problems in the form of a report resulting from a few consultations, and would see three recurring questions that reflect their daily reality concerning what the condition might be and what could be done about it, II) they were asked to base their considerations on the available information only. We aimed to use gender-neutral names for the vignettes to prevent that gender confounds decision-making. In an independent coding of 25 gender-neutral names by six coders, at most three names were labelled by half of the coders as gender-neutral. Given this low agreement and taking into account discrepancies in

literature on gender differences and recognition,[29 36] we mentioned III) that the named cases were intentionally gender-neutral. Attendees were also informed IV) that by responding they voluntarily and anonymously agreed to analysis of their responses for scientific purposes. Preceding the first vignette, the recurring questions were verbally introduced and shown (Box 1). Thereafter, the questions were shown sequentially, each with a progress bar displaying the response time set to a maximum of ten seconds.

Vignette development

In order to quantify to what extend GPs consider anxiety in their early diagnostic opinion, we constructed vignettes that capture various symptoms (supplementary material). Guided by knowledge of the clinical presentation of ADs and the prevalence of common mental health disorders, we had decided to build five vignettes around the combined presentation of anxiety with, respectively, physical complaints, difficult behaviour, mood, developmental, and school attendance problems. On the level of the key aspect *mixed-anxiety*, each vignette represented symptoms related mostly to one of these domains while also depicting symptoms of anxiety that do not overlap with these mental health problems.

To theoretically support the process, we first consulted clinical handbooks and questionnaires.[40-42] Relevant symptom expressions and idioms were obtained from actual clinical referral letters written by GPs of children and adolescents diagnosed in specialised MHC. This linked text-data enabled working back and forth from children's complaints mentioned by their GPs and their descriptive diagnoses, and hence facilitated a controlled development of vignettes that are also natural. The extracted descriptions were grouped into the five domains and an initial set of 11 vignettes (as advised)[43] was formed. In an iterative process five vignettes were selected, further refined, and for readability purposes named as V1-Somatic, V2-Behavioural, V3-Mood, V4-Developmental, V5-School Attendance. In each vignette more cues were given for anxiety compared with any other specific disorder group (for example, in V2-Behavioural vignette we counted the total number of cues as 13, with eight cues that could be related to anxiety, five to behavioural problems, three to attention-hyperactivity and trauma, two to somatic and one to

Box 1. Survey items as shown and verbally introduced

Questions and options as shown	Questions and options verbally introduced as
A1 What is the main complaint?	Where do you think this description fits in? To which symptom-profile would you ascribe this vignette?
1. Typical development	1. Option one if it is probably an example of typical development.
2. Behavioural problems	2. Option two for difficult behaviour: examples include aggressive behaviour or antisocial behaviour.
3. Complaints regarding establishing contact	3. If problems likely indicate an autism spectrum disorder, you can choose option three.
4. Mood problems	4. Mood problems and problems that could be related to depressive disorders.
5. Somatic complaints	5. For physical symptoms choose option five, also if a problem might be psychosomatic in nature.
6. Eating problems	6. Option six for eating problems and probable eating disorders.
7. Anxiety-related complaints	7. Option seven for problems related to anxiety and anxiety disorders.
8. Complaints regarding attention and activity	8. Option eight for attention-related complaints that might indicate attention deficit hyperactivity disorder or attention deficit disorder.
9. Complaints related to the experience of a traumatic event	9. Option nine for problems related to the experience of a traumatic event.
A2 Second complaint group?	If you would like to add a second problem to the main complaint groups you can select it here.
10. No second complaint group	10. Please choose option ten on your voting machine if you do not see another complaint.
A3 Where should this child ideally go?	Then, for each vignette, the triage question: where can this child and the family get the most adequate professional support?
1. Practice Nurse	
2. Local Youth teams	
3. Generalized Mental Healthcare	
4. Specialized Mental Healthcare	
5. Somatic Healthcare/ Hospital	
Where do you think that children with this type of complaint are best helped?	The eight mental health groups will be shown again. Please indicate for each of these how you generally tend to refer children when you suspect these complaints.
B1. Behavioural problems	
B2. Complaints regarding establishing contact	
B3. Mood problems	
B4. Somatic complaints	

B5. Eating problems
B6. Anxiety-related complaints
B7. Complaints regarding attention and activity
B8. Complaints related to the experience of a traumatic event

1. Watchful waiting
2. Practice Nurse
3. The local youth teams
4. General mental healthcare
5. Specialized mental healthcare

If you often tend to first monitor such complaints for a while, you can opt for ”1”. If you often ask your mental health practice nurse to become involved: option ”2”, and so on, up to the point where you feel that a more intense treatment option is adequate. Then you can choose option ”5”: specialized mental healthcare.

Survey items A1 to A3 were presented after each vignette. Given their relevance for early recognition, the response options for the identification questions (A1 and A2) were formulated in terms of “complaints” and “problems” rather than “disorders”. We presented two identification questions to provide sufficient opportunity to identify an AD in the mixed vignettes. The third recurring question queried how the respondents would have referred such a child or family in daily practice, and was presented with response options that reflect the Dutch “stepped care” services. The practice nurse is a mental health professional who works in general practice. The option “hospital” was presented since the vignettes also depicted physical problems. As we introduced the study as one about children *with* psychosocial problems, the option watchful waiting was not presented for the vignettes. The survey items B1 to B8 were presented after the last survey item A3. Here, the option watchful waiting was presented following our interest in whether ADs are viewed as something children grow out of.

autism). To mimic actual consultations, the vignettes also included contextual factors[43] such as academic or home functioning or background information about the GP-child relationship. Case age and vignette length had a limited range (10-12 years, 165-172 words) to avoid possible confounding effects.[43]

During the final stage of vignette development, GPs affiliated with Leiden University Medical Centre were invited to participate in an individual pilot interview (six interviews held in November, 2018). GPs were asked to think aloud and provide verbal feedback while reading the transcript of each vignette and answering the identification and referral questions. We also asked GPs about the realism of the vignettes: they responded positively and felt that the descriptions corresponded to their practice. The pilot led to minor changes, most concerning the “contextual aspects”,[43] for example, initially we alternated mothers and fathers as the parent who went along to the consultation. Several GPs evaluated this as odd, hence we changed it to “mother”.

Validation of vignettes

To validate the recognisability of anxiety, 24 mental health professionals (MHPs) were invited to participate in a pilot presentation, of whom 11, all with extensive clinical experience (median 15-19 years), agreed to participate. All MHPs were employed at Curium-LUMC, a university-affiliated inpatient and outpatient psychiatric clinic for children and adolescents. We surveyed MHPs using the same procedure as in the GP-survey, with the exception of written responses and two additional questions regarding their psychiatric specialisation. The presentation resulted in 55 answers concerning the first identification question, and 53 answers on the second identification question. In line with our aim to develop mixed-anxiety vignettes, MHPs selected different disorder groups: anxiety (n=22), mood (n=19), attention-hyperactivity (n=14), behavioural (n=12) and somatic (n=10) problems were selected frequently, and typical development was not selected (supplementary material). Anxiety was the most or second most frequently selected disorder group in V1-Somatic, V2-Behavioural and V4-Developmental, and the third most frequently selected in V3-Mood and V5-School Attendance.

Outcome definition and analysis

Recognition rate was defined as the ratio of the total number of times anxiety was selected and the total number of times anxiety could have been selected. GPs' and MHPs' recognition was compared following Newcombe's method 10[44] using the Epi package[45] in the R-statistics environment. This method tests for independence as in the chi-square test, yet provides a robust method to estimate confidence intervals. GPs' selection rate of the different disorder categories and the referral indications they made for the vignettes were analysed using descriptive statistics. Whether the likelihood of a MHC referral was lower when anxiety was recognised in a vignette was tested in a logistic multilevel regression using the lme4 package.[46] That GPs would report a MHC referral for ADs less often compared with referral for behavioural, developmental and mood problems was also tested in a logistic multilevel regression, with disorder groups (anxiety versus behaviour, development, and mood problems) as the repeated measures. Respondents with missing responses on more than half of the survey were excluded from all analyses. Data were visualized using the ggplot2 package.[47]

Results

Participants

A total of 299 attendees answered at least one survey question. Of these, 239 reported being a GP. Data from ten GPs were excluded due to missing responses. As a missing response on both identification questions of a vignette occurred 17 times, the number of times anxiety could have been selected was 1128 (229*5-17).

The distribution of the sample's gender and years of experience was similar to the national population[48] (Table 1). These demographics showed no association with the number of times anxiety was recognised, the referral indications GPs made for each vignette, or with their reported referral preferences for ADs (supplementary material).

Table 1. Demographics of the study sample and the national population

	Study sample N=229	Population N=9798	
Gender			
Male	116 (50.7 %)	4799 (49.0%)	
Female	95 (41.5 %)	4999 (51.0%)	
Unknown	18 (7.9 %)	-	
Experience in years		**Age distribution in national population**	
0-2	4 (1.7 %)	< 30	76 (0.8%)
3-5	22 (9.6 %)	30-34	976 (9.9%)
6-9	27 (11.8 %)	35-39	1428 (14.6%)
10-14	16 (7.0 %)	40-44	1518 (15.5%
15-19	35 (15.3 %)	45-49	1396 (14.2%)
>20	124 (54.1 %)	> 50	4357 (44.5%)
Unknown	1 (0.4 %)	Unknown	56 (0.6%)

Recognition

By selecting 167 times anxiety in the 1128 possibilities to opt for a specific disorder group, GPs reached a recognition rate of 14.8%. GPs were less likely to recognise anxiety compared with MHPs (OR= 0.26, 95% CI 0.15 to 0.46, Figure 1). Whereas 55.9% (n=128) of the GPs labelled one or more vignette as anxiety-related, over two fifths did not in any. On average, GPs often selected behavioural, mood, developmental, and trauma-related problems, and typical development (Table 2). GPs and MHPs did not differ in their selection rate of these disorder groups (supplementary material).

Exploration of recognition per vignette revealed a major difference between GPs and MHPs for V1-Somatic (GPs: 5.7%, MHPs: 54.5%). Differences were not statistically significant for V2-Behavioural (GPs: 33.8%, MHPs: 45.5%), and for V3-Mood (GPs: 13.7%, MHPs: 18.2%). GPs recognised anxiety significantly less often than MHPs in V4-Developmental (GPs: 9.0%, MHPs: 36.4%) and in V5-School Attendance (GPs: 11.9%, MHPs: 45.5%).

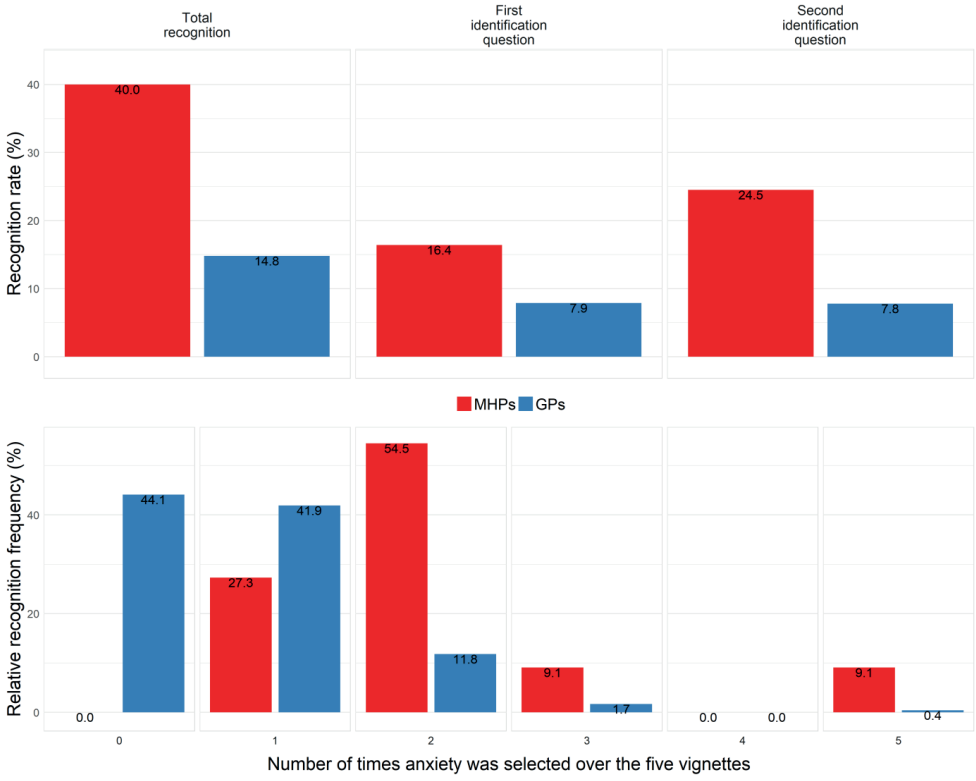


Figure 1. Overall recognition rate and recognition frequency of anxiety

The upper figure depicts MHPs' and GPs' recognition rate of anxiety. MHPs selected anxiety in nine of their 55 responses on the first identification question and in 13 of their 53 responses on the second identification question. GPs selected anxiety in 84 of their 1060 responses on the first identification question and in 83 of their 1067 responses on the second identification question. Differences between the two groups were statistically significant overall (OR= 0.26, 95% CI 0.15 to 0.46), in the first identification question (OR= 0.44, 95% CI 0.21 to 0.93), and in the second identification question (OR=0.26, 95% CI 0.13 to 0.50). The lower figure depicts the percentage of GPs and MHPs that recognised anxiety in none, one, two, three, four, or all five vignettes. The difference between GPs' and MHPs' recognition frequency was statistically significant, $\chi 2_{(5, n=240)}$= 42.94, p< .001.

Table 2. GPs' selection rate of each disorder category

	V1-Somatic		V2-Behavioural		V3-Mood		V4-Developmental		V5-School Attendance	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Anxiety	4.0	1.9	22.4	14.2	4.6	9.8	3.4	6.1	5.8	6.9
Trauma	0.9	1.4	46.3	29.2	0.5	-	1.0	0.5	9.7	24.1
Mood	1.3	1.9	1.0	2.8	52.8	22.8	-	2.8	38.6	11.6
Somatic	8.8	14.6	-	0.5	0.9	0.9	0.5	-	9.2	5.1
Eating	5.3	10.4	-	-	-	-	-	-	4.3	7.9
Autism	10.6	11.8	1.0	-	32.6	27.9	8.3	17.9	4.8	4.2
Attention-hyperactivity	7.1	3.3	6.3	11.8	4.1	13.0	67.6	16.5	1.9	1.9
Difficult behaviour	29.6	14.2	22.0	32.1	3.7	6.0	15.7	22.2	6.8	3.7
Typical development	32.3	12.7	1.0	-	0.9	0.9	3.4	7.5	18.8	4.2
No 2nd complaint group		29.2		9.4		18.6		26.4		30.6

Selection rate of each disorder group shown in percentages, per identification question (first and second) and per vignette (V1 to V5). Missing responses resulted in the following sample sizes in V1-Somatic= 226 and 212, V2-Behavioural= 205 and 212, V3-Mood= 218 and 215, V4-Developmental= 204 and 212, V5-School Attendance= 207 and 216, with the first value depicting the sample size in the first identification question, and the latter the second identification question of each vignette.

Referral

When asked how they would refer a child with a similar profile to the vignettes, the majority responded that they would keep the child in general practice (practice nurse= 40%, Figure 2). GPs who indicated that they would refer most often chose primary healthcare (local youth teams= 24%, primary MHC= 13%). The vignette which most often triggered referral to MHC was V3-Mood.

Those GPs who recognised anxiety reported referral of a child to MHC slightly less often compared with those who did not recognise anxiety (Table 3, supplementary material).

However, multilevel analysis revealed that the likelihood of a referral to MHC was not significantly related to the recognition of anxiety (OR= 0.70, 95% CI 0.42 to 1.18).

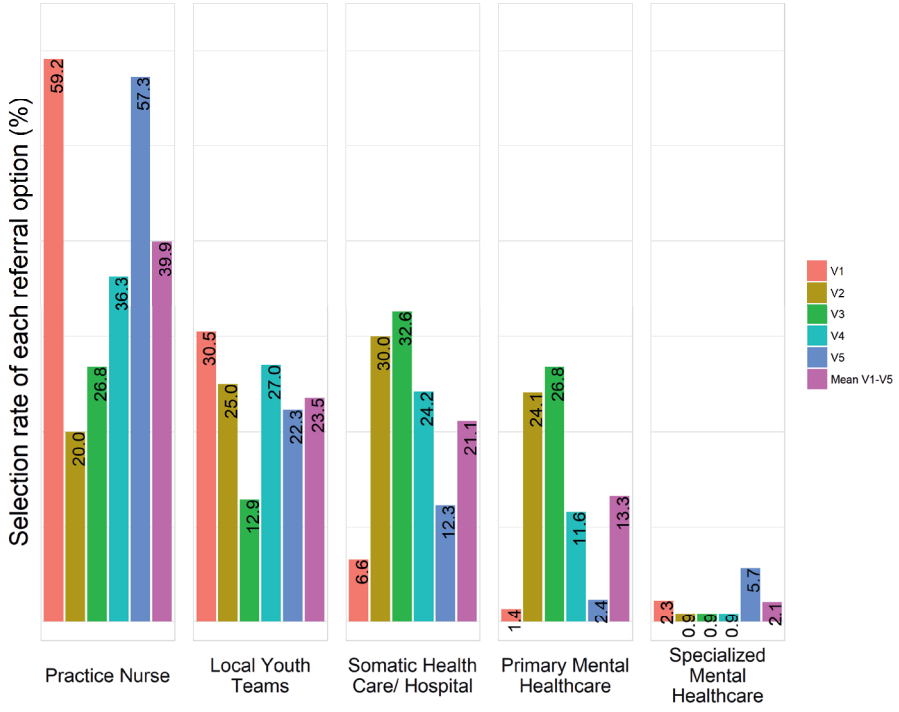


Figure 2. GPs' referral decisions following each of the five vignettes

Sample size was as follows in V1-Somatic= 213, V2-Behavioural= 220, V3-Mood= 224, V4-Developmental= 215, V5-School Attendance= 211.

When asked about their general referral preferences, a majority reported a tendency to refer children with suspected ADs to primary MHC (Figure 3). More GPs reported treatment of suspected anxiety in MHC (OR= 1.79, 95% CI 1.35 to 2.37) compared with suspected developmental, behavioural or mood problems (OR= 0.53, 95% CI 0.39 to 0.72). Analysis per disorder group revealed that GPs' reported referral preferences for anxiety, did not differ from their preferences for mood problems. A referral to MHC was reported less often for difficult behaviour, problems related to autism, attention-hyperactivity and (psycho-)somatic problems, and more often for eating problems and trauma (supplementary material).

Table 3. Referral to mental healthcare by those GPs who recognised anxiety and those who did not

		Other referral options	Mental healthcare
V1-Somatic	Selected	13 (100%)	0 (0%)
	Not selected	192 (96.0%)	8 (4.0%)
V2-Behavioural	Selected	59 (78.7%)	16 (21.3%)
	Not selected	105 (73.9%)	37 (26.1%)
V3-Mood	Selected	24 (80.0%)	6 (20.0%)
	Not selected	137 (71.0%)	56 (29.0%)
V4-Developmental	Selected	17 (89.5%)	2 (10.5%)
	Not selected	166 (86.9%)	25 (13.1%)
V5-School Attendance	Selected	22 (91.7%)	2 (8.3%)
	Not selected	170 (91.9%)	15 (8.1%)
Mean V_1-V_5	Selected	88.0%	12.0%
	Not selected	83.9%	16.1%

Selection frequency of each referral option per vignette partitioned by GPs who selected anxiety and who did not. Data revealed no significant associations between recognition of anxiety and referral to MHC (OR= 0.70, 95% CI 0.42 to 1.18, p= .19). A MHC referral includes referral to primary MHC and specialized MHC, other referral options includes the mental health practice nurse in general practice, local youth teams and somatic healthcare (selection rates of each specific referral option are depicted in supplementary material).

Discussion

To investigate GPs' sensitivity towards ADs, we presented vignettes describing the mixture of problems that might be disclosed during consultations. GPs noticed anxiety sufficiently in the vignette that depicted a prototypic sign ("fears") but infrequently when the depicted symptoms were less explicit. Possible alternative mental health problems were recognised also by MHPs who helped to validate the vignettes but they differed from GPs only in their better recognition of anxiety. Whether a GP indicated a referral to MHC for a given vignette showed no significant association with recognition of anxiety in that vignette. MHC was selected by GPs as the appropriate referral option in less than one fifth of these vignettes depicting problems of mild severity. Interestingly, when asked about their general referral preferences for children with suspected ADs, over two thirds of the GPs reported a preference for intervention in MHC.

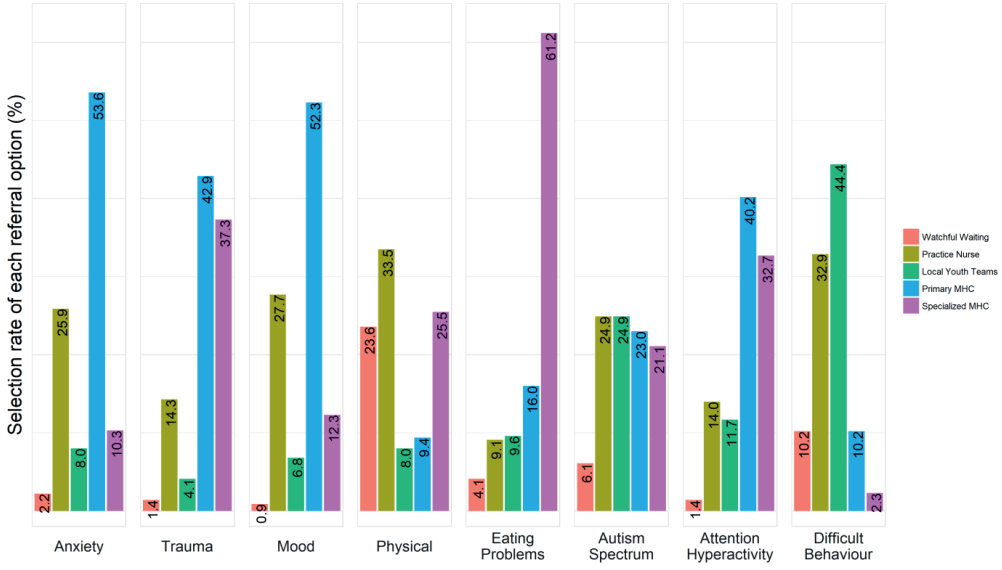


Figure 3. GPs' reported referral preferences for each of the eight disorder groups

Excluding cases with missing responses left the following sample sizes: anxiety n= 224, trauma n= 217, mood n= 220, somatic n= 212, eating problems n= 219, autism n= 213, attention-hyperactivity n= 214, difficult behaviour n= 216.

Strengths and limitations of the study

To the best of our knowledge this is the first study to investigate GPs' interpretations of the various problems typical in children with ADs. We surveyed a large sample but we do not have insight into the representativeness of the attending GPs. Nevertheless, the years of experience and gender distribution of the sample were similar to the national GP population, these demographics did not show an association with the outcome measures.[32] Also, literature is inconclusive in whether GPs' attitudes influence decision-making.[30 31] The presentation of the vignettes to all participants at once yielded a risk of carryover effects. It also incited a game aspect that may have decreased observer effects and improved access to participants' honest and direct considerations.[43] A limitation is the small number of MHPs and the multiple statistical comparisons made between GPs and MHPs. Notwithstanding the explorative nature of the study on the level of the individual vignettes, we applied conservative Bonferroni corrections to avoid false positives. MHPs selected anxiety infrequently in the mixed anxiety-depression vignette. This low recognition rate might be a

consequence of similar number of cues for anxiety (8) and depression (7) in this vignette, the large conceptual overlap between anxiety and depression, and that MHPs might also be more attuned towards depression.[49] GPs could not indicate watchful waiting for the vignettes in this study overestimating the likelihood that GPs take any action when presented with children similar to the vignettes. In conversations with GPs we had found that in case of any ambiguity they would make the rational decision for further enquiry. Given our aim to gain insight in their tendencies, we decided to prevent selection of this safe option. Nevertheless, it is unlikely that addition of this option would change the results considering referral to MHC. Those who would have selected watchful waiting now probably selected the practice nurse. A related caveat is the interpretation of the findings as representing recognition and referral in everyday practice. Decision-making in general practice is shaped by a complex interplay of multiple individual and systemic factors over time. With this novel approach we aimed to isolate the extent to which GPs mind anxiety in their early diagnostic opinion. In this regard, a major strength of the study was the use of actual patient medical records to create the vignettes. This enabled wording similar to that of GPs, improving the construct and external validity of the case descriptions, and diminishing uncontrollable variability.[50]

Comparison with other studies

Overall recognition of anxiety was within the lower range of what is found in literature (9-60%).[27 28 51] This might be a consequence of the mild severity portrayed in the vignettes, a factor limiting recognition of all disorder groups.[32] This design, however, does not explain why GPs selected the other mental health problems yet failed to recognise anxiety proportionately. In V1-Somatic depicting a consultation for recurring earaches, for example, GPs picked up on the probability of behavioural, somatic, or autism spectrum-related problems. They were also sensitive to whether typical development was depicted. Nonetheless, the portrayal should have also raised the prospect of an AD, a possibility that was recognised by very few GPs. GPs recognised anxiety sufficiently in one vignette. This being the vignette that literally mentioned the core feature of anxiety (“fears”) supports that unfamiliarity with symptom presentation might be the driving factor behind the low recognition rates.

The recognition rate was realized by just over half of GPs. Others did not select anxiety in any vignette. A question is what underlies these individual differences. Previous studies also found no association of recognition with gender or experience of the professional, and literature is not one-sided about the role of attitude.[31] Observations we made during the pilot with six GPs suggests that differences might be caused by how they "perceive, read and interpret"[37] presenting symptoms. For instance, all GPs commented about the word "temperamental", yet many neglected "timid" although both words were in the same sentence. GPs also varied in the problem areas they gave attention to. Most GPs expressed a focus on school and home burden (e.g. "School is a benchmark for me and that is going well. If it wasn't alarm bells would be going off"). By contrast many did not reveal a need to enquire further with children themselves or about leisure activities and social relations. If a GP tends to piece out some symptom-profiles and neglect others based on context,[52 53] this might be of particular consequence for the recognition of ADs characterized by a covert and less disruptive profile.

An important finding concerned the final vignette which portrayed a child with social anxiety disorder. In response to this vignette, GPs frequently selected typical development and the mildest referral options. Although a similar response pattern was obtained for the first vignette, both the vignette developers and MHPs evaluated the final vignette as a more severe case description. This finding is in line with studies suggesting low recognition of social anxiety.[29 35 54] Recognition of social anxiety by professionals is of particular importance since these patients are known to underreport symptoms, and since diagnosis-specific interventions are advised in clinical guidelines.[55]

Contrary to hypothesis, GPs did not prove less likely to prefer intervention in MHC for ADs compared with problems related to difficult behaviour, developmental disorders or depression. Results of this normative part of the study contrast with clinical studies demonstrating that a substantial proportion of children with ADs are unlikely to be referred to MHC.[25 30 56] This disparity with clinical practice might be related to factors other than GPs' conscious evaluations, such as strain on parents and their treatment desires that differ between disorders.[57] Another explanation might be a discrepancy in GPs conscious evaluations on the one hand, and their skills and implicit decision-making tendencies on the other hand. Our data supports this explanation to some extent, although to our knowledge

no previous study examined the evaluations and implicit tendencies GPs hold towards treatment of children with ADs. We found a discrepancy in GPs' responses in that the majority of the GPs reported to consider a referral when they suspect ADs, but generally chose management in primary healthcare rather than referral to MHC in the vignettes. This comparison of reported referral preferences to the referral indications made for the vignettes might be invalid if GPs assumed a higher severity when asked about referral preferences versus the mild severity depicted in the vignettes. However, in both, we emphasised that the survey considers probable mental health problems, including subthreshold levels of severity. The variations in GPs answers between disorder groups supports that they understood the question as it was meant. Also, GPs were inclined to refer the vignette scenario with symptoms that overlapped with mood problems to MHC, in agreement with their reported normal practice. This leaves open whether their threshold to evaluate anxiety as an AD differs from when they consider mood problems as depression,[54] yet supports our hypothesis that differences in the likelihood of a referral may depend on symptom presentation. Notably, during the individual pilot interviews, we observed a focus on symptoms related to depression "because you would not want to leave it untreated" and varying heuristics GPs used when evaluating the relevance of symptoms.

Meaning of the study

Moving beyond sensitivity as the endpoint of diagnostic accuracy, the results suggest that the repeatedly found underrecognition of anxiety goes back to the initial considerations GPs make. Our findings confirm the conjecture that recognition of ADs is often confounded by their overlap with other common and well-known mental health problems and similarity with typical developmental variations.[26] However, this finding per se does not distinguish between the alternate possibilities that GPs prioritise other disorder groups or lack sufficient knowledge of the presentation and prevalence of child anxiety. In fact, it has been suggested that there might be a prioritisation issue given the seemingly low burden and impact of ADs.[38] In the current study, the presentation of the vignettes using subtitled audio fragments precluded the possibility of rereading and elaborate evaluation. By verbally presenting vignettes that do not indicate a distinctive mental health disorder, we aimed to circumvent knowledge of diagnostic criteria and the question of a primary or differential diagnosis. The study results therefore imply that, even before rational prioritisation, there could be a

significant bottleneck in first interpretation that might have otherwise led to early recognition. Anxiety might not resonate sufficiently as the likely problem when GPs encounter children with ADs. As a result, GPs might be inclined to dismiss the initial symptoms as insignificant or to investigate further but with a focus on another disorder group, possibly at the expense of exploring and eliciting relevant symptoms of anxiety. Leaving aside the discussion of whether it is a GP's task to assess the specifics of a child's mental health problem, we argue that it is worthwhile to improve knowledge of early symptom presentation in children with problems that are relatively poorly recognised by parents and teachers. GPs face the difficult task to recognise mental health problems in a limited amount of consult time, yet report difficulties with the use of structured assessment methods in clinical practice.[58] To avail the benefits of their unguided decision-making, it is necessary that they become more attuned and aware of the prevalence, the breadth, and relevance of typical signs of anxiety. Understanding how professionals comprehend and act on presenting symptoms is important, even more so in light of the high base-rate of ADs,[8] their early age of onset (11 years),[18] and risk of persistence into adulthood.[1-3] The finding that majority of the GPs reported to prefer an MHC referral indicates that they do acknowledge anxiety as a treatment demanding problem, and leaves opportunity for targeted improvements.

Unanswered questions and future research

Why did some GPs recognise anxiety in multiple vignettes whereas others did not? Another open question concerns how GPs would have responded had the referral option "no action needed" been available for the vignettes. Causal inferences as regards symptom presentation and subsequent recognition and referral could be made if the many symptoms associated with ADs could be systematically varied using a factorial vignette design.[43] If combined with qualitative methods, future studies could further enhance our understanding of the decision-making process.

Reference list Study 1

1. Beesdo, K., Knappe, S., and Pine, D. S., *Anxiety and anxiety disorders in children and adolescents: developmental issues and implications for DSM-V.* Psychiatr Clin North Am, 2009. **32**(3): p. 483-524.
2. Bittner, A., Egger, H. L., Erkanli, A., Jane Costello, E., Foley, D. L., and Angold, A., *What do childhood anxiety disorders predict?* J Child Psychol Psychiatry, 2007. **48**(12): p. 1174-83.
3. Copeland, W. E., Angold, A., Shanahan, L., and Costello, E. J., *Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study.* J Am Acad Child Adolesc Psychiatry, 2014. **53**(1): p. 21-33.
4. Haller, H., Cramer, H., Lauche, R., Gass, F., and Dobos, G. J., *The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review.* BMC Psychiatry, 2014. **14**: p. 128.
5. Woodward, L. L. and Fergusson, D. M., *Life Course Outcomes of Young People with Anxiety disorders in Adolescence.* Journal of the American Acadaemy of Child and Adolescent Psychiatry 2001. **40**(9): p. 8.
6. Bosman, R. C., Ten Have, M., de Graaf, R., Muntingh, A. D., van Balkom, A. J., and Batelaan, N. M., *Prevalence and course of subthreshold anxiety disorder in the general population: A three-year follow-up study.* J Affect Disord, 2019. **247**: p. 105-113.
7. Pape, K., Bjorngaard, J. H., Holmen, T. L., and Krokstad, S., *The welfare burden of adolescent anxiety and depression: a prospective study of 7500 young Norwegians and their families: the HUNT study.* BMJ Open, 2012. **2**(6).
8. Baxter, A. J., Scott, K. M., Vos, T., and Whiteford, H. A., *Global prevalence of anxiety disorders: a systematic review and meta-regression.* Psychol Med, 2013. **43**(5): p. 897-910.
9. Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., et al., *Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A).* J Am Acad Child Adolesc Psychiatry, 2010. **49**(10): p. 980-9.
10. Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., et al., *Scaling-up treatment of depression and anxiety: a global return on investment analysis.* The Lancet Psychiatry, 2016. **3**(5): p. 415-424.
11. Bennett, K., Manassis, K., Duda, S., Bagnell, A., Bernstein, G. A., Garland, E. J., et al., *Preventing Child and Adolescent Anxiety Disorders: Overview of Systematic Reviews.* Depress Anxiety, 2015. **32**(12): p. 909-18.
12. Warwick, H., Reardon, T., Cooper, P., Murayama, K., Reynolds, S., Wilson, C., et al., *Complete recovery from anxiety disorders following Cognitive Behavior Therapy in children and adolescents: A meta-analysis.* Clin Psychol Rev, 2017. **52**: p. 77-91.
13. Moreno-Peral, P., Conejo-Ceron, S., Rubio-Valera, M., Fernandez, A., Navas-Campana, D., Rodriguez-Morejon, A., et al., *Effectiveness of Psychological and/or Educational Interventions in the Prevention of Anxiety: A Systematic Review, Meta-analysis, and Meta-regression.* JAMA Psychiatry, 2017. **74**(10): p. 1021-1029.
14. Teubert, D. and Pinquart, M., *A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents.* Journal of Anxiety Disorders, 2011. **25**(8): p. 1046-1059.
15. Garralda, E. Child and adolescent psychiatry in general practice. Aust N Z J Psychiatry, 2001. 35(3):308-14.

16. World Health Organization, *The World health report: 2001 Mental health: new understanding, new hope.* 2001: Geneva: World Health Organization.
17. Garralda, M. E., Bowman, F. M., and Mandalia, S., *Children with psychiatric disorders who are frequent attenders to primary care.* European Child & Adolescent Psychiatry, 1999. **8**(1): p. 34-44.
18. Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E., *Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication.* Arch Gen Psychiatry, 2005. **62**(6): p. 593-602.
19. Gieteling, M. J., Lisman-van Leeuwen, Y., Passchier, J., Koes, B. W., and Berger, M. Y., *The course of mental health problems in children presenting with abdominal pain in general practice.* Scand J Prim Health Care, 2012. **30**(2): p. 114-20.
20. Rockhill, C., Kodish, I., DiBattisto, C., Macias, M., Varley, C., and Ryan, S., *Anxiety disorders in children and adolescents.* Curr Probl Pediatr Adolesc Health Care, 2010. **40**(4): p. 66-99.
21. Zwaanswijk, M., Verhaak, P. F., van der Ende, J., Bensing, J. M., and Verhulst, F. C., *Consultation for and identification of child and adolescent psychological problems in Dutch general practice.* Fam Pract, 2005. **22**(5): p. 498-506.
22. Bayram, C., Harrison, C., Charles, J., and Britt, H., *'The kids are alright' – Changes in GP consultations with children 2000–15.* Australian Family Physician. **44**(12): p. 2.
23. Roberts, J., Sanci, L., and Haller, D., *Global adolescent health: is there a role for general practice?* Br J Gen Pract, 2012. **62**(604): p. 608-10.
24. Netherlands., S. *Health and health care; General practice consults by personal characteristics*. 2019 [cited 2019 18 may]; Available from: https://opendata.cbs.nl/statline/#/CBS/en/dataset/83005ENG/table?dl=20824.
25. Hyde, J., Evans, J., Sharp, D., Croudace, T., Harrison, G., Lewis, G., et al., *Deciding who gets treatment for depression and anxiety: a study of consecutive GP attenders.* British Journal of General Practice, 2005. **55**(220): p. 846-853.
26. Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P.O., Löwe, B. Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. Annals of Internal Medicine. 2007;146(5).
27. Olariu, E., Forero, C. G., Castro-Rodriguez, J. I., Rodrigo-Calvo, M. T., Alvarez, P., Martin-Lopez, L. M., et al., *Detection of Anxiety Disorders in Primary Care: A Meta-Analysis of Assisted and Unassisted Diagnoses.* Depress Anxiety, 2015. **32**(7): p. 471-84.
28. Chavira, D. A., Stein, M. B., Bailey, K., and Stein, M. T., *Child anxiety in primary care: prevalent but untreated.* Depress Anxiety, 2004. **20**(4): p. 155-64.
29. Janssen, E. H., van de Ven, P. M., Terluin, B., Verhaak, P. F., van Marwijk, H. W., Smolders, M., et al., *Recognition of anxiety disorders by family physicians after rigorous medical record case extraction: results of the Netherlands Study of Depression and Anxiety.* Gen Hosp Psychiatry, 2012. **34**(5): p. 460-7.
30. Steele, M. M., Lochrie, A. S., and Roberts, M. C., *Physician identification and management of psychosocial problems in primary care.* J Clin Psychol Med Settings, 2010. **17**(2): p. 103-15.
31. Koning, N. R., Buchner, F. L., Verbiest, M. E. A., Vermeiren, R., Numans, M. E., and Crone, M. R., *Factors associated with the identification of child mental health problems in primary care-a systematic review.* Eur J Gen Pract, 2019. **25**(3): p. 116-127.
32. Tylee, A., Walters, P. Underrecognition of anxiety and mood disorders in primary care: Why does the problem exist and what can be done? Journal of Clinical Psychiatry. 2007;68(2):27-30.

33. Reardon, T., Harvey, K., Young, B., O'Brien, D., and Creswell, C., *Barriers and facilitators to parents seeking and accessing professional support for anxiety disorders in children: qualitative interview study.* Eur Child Adolesc Psychiatry, 2018. **27**(8): p. 1023-1031.
34. Neil, L. and Smith, M., *Teachers' recognition of anxiety and somatic symptoms in their pupils.* Psychology in the Schools, 2017. **54**(9): p. 1176-1188.
35. Herbert, J. D., Crittenden, K., and Dalrymple, K. L., *Knowledge of social anxiety disorder relative to attention deficit hyperactivity disorder among educational professionals.* J Clin Child Adolesc Psychol, 2004. **33**(2): p. 366-72.
36. Furnham, A. and Lousley, C., *Mental health literacy and the anxiety disorders.* Health, 2013. **05**(03): p. 521-531.
37. Elstein, A. S. and Schwartz, A., *Clinical problem solving and diagnostic decision making: selective review of the cognitive literature.* BMJ, 2002. **324**(7339): p. 729-32.
38. O'Brien, D., Harvey, K., Young, B., Reardon, T., and Creswell, C., *GPs' experiences of children with anxiety disorders in primary care: a qualitative study.* Br J Gen Pract, 2017. **67**(665): p. e888-e898.
39. Aydin, S., Westenberg, P. M., and et al. *Psychosocial problems in youths: What does the GP decide?* in *Boerhaave Continuing Medical Education conference: Progress and Practice 13 and 14 December*. 2018. Leiden, The Netherlands.
40. Carr, A., *The Handbook of Child and Adolescent Clinical Psychology: A Contextual Approach.* 1999, Routledge Taylor & Francis Group: London and New York.
41. DAWBA. *Dutch version of the Development and Well-Being Assessment*. [cited 2019 22 january]; Available from: https://www.dawba.info/py/dawbainfo/f1ned.py#b.
42. Goodman, A., Heiervang, E., Collishaw, S., and Goodman, R., *The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples.* Soc Psychiatry Psychiatr Epidemiol, 2011. **46**(6): p. 521-32.
43. Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., et al., *Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies.* Int J Clin Health Psychol, 2015. **15**(2): p. 160-170.
44. Newcombe, R. G., *Interval estimation for the difference between independent proportions: Comparison of eleven methods.* Statistics in Medicine, 1998. **17**(8): p. 873-890.
45. Stevenson, M., *epiR: Tools for the Analysis of Epidemiological Data.* 2019.
46. Bates, D., Machler, M., Bolker, B. M., and Walker, S. C., *Fitting Linear Mixed-Effects Models Using lme4.* Journal of Statistical Software, 2015. **67**(1): p. 1-48.
47. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2016: Springer-Verlag New York.
48. Velden, L. F. J. v. d., Kasteleijn, A., and Kenens, R. J. *Cijfers uit de registratie van huisartsen*. 2016; Available from: https://nvl004.nivel.nl/nivel-2015/sites/default/files/cijfers-uit-de-registratie-van-huisartsen-peiling-januari-2016.pdf.
49. Asp, M., Lindqvist, D., Fernstrom, J., Ambrus, L., Tuninger, E., Reis, M., et al., *Recognition of personality disorder and anxiety disorder comorbidity in patients treated for depression in secondary psychiatric care.* PLoS One, 2020. **15**(1): p. e0227364.
50. Elstein, A.S., Rovner, D. R., Holzman, G. B., Ravitch, M. M., Rothert ML, Holmes MM. Psychological Approaches to Medical Decision-Making. Am Behav Sci. 1982;25(5):557-84.
51. Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., et al., *Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity*

Survey-Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry, 2011. **50**(1): p. 32-45.
52. Lykke, K., Christensen, P., and Reventlow, S., *"This is not normal ... "--signs that make the GP question the child's well-being.* Fam Pract, 2008. **25**(3): p. 146-53.
53. Wissow, L. S., Zafar, W., Fothergill, K., Ruble, A., and Slade, E., *Using vignettes to assess contributions to the work of addressing child mental health problems in primary care.* BMC Health Serv Res, 2016. **15**: p. 584.
54. Paulus, D. J., Wadsworth, L. P., and Hayes-Skelton, S. A., *Mental Health Literacy for Anxiety Disorders: How perceptions of symptom severity might relate to recognition of psychological distress.* J Public Ment Health, 2015. **14**(2): p. 94-106.
55. National Institute for Health and Care Excellence. *Social anxiety disorder: recognition, assessment and treatment Clinical guideline [CG159]*. 2013 [cited 2019 31 july]; Available from: https://www.nice.org.uk/guidance/cg159/chapter/1-Recommendations#interventions-for-children-and-young-people-with-social-anxiety-disorder-2.
56. Wang, P. S, Angermeyer, M., Borges, G., et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry. 2007;6(3):177-85.
57. Chavira, D. A., Garland, A., Yeh, M., McCabe, K., Hough, R. L. Child anxiety disorders in public systems of care: comorbidity and service utilization. J Behav Health Serv Res. 2009;36(4):492-504.
58. Wittchen H. U., Mühlig S, and Beesdo, K., *Mental disorders in primary care.* Dialogues Clin Neurosci, 2003. **5**(2): p. 115-128.

How words relate to our experiences is like catching a melody on an analogue radio: it changes whilst searching for the right frequency. This is the impossibility of psychiatric classifications and a blessing for psychotherapy

Chapter 3

Informative value of referral letters from general practice for child and adolescent mental healthcare

Published as: Aydin S., Crone M.R., Siebelink B.M., Numans M.E., Vermeiren R.R.J.M., Westenberg P.M. Informative value of referral letters from general practice for child and adolescent mental healthcare. European Child and Adolescent Psychiatry, 2021, doi: 10.1007/s00787-021-01859-7

Children's mental health is an acknowledged key area of concern for overall health, as is the adequate and appropriate allocation of resources available for mental healthcare.[1-5] In many countries the general practitioner (GP) is at the heart of this challenge with its key role in the recognition and referral of those in need of specialised care.[6] The bridge to specialised healthcare is formed mostly by referral letters (RLs). In fact, the RL represents the only substantive information transfer and the starting point for decision-making by the receiving services in a considerable number of cases. Evidently, RLs are central to a patient's transition and can potentially contribute to the diagnostic work-up and subsequent adequate provision of healthcare.[7-13] Notwithstanding, it is a widespread assumption that RLs hold very limited or no substantive value and are merely an administrative task.[5]

Several studies across various fields of medicine have analysed the information content of RLs, but little is known concerning the average RL for children and adolescents accessing mental health services.[14 15] RLs to psychiatric services could potentially guide institutions as regards the urgency of registration or even which subspecialty may be appropriate (e.g. emotional disorders). Studies concerning the recognition of psychosocial problems show variation depending on the type of disorder, generally with lower recognition rates for emotional disorders compared to externalizing or developmental disorders. Within emotional disorders, anxiety disorders are often less well recognised than depressive disorders.[16-21] These variations may well hold when considering the informative value of RLs. Nonetheless, as per our knowledge no study has provided a comprehensive overview of the full range of common reasons for referral, or has addressed the question of the informative value of RLs for child and adolescent mental healthcare.

Objectives

To increase understanding of the informative value of RLs, in this study we compared information found in children's and adolescents' RLs to the later diagnostic classifications made in specialised mental healthcare. First, we asked if RLs demanding urgency were associated with higher levels of functional impairment. Next, we inspected predictive values for the full breadth of diagnostic categories covering higher-order level emotional and developmental disorders, and specifically for the common disorder groups: anxiety disorders, depressive disorders, post-traumatic stress disorders (PTSD), eating disorders,

autism spectrum disorders (ASD), attention-deficit (hyperactivity) disorders (ADHD), and behavioural disorders. In an explorative approach, we also inspected cross-relations between these categories and indications made in RLs. Thirdly, we aimed to relate the predictive value of RLs to age, gender, levels of functional impairment, and length of psychiatric treatment history. In addition, finally, to gain broader insight into the reasons for referral, we examined the informative value of more general reasons for referral mentioned in RLs,[5] such as physical ailments or educational and parental difficulties.

Methods

Study design and sample

We conducted a retrospective chart review of the electronic medical records (EMRs) at Curium-LUMC, a clinic for mental health treatment affiliated to Leiden University Medical Centre (LUMC). Curium-LUMC receives referrals from a quarter of all municipalities in the Netherlands, and offers outpatient, day patient, and inpatient treatment for minors aged three to 18 years. Outcomes were based on institutional protocols designed to classify DSM-5 diagnoses following the gold standard assessment procedure in child and adolescent psychiatry. The diagnostic work-up facilitates combining structured information from various informants (youth themselves, caregivers and/or teachers), as well as the clinicians' judgement after interview and observation.[16 22-25]

For the purposes of feasibility we set a two-year limit and included cases that registered between January, 2015 and December, 2017. To improve the reliability of the reference standard,[22] we only included data on cases classified using a comprehensive assessment including interview with a clinician, observation, and a structured multi-informant assessment. The latter was provided by the Development and Well-Being Assessment (DAWBA[26]) which is part of the institution's intake protocol. Yearly about 30% of the total caseload of the institution follows a different intake and assessment procedure. Those are patients that register for inpatient care or in a critical situation, and are not included in this study. Within the set time period 1268 patients and/or caregivers had completed the comprehensive intake procedure. Three cases were excluded because of an illegible RL, and six owing to the absence of an RL in the EMR. This resulted in a sample of 1259 extracted

RLs, of which the 723 (57.4%) from general practice could be included in the study. As this is the first study to investigate RLs for a wide a range of reasons of referral, we decided a priori to analyze only RLs from the most frequent referrer. In the Netherlands, as in many other European countries, this is the general practitioner.[27 28] An overview of referrers can be found in the supplementary material.

Data and measures

We coded and then compared which of the various mental health disorders were indicated in RLs and whether they corresponded to the final clinical classifications including comorbidities. Coding followed the DSM-5 chapter structure, e.g. post-traumatic stress disorder (PTSD) and obsessive compulsive disorders were separated from anxiety disorders, whereas phobias were included.[29] For common disorders in psychiatric services, such as ASD and ADHD, we present values for individual classifications rather than a whole chapter (e.g. the neurodevelopmental disorders) combined. Regarding the higher-order disorder groups, we present metrics for both internalizing disorders and developmental disorders, rather than the common dichotomisation of internalizing versus externalizing problems. This approach was based on the high prevalence of ASD and ADHD and the very low prevalence of conduct disorders in the study sample, as well as the fact that ADHD is conceptually related to both externalizing and neurodevelopmental disorders. All data were handled in compliance with regulatory requirements and the code of conduct for research using health data. Based on the retrospective nature of the study, the Medical Ethics Committee of the LUMC provided an exemption for written informed consent (G18.080).

Extraction of referral letter data

RLs were extracted from individual EMRs. Two graduate students transcribed the clinical texts from RLs into a digital data extraction form. To achieve consistency in data extraction, the students and author SA independently extracted an initial set of 30 RLs. After achieving consensus, for each 100th transcribed RL, five selections were examined and discussed to prevent variation developing over time.

An EMR login code that only gave access to filed correspondence was created to ensure blinding for diagnoses recorded elsewhere in a patient's EMR. The data extraction form included the following: a transcription of the main reason for referral, other contextual

information relayed with the RL, whether an ICPC code (International Classification of Primary Care code[30-32]) was included, which ICPC codes were present (together with the year and textual description of the code), the referring healthcare institution, and whether the data extraction should be discussed. The form also captured an approximate summary of the patient's psychiatric treatment history (no other previous mental health treatments, short-term treatment of up to a year including primary healthcare, or a relatively long treatment history). This is an estimation for whether patients were diagnosed earlier, as an approximation for whether the referrer might have used a formal diagnosis in the RL. To better estimate treatment history, RLs were not our only source to estimate treatment history. Where necessary, students were asked to search for additional information in other correspondence present in the EMR. If RLs were sent and filed with attached reports from earlier treatments, these attachments were not extracted.

Coding of the referral letters

Regarding indications of urgency, we distinguished three groups of RLs: those in which priority was explicitly requested (including the words "urgent" or "emergency"), in which a serious need was indicated explicitly ("ASAP", "major" or "serious" [problems]), and those without any such statement.

With respect to reasons for referral, we labelled the transcribed RLs using codes from the ICPC-01 classification system currently used in general practice in the Netherlands.[32] The ICPC system provides codes for reported symptoms and contextual problems, in addition to codes for physician's (tentative) diagnoses. To aid the coding process, an extensive manual including a glossary of probable reasons for referral and corresponding ICPC codes was compiled and discussed with a GP who has extensive experience with mental healthcare and research using the ICPC coding system. Besides codes from chapter P (for psychological problems), the manual also included codes from chapter Z (for psychosocial problems), as well as some general codes for physical ailments (e.g. A04-Weakness/tiredness, N01-Headache, D01-Abdominal pain/cramps). This manual was refined over the course of five meetings based on the discussion of 20 RLs that were individually coded by SA, PMW, BMS and MRC. During this iterative process some extra codes that are not covered by the ICPC system were added due to their high prevalence in RLs (e.g. self-harm, being bullied, school attendance problems). Based on the length and information load of the RLs, we labelled

each RL with up to five ICPC codes and coded in order of decreasing importance (from the main reason for referral to more peripheral symptoms and problems mentioned in RLs).

To evaluate consistency in coding, a random selection of 150 RLs was made and the weighted average agreement was computed between the first author who coded all RLs and the three second coders who each coded a set of 50 letters. Weighted average agreement between coder 1 and the three 2nd coders was 82% (lowest 79%, highest 83%), suggesting generally reliable coding. Chance-corrected agreement on the frequency of specific reasons for referral was also high, for example, excellent agreement was reached on whether anxiety was coded or not, with an overall κ=.81 (95% CI κ=.73 to .86, supplementary material).

The reference standard and clinical context

The diagnostic process starts immediately upon registration of a patient and receipt of an RL. RLs are scanned and filed in EMRs. A designated employee then conducts a short telephone interview with parents or caregivers, and provides them with an admission package that includes a login code for the online multi-informant DAWBA tool.[26] Parents, teachers and youth over the age of 11 years are invited to respond, except in case of an inpatient referral. In the online DAWBA environment informants' responses to closed-ended questions generate scale scores which, together with their responses to open-ended questions, can be remotely reviewed by a clinical rater. A report on this review is then copied to the EMR to facilitate reliability during a face-to-face intake interview that is often led by a senior psychologist. Therein the professional is free in how to incorporate the DAWBA data or to supplement with additional assessment methods. The intake assessment is followed by a psychiatric assessment, after which a classification and a CGAS score[33] is entered in the EMR. CGAS (Children's Global Assessment Scale) scores are an estimation of the level of functional impairment and range between zero and 100, with lower scores indicating more impairment. Depending on complexity and needs, variations to this protocol are common in daily clinical care. The administration of a classification can be postponed when further assessment is needed or the endorsement of a DAWBA is passed when a case enters with emergence. In addition, classifications can be adjusted following insights obtained during treatment. We found, in line with the available literature,[34] that such adjustments in classifications were made in about a tenth of cases, over half of which considering minor changes (for example a deletion of a V-code: other conditions). In these instances the last

entry was kept as reference. Contrary to the reasons for referral, outcome measures could be extracted groupwise and concurrently from the EMR system.[35]

Secondary measures

To better understand sample characteristics, we obtained data on a patient's age and gender, their neighbourhood socioeconomic status (nSES) score and the type of care (outpatient, daycare or inpatient). Age and gender were extracted from the DAWBA data, whereas nSES and type of care were derived from the EMR. nSES is a normalised and standardised score based on the income, education and occupation of inhabitants for each postal code area in the Netherlands.[36]

Statistical analysis

First, the demographics of sample and excluded cases were compared in an ANOVA, with nSES and CGAS scores as dependent variables, and sample and type of care as main and interaction effect. This was followed by an analysis of descriptive statistics to gain insight into the content of the average RL.

Using ANOVA, we compared impairment levels (as approximated by CGAS scores) between the three types of referral letters (priority requested, serious problems indicated or normal referral).

The reasons for referral and the final clinical diagnoses were then cross-tabulated for the various classifications. We noted the number of RLs that accurately predicted outcome as a ratio of the total frequency of a psychiatric outcome. This represents the sensitivity of a test and when plotted against the specificity of an instrument the area under the receiver operating curve (AUROC) value is obtained. AUROC values are considered to be insensitive to sample prevalence and indicate the strength of discriminative ability, being graded as fair (.50-.70), fair to moderate (.70-.80), good (.80-.90) and excellent (.90-1.00).[37] Plots were created for those with and without multiple classifications to obtain values representative for the daily clinical cohort (including those with comorbidity) and to provide insight into the potential effects of comorbidity on the metrics. AUROCs were plotted using pROC [38] and 95% CIs of the diagnostic metrics were computed in EpiR.[39]

We computed positive predictive values (PPVs), negative predictive values (NPV) and likelihood ratios of positive and negative RLs (LR^+ and LR^-) to quantify the likelihood of classifications being made. PPVs are computed as the number of RLs classified with their reason for referral as a ratio of the total frequency of that reason for referral. Similarly, NPVs represent the percentage of those who were not referred for a particular problem and were not classified as such, expressed as a ratio of the number of RLs without that particular reason for referral. As a percentage, predictive values are very intuitive. Nonetheless, they depend on the prevalence of the outcome and are, therefore, not easily generalisable. LR^+ and LR^- values, on the other hand, are less susceptible to sample distribution[40] as they represent the actual likelihood of a particular outcome for those positive (LR^+) or negative (LR^-) on a test. For LR^+, values >2 indicate a slight increase in post-test probability of about 15% in the likelihood of a positive outcome, and >10 indicates a large increase of approximately 45%. LR^- values <0.5 point towards a slight decrease of 15%, and <0.1 a decrease of 45%, interpreted as a strong indicator of absence. Tests with an LR^+>20 or LR^- <0.05 are deemed diagnostic in clinical practice.[41]

Lastly, in a logistic regression analysis, we analysed whether the predictive value of RLs differed depending on age, gender, CGAS score or treatment history.

Results

Sample characteristics

Demographics of the sample are depicted in Table 1. On average, girls (43.6%) were 13 years old and boys were 10 years old. Around a third of the sample had no history of previous mental health treatment. The majority had one (47.4%) or two (27.9%) DSM-5 classifications (Table 2). The study sample had an average nSES score (M=0.47, SD=0.77) and moderate to serious dysfunctioning as approximated by CGAS scores (M=51.01, SD=7.61, n=689). The included study sample was similar to the not included caseload of the institution regarding nSES score ($F_{(2, 2032)}$=0.58, p=.56, $\eta^2_{partial}$<0.000), but showed a higher CGAS score ($F_{(2, 1804)}$=14.53, p<.000, $\eta^2_{partial}$=0.016).

Table 1. Sample Characteristics N=723		
		n (%)
Age	5-7	131 (18.1)
	8-10	189 (26.1)
	11-13	153 (21.2)
	14-15	147 (20.3)
	16-18	103 (14.2)
Gender	Male	408 (56.4)
	Female	315 (43.6)
Mental health treatment history	None	202 (27.9)
	Short/Limited	228 (31.5)
	Long	284 (39.3)
	Unknown	9 (1.2)
Medical conditions	None classified	577 (79.8)
	Singular	47 (6,5)
	Complex	18 (2.5)

The Mental health treatment history variable is an estimation based on the information available in the medical record, see below section "data extraction".

Content of referral letters

The average extracted reason for referral consisted of 59 words (SD=41, range 2 to 246) and depicted problems regarding psychological functioning as well as contextual information. Priority was requested in 36 RLs (5.0%), and a serious need was explicitly indicated in another 50 RLs (6.9%). A few RLs stated only a general request for psychiatric evaluation or treatment without any other additional information (1.2%, Table 3). Most RLs contained one (25.0%), two (32.2%) or three (24.8%) symptoms or tentative diagnoses. The majority of reasons for referral concerned psychological problems. Next to the textual description of the problems which we coded ourselves, 45.8% (n=331) of RLs contained an ICPC code registered by the referrer, most of which were from the P chapter (supplementary table).

Informative value of referral letters

The average CGAS score of youth with an RL not explicitly indicating urgency or a severe status (M=51.35, SD=7.12) was only slightly higher when compared to those with an RL that explicitly mentioned urgency (M=47.27, SD=8.12) or an RL stating the seriousness of the condition (M=48.83, SD=8.01; $F(2,686)=7.71$, $p<.001$). Whereas 41.6% of RLs did not mention any of the later clinically-established classifications, the majority of RLs (50.8%) mentioned one, two (7.3%) or even three (0.3%) provisional diagnoses that were in line with the outcome.

Table 2. Prevalence of the various clinical classifications

	n (%)
Clinical classifications	
Neurodevelopmental disorders	425 (58.8)
Intellectual disability	21 (2.90)
Communication disorder	18 (2.49)
Motor disorders	14 (1.94)
Autism spectrum disorder	214 (29.60)
Attention-deficit hyperactivity disorder	243 (33.61)
Specific learning disorder	38 (5.26)
Other Neurodevelopmental Disorders	25 (3.46)
Schizophrenia spectrum and other psychotic disorders	2 (0.28)
Depressive disorders	92 (12.72)
Anxiety disorders	105 (14.5)
Separation anxiety disorder	8 (1.11)
Specific phobia	6 (0.83)
Social anxiety disorder	16 (2.21)
Panic disorder	8 (1.11)
Agoraphobia	1 (0.14)
Generalized anxiety disorder	47 (6.50)
Anxiety disorder not otherwise specified	28 (3.87)
Obsessive-compulsive and related disorders	8 (1.11)
Trauma and stressor-related disorders	39 (5.4)
Post-traumatic stress disorder	21 (2.90)
Adjustment disorder	4 (0.55)
Reactive attachment disorder	15 (2.10)
Disinhibited social engagement disorder	1 (0.14)
Disorder of infancy, childhood, or adolescence NOS	24 (3.32)
Somatic symptom and related disorders	17 (2.35)
Feeding and eating disorders	27 (3.73)
Elimination disorders	8 (1.11)
Gender dysphoria	6 (0.83)
Disruptive, impulse-control, and conduct disorders	43 (5.95)
Oppositional defiant disorder	15 (2.10)
Intermittent explosive disorder	2 (0.28)
Conduct disorder	2 (0.28)
Other specified- or Unspecified DIC and CD	24 (3.32)
Substance-related and addictive disorders	2 (0.28)
Personality disorders	34 (4.70)
Number of clinical classifications	
0	91 (12.6)
1	343 (47.4)
2	202 (27.9)
3	71 (9.8)
4	15 (2.1)
5	1 (0.1)

The distribution of the clinical classifications is depicted as per the DSM-5 chapters, excluding the classified V-codes. NOS=not otherwise specified. DIC, and CD= Disruptive, impulse-control and conduct disorders. There were no cases classified with Bipolar and related disorders, Mutism, Body dysmorphic disorder, Dissociative disorders, Acute stress disorder or Sleep-wake disorders. Cases could be classified with more than one diagnosis.

When we considered the informative value in relation to higher-order internalizing and developmental disorders, we found that just over half of the RLs suggesting anxiety, depression and/or trauma accurately predicted subsequent classifications (Table 4). Indications of autism-related, attention-hyperactivity and/or behavioural problems were predictive in over two-thirds of cases. How well the indications in RLs correlated with later higher-order classifications did not differ between girls and boys, different age groups or based on whether there was a previous mental health treatment history (supplementary material).

Table 3. Frequencies of problem areas in referral letters

	First	Second	Third	Fourth	Fifth
Psychological	685 (94.7)	402 (55.6)	196 (27.1)	82 (11.3)	29 (4.0)
Social	26 (3.6)	113 (15.6)	95 (13.1)	30 (4.1)	12 (1.7)
Physical	3 (0.4)	18 (2.5)	10 (13.8)	9 (1.2)	3 (0.4)
No code labelled at this spot	9 (1.2)	190 (26.3)	422 (58.4)	602 (83.3)	679 (93.9)

Depicted are the frequencies (%) of the ICPC codes, per domain, per coding spot (N=723). Psychological=codes from the P and T chapters (eating disorders and symptoms) combined. Social=Z chapter. Physical=all other labels given. On some of the RLs referrers had written ICPC codes themselves - these can be found in the supplementary material.

Differences were found with regard to the percentage of specific classifications indicated in RLs (Figure 1). Youth with anxiety disorders were infrequently referred as such (sensitivity=41.9%, 95% CI 32.4-51.4), with somewhat higher values for PTSD (52.4%, 95% CI 33.3-71.4) and ASD (54.7%, 95% CI 48.1-61.2). Confidence intervals overlapped for many disorder groups. A notable exception was eating disorders, which were referred with greatest accuracy (sensitivity=92.9%, specificity=98.4%). To explore whether the metrics are a result of comorbidity, AUROC values were inspected after removal of those with co-occurring classifications (lower Figure 1). In absolute terms, sensitivity increased for depressive, eating, and attention-deficit hyperactivity disorders while at the same time sample size decreased considerably, limiting the value of these findings.

We then investigated the predictive value of various reasons for referral (Table 5). The highest PPV was found for eating problems, where 67.6% of RLs were concordant with an ensuing eating disorder classification. PPVs varied, with behavioural problems showing the lowest PPV value, followed by trauma, anxiety, depression, autism and attention-hyperactivity problems. The value of the RL in predicting specific disorder groups did not

differ between girls and boys, different age groups or depending on treatment history (supplementary material), with the exception of a small age effect for the indication ADHD. Information in the RLs predicted the diagnosis of ADHD better with increasing age (OR=1.14, 95% CI 1.03-1.27, *p*=.026).

Broader investigation of the reasons for referral revealed that a quarter of children referred for mood problems were later classified with an anxiety disorder (24.3%, supplementary material). The reverse association, i.e. referred for anxiety then classified with depression, was not found. A similar pattern was seen for those eventually diagnosed with behavioural disorders, as they were equally likely to be referred for suggested behavioural problems (14.3%) or trauma (13.9%). Although high raw values were found for some other disorder groups, the frequencies were no more than expected by chance.

Finally, we investigated the informative value of other general problems frequently indicated in RLs (Table 6). Those referred with academic problems were often classified with ADHD (46.4%), and those referred for school attendance problems with an anxiety disorder (42.9%). Half of children referred with possible learning disorders were diagnosed with ADHD. Referral with physical symptoms was significantly associated with a subsequent diagnosis of a depressive disorder (34.4%), and relatively high percentages were also found for anxiety, ASD and ADHD (25%, 25% and 12.5%, respectively). Similarly, around 40% of indications for suicidal ideation or self-harm were subsequently related to a diagnosis of a depressive disorder. Over 80% of children with an indication of bullying or related problems in the social environment were classified with an ASD or ADHD. Other infrequently mentioned problems can be found in the supplementary material.

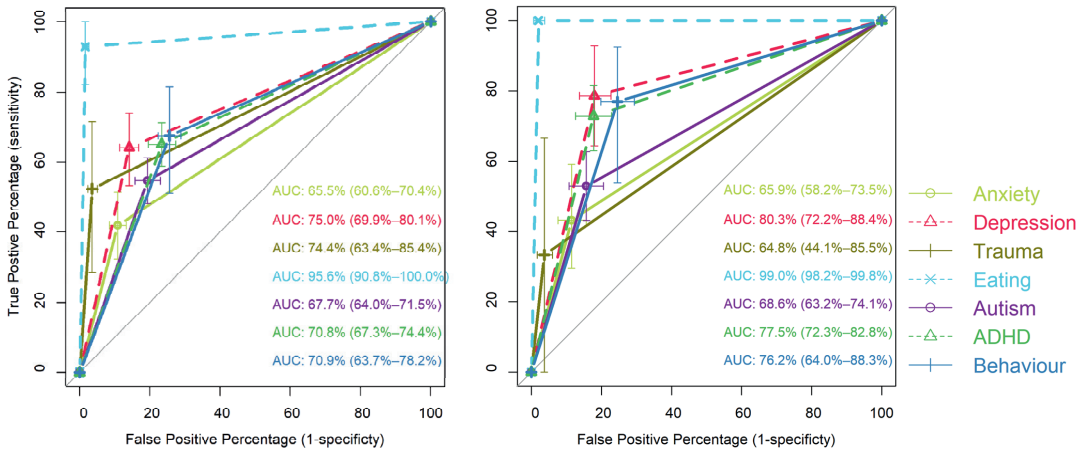


Figure 1. AUROC values of indications made in RLs by disorder group and sample

Plotted are the 95% confidence intervals of the sensitivities and specificities, depicted together with the 95% confidence intervals of the AUROC values. The figure on the left presents values of the complete sample (N=723), thus including those with multiple classifications. The figure on the right depicts values in a sample created by excluding cases with co-occurring diagnoses. Note that here the sample size decreased substantially (n=306) as did the number of cases (anxiety disorders n=44, depressive disorder n=28, PTSD n=6, eating disorders n=13, ASD n=102, ADHD n=92, behavioural disorders n=13).

Table 4. Informative value of the referral letter for higher-order categories

	Cases/ Positive RLs	PPV (95% CI)	Non-cases/ Negative RLs	NPV (95% CI)	Sensitivity	Specificity	LR^+ (95% CI)	LR^- (95% CI)
Anxiety/Depression n=179	121/224	54.0 (49.0-59.0)	58/499	88.4 (86.0-90.4)	67.6 (60.2-74.4)	81.1 (77.5-84.3)	3.57 (2.92-4.37)	0.40 (0.32-0.50)
Anxiety/Depression/ PTSD n=195	137/249	55.0 (50.3-59.6)	58/474	87.8 (85.2-89.9)	70.3 (63.3-76.6)	78.8 (75.1-82.2)	3.31 (2.74-4.00)	0.38 (0.30-0.47)
ASD/ADHD n=391	297/419	70.9 (67.7-73.9)	94/304	69.1 (64.8-73.1)	76.0 (71.4-80.1)	63.3 (57.8-68.5)	2.07 (1.78-2.41)	0.38 (0.31-0.46)
ASD/ADHD/Behavioural disorders n=412	355/505	70.3 (67.7-72.8)	57/218	73.9 (68.5-78.6)	86.2 (82.5-89.4)	51.8 (46.1-57.4)	1.79 (1.58-2.02)	0.27 (0.21-0.35)
All neurodevelopmental/ Behavioural disorders n=444	383/519	73.8 (71.3-76.2)	61/204	70.1 (64.4-75.2)	86.3 (82.7-89.3)	51.3 (45.2-57.3)	1.77 (1.56-2.01)	0.27 (0.21-0.35)

Depicted are the accuracy metrics in numbers for the combined higher-order disorder groups, e.g. Anxiety/Depression depicts the accuracy metrics between RLs containing anxiety and/or depression and the final clinical classification anxiety and/or depression. PTSD=post-traumatic stress disorder, ASD=autism spectrum disorders, ADHD=attention-deficit hyperactivity disorder. PPV=positive predictive value: the number of children with an issued reason for referral who were also classified with that reason for referral as a ratio of the total frequency of that reason for referral. NPV=negative predictive value: the number of RLs without any indication of the disorder and no final classification, as a ratio of the total number of RLs. Sensitivity=number of children with an issued reason for referral who were also classified with that reason for referral as a ratio of the total prevalence of that diagnostic classification. Specificity=number of RLs without an indication that were also not classified with it as a ratio of the total sample without that diagnostic classification.

Table 5. Informative value of RLs for the seven most widespread mental health disorders

Classification	Cases/Positive RLs	PPV (95% CI)	Non-cases/ Negative RLs	NPV (95% CI)	LR^+ (95% CI)	LR^- (95% CI)
Anxiety n=105	44/111	39.6 (32.3-47.5)	551/612	90.0 (88.5-91.4)	3.87 (2.81-5.32)	0.65 (0.55-0.77)
Depressive n=92	59/148	39.9 (34.1-45.9)	542/575	94.3 (92.6-95.6)	4.55 (3.56-5.81)	0.42 (0.32-0.55)
PTSD n=21	11/36	30.6 (20.1-43.5)	677/687	98.5 (97.7-00.1)	14.71 (8.40-25.77)	0.49 (0.32-0.77)
Eating n=27	25/37	67.6 (54.1-78.7)	684/686	99.7 (98.9-99.9)	53.70 (30.34-95.05)	0.08 (0.02-0.29)
ASD n=214	117/215	54.4 (49.0-59.7)	411/508	80.9 (78.4-83.2)	2.84 (2.29-3.52)	0.56 (0.48-0.65)
ADHD n=243	158/270	58.5 (53.9-63.0)	368/453	81.2 (78.4-83.8)	2.79 (2.31-3.36)	0.46 (0.38-0.55)
BD n=43	29/203	14.3 (11.6-17.5)	506/520	97.3 (95.9-98.2)	2.64 (2.06-3.36)	0.44 (0.28-0.67)

RL=referral letter. PPV=positive predictive value in percentages. NPV=negative predictive value in percentages. LR^+=positive likelihood ratio, LR^-=negative likelihood ratio. PTSD=post-traumatic stress disorder, ASD=autism spectrum disorders, ADHD=attention-deficit (hyperactivity) disorders. BD=behavioural disorders: conduct disorder and oppositional defiant disorder. Frequencies and cross-tabulations are depicted in the supplementary material, as are metrics for the less prevalent disorder groups, the various neurodevelopmental and specific anxiety disorders.

Table 6. Frequency of general reasons for referral per disorder group

	Anxiety disorders	Depressive disorder	PTSD	Eating disorders	ASD	ADHD	Behavioural disorders
Academic problems n=84	11 (13.1%)	5 (6.0%)	1 (1.2%)	2 (2.4%)	32 (38.1%)	**39 (46.4%)**	7 (8.3%)
st.adj.res.	-0.4	-2.0	-1.0	-0.7	1.8	**2.6**	1.0
School attendance n=28	**12 (42.9%)**	**8 (28.6%)**	0	0	7 (25.0%)	4 (14.3%)	0
st.adj.res.	**4.3**	**2.6**	-0.9	-1.1	-0.5	-2.2	-1.4
Learning disorders n=30	0	0	0	0	8 (26.7%)	**16 (53.3%)**	1 (3.3%)
st.adj.res.	-2.3	-2.1	-1.0	-1.1	-0.4	**2.3**	-0.6
Somatic symptoms n=32	8 (25.0%)	**11 (34.4%)**	0	0	8 (25.0%)	4 (12.5%)	0
st.adj.res.	1.7	**3.8**	-1.0	-1.1	-0.6	-2.6	-1.5
Problems Sleeping n=18	4 (22.2%)	4 (22.2%)	2 (11.1%)	1 (5.6%)	2 (11.1%)	7 (38.9%)	0
st.adj.res.	0.9	1.2	2.1	0.4	-1.7	0.5	-1.1
Suicidal ideation n=53	10 (18.9%)	**23 (43.4%)**	1 (1.9%)	0	14 (26.4%)	8 (15.1%)	2 (3.8%)
st.adj.res.	0.9	**7.0**	-0.5	-1.5	-0.5	-3.0	-0.7
Self harm n=28	7 (25.0%)	**12 (42.9%)**	**3 (10.7%)**	**3 (10.7%)**	8 (28.6%)	6 (21.4%)	1 (3.6%)
st.adj.res.	1.6	**4.9**	**2.5**	**2.0**	-0.1	-1.4	-0.5
Problems with parents n=87	15 (17.2%)	14 (16.1%)	**8 (9.2%)**	3 (3.4%)	16 (18.4%)	25 (28.7%)	**11 (12.6%)**
st.adj.res.	0.8	1.0	**3.7**	-0.2	-2.4	-1.3	**2.8**
Bullied/Social relatedness n=51	5 (9.8%)	6 (11.8%)	1 (2.0%)	1 (2.0%)	20 (39.2%)	22 (43.1%)	1 (2.0%)
st.adj.res.	-1.0	-0.2	-4.0	-0.7	1.6	1.5	-1.2

Frequency of the general reasons for referral per disorder group (as a percentage of the total frequency of that reason for referral). Standardized adjusted residuals depict the discrepancy between observed and expected values and suggest statistical significance at the level of $p<0.05$ if $>|1.96|$ (depicted in bold). A case could be referred for multiple reasons, as well as be classified with multiple diagnoses. Academic problems (ICPC code Z07) were coded when a decline in academic functioning was indicated. Learning disorders (ICPC code P24) were coded when more specific indications were made, such as indications of specific learning disorders, dyslexia, language and speech disorders or developmental coordination disorder. Social relatedness was coded when Loneliness (ICPC code Z04.03) and Relationship problem with friends (Z24) were indicated.

Discussion

The adequate provision of mental healthcare is an ongoing topic and any additional role for RLs beyond an administrative process is a subject of debate within the field. Nonetheless, over half of children in this clinical sample were subsequently classified with at least one condition mentioned in their RL. For higher-order combined categories we found PPVs of over 50% for internalizing disorders and over 70% for developmental disorders. Scrutinising PPVs for each of the common diagnostic categories, we found that over two thirds of RLs that suggested eating disorders were in concordance with the outcome. Half of RLs that suggested autism or ADHD as the underlying problem concurred with the later classification. Around two fifths of RLs that mentioned anxiety or depression were later classified as such, and a third of RLs indicating trauma resulted in a classification of PTSD. The least informative reason for referral was behavioural problems, with only one in seven RLs with this indication resulting in the classification of behavioural disorders. Considering sensitivities the highest value was found for eating disorders and the lowest for anxiety disorders. We found no support for an association between predictive value of RLs and estimated length of treatment history, gender, age or level of functional impairment, except for a weak association between higher age and ADHD. Exploration of the reasons for referral more broadly revealed that some general problems such as learning or family problems were often indicated and associated with different outcomes.

Our findings are in line with the broader medical literature, research in various medical specialties suggests that RLs yield some useful information but improvements are necessary. The predictive values we found are similar to the two studies on RLs concerning autism spectrum disorders[14] and non-obsessive-compulsive anxiety disorders.[15] To the best of our knowledge no other studies have been published on the value of RLs in predicting the full range of common diagnostic categories. The differences between disorder groups were, however, mostly parallel to those from general literature on recognition of psychosocial problems. For instance, whilst RLs are less concordant considering those with anxiety disorders, they were better in including those with depressive disorders.[21] Considering that behavioural problems are mentioned at a five-fold higher frequency in RLs compared to their prevalence at diagnosis, one might legitimately ask whether referrers label difficult

behaviour that may be common in various disorder groups as a behavioural disorder. It is a question for future studies to investigate to what extent it is that referrers pick up or zoom in on these rather externalizing manifestations, or in how far it is a terminological issue and their way to state problems in behaviour.

We also coded and analysed some indications made beyond diagnostic labels. Here we found that problems at school and within the family environment were frequently mentioned. This relates to what is described earlier as the strength and weakness of GPs; ability to adopt a contextual and systemic approach on the one hand[42 43] and on the other hand focusing less on the internal experience of youth which might impede noticing and recording covert problems such as anxiety disorders.[18] In the context of the many somatic manifestations of psychosocial problems in youth, a surprising finding was the low prevalence of physical symptoms in RLs. This has been reported earlier in literature on adult mental health[44] A possible explanation is that once the decision to refer to psychiatry is made, referrers might perceive somatic symptoms as irrelevant[5] This may also relate to our observation during coding that many RLs seemed to be written as a concise justification of referral,[5] rather than a description of the circumstances with the goal of maximum information transfer. Nonetheless, we did not structurally investigate this interesting issue.

About one in ten cases in this referred sample were not classified with a DSM-label and sent back to the referrer or another institution, often primary care. This "wrong referral rate" is up to half the amount suggested in other studies[45] which we relate to the protocol of the institution including pre-intakes by phone. As from the point of view of families and referrers a "back referral" is impactful, we prefer to interpret each registration as a request for help following exhausted resources in general practice.[5] Inspecting RLs within such a perspective could contribute to a mutual understanding of the language and decision-making in both ends of healthcare.

Strengths and limitations

In a relatively large registration cohort we related information conveyed in RLs to the full breadth of diagnostic outcomes. A strength inherent to the study is that the results present values from everyday practice. A possible concern is the extent to which this single institutional sample reflects the needs of children that register with specialised mental

healthcare, as our findings might not be generalisable to centres that operate on another institutional level. Notwithstanding, one could reasonably argue that when investigating the predictive value of RLs it is less the centre's diversity but the referrers that matters. Since the institution receives referrals from a broad range of referring practices, our results might usefully inform specialised mental health institutions. That said, the current results should be viewed as a first thorough endeavour to the issue of RLs. We choose a priori to examine RLs from general practice only as they are the most frequent referrer and usually the first families turn to for help.[46] Future studies might investigate RLs from other referrers such as medical specialists and paediatricians to shed light on what differences between referrers exist.

A major strength of our study is the use of the best estimate approach in psychiatry as outcome measure.[47] We included data on patients that were diagnosed using structured assessment as well as clinician judgement following face-to-face interview. The criterion of available structured assessment might have led to a selected sample as those registering in a critical situation are not asked to fill in the DAWBA before the intake interview. Exclusion of these tertiary care patients might either have inflated or deflated agreement. The excluded cases might have had a more complex presentation and thus less concordance between reason of referral and outcome. A part might even not have had a RL as it is not planned care and they arrive through a different route. On the other hand, these youths might have had more marked problems and therewith problems that were better recognizable for the referrer. Nonetheless, a focus on outpatient referrals is preferable in regards generalisability of this first investigation as referrals to specialised healthcare are generally more common and we aimed to gain insight in the value of RLs to child and adolescent mental healthcare.

A downside to extracting clinical data is that clinicians who made the diagnoses had access to RLs, which could potentially inflate agreement between predictor and outcome ("incorporation bias"), although there is insufficient empirical evidence for such effects.[48 49] In fact, existing literature suggests that most mental health professionals tend to view RLs as incomplete and do not automatically accept information contained in RLs.[50] Moreover, we found PPVs similar to those found in the few available studies. Last but not least, the clinical diagnostic process of the clinic is extensive and elaborate, embracing interviews and

questionnaires endorsed by multiple informants and professionals. It is likely that in the presence of this information clinicians will not rely on RLs. That said, replication of current findings in a study setting that ensures complete independence between RLs and diagnoses would lend stronger support to the quantified values.

Another strength of the study was the rigorous coding of information contained in the RLs. We reached good reliability despite multiple labels given to most RLs. In line with the clinical nature of the research question, we aimed to keep the sample as natural as possible, meaning that youth with co-occurring disorders and multiple reasons for referral were included. However, we did not differentiate the main reason for referral or the tentative diagnosis from secondary problems, context, or other symptoms and problems mentioned in RLs. This was impossible given the retrospective design of the study and the differences between RLs in terms of layout and writing style. Basing our coding on first-mentioned issues would have been inadequate since some GPs first provide extensive background to the referral, others only outline the current situation without providing a clear diagnostic interpretation, whereas many others prefer very short and concise description. Differentiating symptoms and diagnoses presented in RLs might be a topic for future studies as a good RL is proposed to contain an explicit indication of a preliminary diagnosis. [43]

A limitation of the study may be the analyses of how the informative value of RLs varies with gender, age, treatment history and level of disfunctioning. Including these four interaction terms in addition to their main effect, together with the imbalance between cases and non-cases, resulted in reduced power. Studies with a larger sample size might differentiate positive and negative agreement between RLs and diagnosis as well, as this might differ depending on these factors. Similarly, our results may have underestimated the informative value of RLs related to the urgency of referral. We differentiated three subsamples of RLs based on the presence of explicit statements of urgency (urgent, serious need, or no explicit statement). Yet we noted descriptions of urgency using more general phrases in RLs that were not included in the two subsamples with explicit statements.

Implications

The study findings suggest that most RLs do contain valuable information. Nevertheless, an important question is what value is sufficient. On the one hand, none of the diagnostic

likelihood ratios we found reached the necessary levels for clinically meaningful use, with the exception of indications for eating disorders. On the other hand, as we might cautiously infer from the moderate AUROCs found in this study, RLs may be almost as valuable as some structured assessment instruments in discriminating psychiatric classifications.[16 51 52] However, this assertion should be placed in perspective of the numbers and the context of referrals. As the study considered prevalent disorders and a selected sample, even the high values we found imply a major cost of false omissions when absolute numbers are considered.[21] From the perspective of referrers,[53] attributing subsequently divergent diagnoses as inaccurate would lack the necessary nuance. Specialised healthcare populations are epitomised by inherently complex problems and the need for elaborate diagnosis is a valid reason for referral. Putting aside expectations of high accuracy, our results support use of RLs as a node of information in the diagnostic work-up. Beyond their effect on diagnosis and allocation, incorporating RLs in the assessment process may have a welcome side-effect as it might potentially ameliorate families perception of fragmented care.[45]

In countries where the GP has a gatekeeper role, content guidelines and formats are defined and embedded in health records to help improve RLs. Accordingly, the RL is an integral component of a GPs' training and continuing medical education.[43 54-56] The sensitivity and specificity values found in this study might help inform curricula.

Another finding with clinical implications concerns the ICPC codes included in RLs. When GPs register a code they also write out a short description, often in just two or three words. We observed that these descriptions often suggested a disorder or symptom that diverged from the ICPC code they had registered and copied to the RL. This suggests that the ICPC codes communicated in a RL have limited significance in specialised mental health services, and in research using automatised analysis in medical records. Finally, guidelines on coding could be improved as the study revealed some limits of the P and Z chapters of the ICPC coding system. There are some inconsistencies between symptom and disorder codes, as some codes for important problems are lacking, whereas multiple codes exist for some less prevalent symptoms. In recent years a sub-code for autism spectrum disorders has been added, for example, and most GPs in our sample seemed to use it as intended.

Conclusion

In this study, we investigated the symptoms and provisional diagnoses described in RLs to child and adolescent mental healthcare. We conclude that, contrary to widespread clinical anecdotes, RLs appear to hold informative value and might add to the clinical process in child and adolescent psychiatry. Future studies of RLs may shed light on other important dimensions of utility and quality. Among these are the clarity and completeness of the information conveyed, the investigation and treatment requested, and how these factors relate to the diagnostic work-up and treatment families eventually receive.[13 57 58] Another essential question relates to the factors explaining individual differences between RLs. Quantification of the complete process between referral and assessment is necessary to stimulate a mutual understanding of strengths and weaknesses – at both the referring and receiving end in healthcare – and thus help inform the day-to-day diagnostic process.

Reference list Study 2

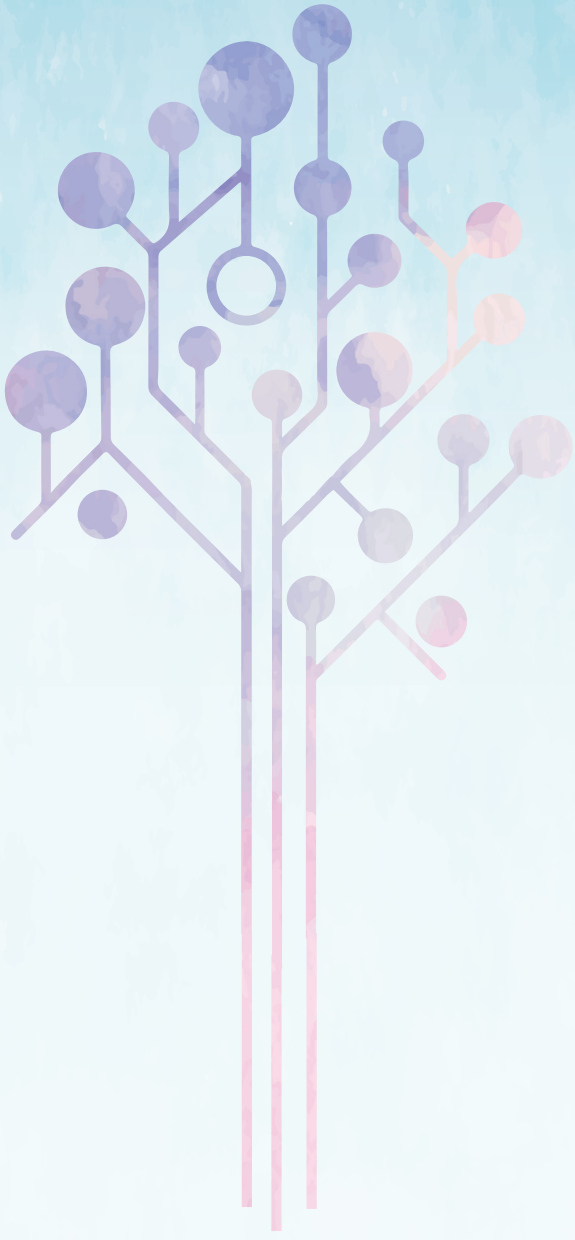
1. World Health Organization. Caring for children and adolescents with mental disorders: Setting WHO directions. https://www.who.int/mental_health/media/en/785.pdf. Published 2003. Accessed 15-01, 2020.
2. Shaffer-Hudkins, E., Suldo, S., Loker, T., March, A. How Adolescents' Mental Health Predicts Their Physical Health: Unique Contributions of Indicators of Subjective Well-being and Psychopathology. *Applied Research in Quality of Life.* 2010;5(3):203-217.
3. Reef, J., van Meurs, I., Verhulst, F. C., van der Ende, J. Children's problems predict adults' DSM-IV disorders across 24 years. *J Am Acad Child Adolesc Psychiatry.* 2010;49(11):1117-1124.
4. Signorini, G., Singh, S. P., Boricevic-Marsanic, V., Dieleman, G., Dodig-Ćurković, K., Franic, T., et al. Architecture and functioning of child and adolescent mental health services: a 28-country survey in Europe. *The Lancet Psychiatry.* 2017;4(9):715-724.
5. Shaw, I., Smith, K. M., Middleton, H., Woodward, L. A letter of consequence: referral letters from general practitioners to secondary mental health services. *Qual Health Res.* 2005;15(1):116-128.
6. Rushton, J., Bruckman, D., Kelleher, K. Primary care referral of children with psychosocial problems. *Arch Pediatr Adolesc Med.* 2002;156(6):592-598.
7. Fickel, J. J., Parker, L. E., Yano, E. M., Kirchner, J. E. Primary care - mental health collaboration: an example of assessing usual practice and potential barriers. *J Interprof Care.* 2007;21(2):207-216.
8. Zuckerman, K. E., Nelson, K., Bryant, T. K., Hobrecker, K., Perrin, J. M., Donelan, K. Specialty referral communication and completion in the community health center setting. *Acad Pediatr.* 2011;11(4):288-296.
9. Stille, C., Primack, W., Savageau, J. Generalist-subspecialist communication for children with chronic conditions: A regional physician survey. *Pediatrics.* 2003;112:1314-1320.
10. Najim, H., Singh, P. Critical Analysis of Psychiatrists' Opinion in GP Referral letter. *Psychiatr Danub.* 2015;27:429-S431.
11. Spaanjaars, N. L., Groenier, M., van de Ven, M. O. M., Witteman, C. L. M. Experience and Diagnostic Anchors in Referral Letters. *European Journal of Psychological Assessment.* 2015;31(4):280-286.
12. Sampson, R., Cooper, J., Barbour, R., Polson, R., Wilson, P. Patients' perspectives on the medical primary-secondary care interface: systematic review and synthesis of qualitative research. *BMJ Open.* 2015;5(10):e008708.
13. Jiwa, M., Dadich, A. M. Referral letter content : can it affect patient outcomes? 2013;19(3S):140-149.
14. Ford, T., Kenchington, R., Norman, S., Hancock, J., Smalley, A., Henley, W., et al. The agreement between the referrer, practitioner and research diagnosis of autistic spectrum conditions among children attending child and adolescent mental health services. *Eur Child Adolesc Psychiatry.* 2019;28(9):1253-1264.

15. Hansen, B. H., Oerbeck, B., Skirbekk, B., Kristensen, H. Non-obsessive-compulsive anxiety disorders in child and adolescent mental health services--Are they underdiagnosed, and how accurate is referral information? *Nord J Psychiatry.* 2016;70(2):133-139.
16. Duncan, L., Comeau, J., Wang, L., Vitoroulis, I., Boyle, M. H., Bennett, K. Research Review: Test-retest reliability of standardized diagnostic interviews to assess child and adolescent psychiatric disorders: a systematic review and meta-analysis. *J Child Psychol Psychiatry.* 2019;60(1):16-29.
17. Wittchen HU, Mühlig S, K., B. Mental disorders in primary care. *Dialogues Clin Neurosci.* 2003;5(2):115-128.
18. Aydin, S., Crone, M. R., Siebelink, B. M., Vermeiren, R., Numans, M. E., Westenberg, P. M. Recognition of anxiety disorders in children: a cross-sectional vignette-based survey among general practitioners. *BMJ Open.* 2020;10(4):e035799.
19. Lempp, T., Heinzel-Gutenbrunner, M., Bachmann, C. Child and adolescent psychiatry: which knowledge and skills do primary care physicians need to have? A survey in general practitioners and paediatricians. *Eur Child Adolesc Psychiatry.* 2016;25(4):443-451.
20. Brattfjell, M. L., Jozefiak, T., Wichstrom, L. Predictors of community versus specialty mental health service use: a prospective cohort study. *Eur Child Adolesc Psychiatry.* 2020.
21. Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., et al. Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry.* 2011;50(1):32-45.
22. Hughes, C. W., Rintelmann, J., Mayes, T., Emslie, G. J., Pearson, G., Rush, A. J. Structured interview and uniform assessment improves diagnostic reliability. *J Child Adolesc Psychopharmacol.* 2000;10(2):119-131.
23. Aebi, M., Kuhn, C., Metzke, C. W., Stringaris, A., Goodman, R., Steinhausen, H. C. The use of the development and well-being assessment (DAWBA) in clinical practice: a randomized trial. *Eur Child Adolesc Psychiatry.* 2012;21(10):559-567.
24. Jensen, A. L., Weisz, J. R. Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. *Journal of Consulting and Clinical Psychology.* 2002;70(1):158-168.
25. Ramirez Basco, M., Bostic, J. Q., Davies, D., Rush, A. J., Witte, B., Hendrickse, W., et al. Methods to improve diagnostic accuracy in a community mental health setting. *Am J Psychiatry.* 2000;157(10):1599-1605.
26. Goodman, R., Ford, T., Richards, H., Gatward, R., Meltzer, H. The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. *Journal of Child Psychology and Psychiatry.* 2000;41(5):645-655.
27. Jorm, A. F., Wright, A., Morgan, A. J. Where to seek help for a mental disorder? National survey of the beliefs of Australian youth and their parents. *Med J Aust.* 2007;187(10):556-560.
28. Kramer, T., Garralda, M. E. Child and adolescent mental health problems in primary care. *Advances in Psychiatric Treatment.* 2018;6(4):287-294.
29. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (5th ed.).* Washington, DC.2013.

30. Working Party on ICPC. *International Classification of Primary Care (ICPC). Manual for use of ICPC in relevance studies.* Amsterdam: Department of General Practice, University of Amsterdam; 1985.
31. Lamberts H, Wood M, eds. *International Classification of Primary Care.* Oxford: Oxford University Press; 1987.
32. Nederlands Huisartsen Genoorschap ["Dutch College of General Practitioners"]. ICPC-online. https://www.nhg.org/themas/publicaties/icpc-online. Accessed January - March, 2018.
33. Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., et al. A children's global assessment scale (CGAS). *Arch Gen Psychiatry.* 1983;40(11):1228-1231.
34. Davis, K. A., Sudlow, C. L., Hotopf, M. Can mental health diagnoses in administrative data be used for research? A systematic review of the accuracy of routinely collected diagnoses. *BMC Psychiatry.* 2016;16:263.
35. Nesvag, R., Jonsson, E. G., Bakken, I. J., Knudsen, G. P., Bjella, T. D., Reichborn-Kjennerud, T., et al. The quality of severe mental disorder diagnoses in a national health registry as compared to research diagnoses based on structured interview. *BMC Psychiatry.* 2017;17(1):93.
36. Sociaal en Cultureel Planbureau (SCP). Statusscores 2016 (Netherlands socioeconomic status by postcode). https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores/statusscores_2016. Published 2016. Accessed 4 June 2018.
37. Hsiao, J. K., Bartko, J. J., Potter, W. Z. Diagnosing diagnoses. Receiver operating characteristic methods and psychiatry. *Arch Gen Psychiatry.* 1989;46(7):664-667.
38. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
39. Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., et al. epiR: Tools for the Analysis of Epidemiological Data. 2019.
40. Baeyens, J. P., Serrien, B., Goossens, M., Clijsen, R. Questioning the "SPIN and SNOUT" rule in clinical testing. *Arch Physiother.* 2019;9:4.
41. Fischer, J. E., Bachmann, L. M., Jaeschke, R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med.* 2003;29(7):1043-1051.
42. Ford, E., Campion, A., Chamles, D. A., Habash-Bailey, H., Cooper, M. "You don't immediately stick a label on them": a qualitative study of influences on general practitioners' recording of anxiety disorders. *BMJ Open.* 2016;6(6):e010746.
43. Hartveit, M., Thorsen, O., Biringer, E., Vanhaecht, K., Carlsen, B., Aslaksen, A. Recommended content of referral letters from general practitioners to specialised mental health care: a qualitative multi-perspective study. *BMC Health Serv Res.* 2013;13:329.
44. Culshaw, D., Clafferty, R., Brown, K. Let's get physical! A study of general practitioner's referral letters to general adult psychiatry--are physical examination and investigation results included? *Scott Med J.* 2008;53(1):7-8.
45. Scottish Association for Mental Health and Information Services Division Scotland, NIHS. Rejected Referrals Child and Adolescent Mental Health Services (CAMHS): A qualitative and quantitative audit. In: Group Scotland, T. S., Edinburgh EH6 5NA PPDAS433246 (06/18), ed.: The Scottish Government; 2018:

https://www.gov.scot/binaries/content/documents/govscot/publications/research-and-analysis/2018/06/rejected-referrals-child-adolescent-mental-health-services-camhs-qualitative-quantitative/documents/00537523-pdf/00537523-pdf/govscot%3Adocument/00537523.pdf.
46. Chui, Z., Gazard, B., MacCrimmon, S., Harwood, H., Downs, J., Bakolis, I., et al. Inequalities in referral pathways for young people accessing secondary mental health services in south east London. *Eur Child Adolesc Psychiatry.* 2020.
47. Kosten, T. A., Rounsaville, B. J. Sensitivity of psychiatric diagnosis based on the best estimate procedure. *Am J Psychiatry.* 1992;149(9):1225-1227.
48. Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Group, Q.-S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clin Epidemiol.* 2013;66(10):1093-1104.
49. Rutjes, A. W., Reitsma, J. B., Di Nisio, M., Smidt, N., van Rijn, J. C., Bossuyt, P. M. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ.* 2006;174(4):469-476.
50. Pfeiffer, A. M., Whelan, J. P., Martin, J. M. Decision-making bias in psychotherapy: Effects of hypothesis source and accountability. *Journal of Counseling Psychology.* 2000;47(4):429-436.
51. You, D. S., Youngstrom, E. A., Feeny, N. C., Youngstrom, J. K., Findling, R. L. Comparing the Diagnostic Accuracy of Five Instruments for Detecting Posttraumatic Stress Disorder in Youth. *J Clin Child Adolesc Psychol.* 2017;46(4):511-522.
52. Vugteveen, J., De Bildt, A., Hartman, C. A., Timmerman, M. E. Using the Dutch multi-informant Strengths and Difficulties Questionnaire (SDQ) to predict adolescent psychiatric diagnoses. *Eur Child Adolesc Psychiatry.* 2018;27(10):1347-1359.
53. Hinrichs, S., Owens, M., Dunn, V., Goodyer, I. General practitioner experience and perception of Child and Adolescent Mental Health Services (CAMHS) care pathways: a multimethod research study. *BMJ Open.* 2012;2(6).
54. Hamburger, E. K., Lane, J. L., Agrawal, D., Boogaard, C., Hanson, J. L., Weisz, J., et al. The referral and consultation entrustable professional activity: defining the components in order to develop a curriculum for pediatric residents. *Acad Pediatr.* 2015;15(1):5-8.
55. Royal College of General Practitioners. Guide to supporting information for appraisal and revalidation. Presenting your evidence: Referral letters and responses. https://www.rcgp.org.uk/training-exams/practice/revalidation/guide-to-supporting-information-for-appraisal-and-revalidation/reflection-and-professionalism.aspx. Accessed May 19, 2020.
56. Nederlands Huisartsen Genootschap. NHG-Richtlijn Informatieuitwisseling huisarts - tweedelijns GGZ,. In: Utrecht2011: https://www.nhg.org/sites/default/files/content/nhg_org/uploads/final_webversie_ha-ggz_sk_lres_10nov11_0.pdf.
57. Tobin-Schnittger, P., O'Doherty, J., O'Connor, R., O'Regan, A. Improving quality of referral letters from primary to secondary care: a literature review and discussion paper. *Prim Health Care Res Dev.* 2018;19(3):211-222.
58. Durbin, J., Barnsley, J., Finlayson, B., Jaakkimainen, L., Lin, E., Berta, W., et al. Quality of communication between primary health care and mental health care: an examination of referral and discharge letters. *J Behav Health Serv Res.* 2012;39(4):445-461.

We keep asking 'how to reduce the gap between science and practice' but how about widening the common space between both?



Chapter 4

The diagnostic process from primary care to child and adolescent mental healthcare services: The incremental value of information conveyed through referral letters, screening questionnaires and structured multi-informant assessment

Published as: Aydin, S., Siebelink, B.M., Crone, M.C., Van Ginkel, J., Numans, M.E., Vermeiren, R.R.J.M., & Westenberg, P.M. The diagnostic process from primary care to child and adolescent mental healthcare services: The incremental value of information conveyed through referral letters, screening questionnaires and structured multi-informant assessment. British Journal of Psychiatry Open, 2022, 8(3), E81, doi:10.1192/bjo.2022.47

The formulation of a clinical diagnosis is critical to the child and adolescent mental healthcare (CAMH).[1-3] The current approaches for the diagnostic process include the judgement of a clinician or the use of structured assessment instruments. Four decades of research support the use of structured instruments, which results in more consistent application of diagnostic criteria, a decrease in information variance and bias, and improved recognition of less obvious or secondary conditions.[4-6] Clinical and evidence-based assessment (EBA) guidelines therefore recommend integration of both methods, to benefit from the nuance and parsimony associated with clinical judgement, combined with the accuracy and reliability intrinsic to structured assessment.[7 8] As in clinical practice with stepped-care and matched-care approaches, assessment is conducted in sequential stages, with EBA the question is raised as to whether instruments meaningfully contribute to the diagnostic work-up and how far each additional information step overlaps. Although the various instruments have been studied for value as stand-alone measures,[7] less is known about the incremental value of the various nodes of information. Given the tension between efficiency of information gathering and reliability in the diagnosticprocess,[9] a better understanding is needed of the value of a validated diagnostic work-up; in this case, a work-up that captures the combined benefits of structured assessment and clinical judgement, suggesting potential for use at the interface between primary and secondary CAMH. Accordingly, the aim of the present study was to investigate the *incremental* value of routinely gathered successive assessments. We investigated the added value of referral letters, a screening questionnaire and a structured multi-informant assessment gathered during the registration procedure at an academic centre for child and adolescent psychiatry

The diagnostic procedure

In several countries it is standard practice for CAMH registration to take place via front-line practitioners such as paediatricians or general practitioners. If a decision is made, based on screening or clinical judgement, to refer to CAMH, a referral letter indicating the probable mental health diagnosis forms a bridge to CAMH. For many children and adolescents, referral letters represent the only form of information transfer from the referrer, and may contribute to the diagnostic and treatment process in CAMH. Although many professionals in the field believe that referral letters have no clinical value, in a recent study, we found that 42-93% of youth reasons for referral saw no change in later psychiatric diagnosis.[10]

Although these numbers are substantial, we also observed considerable variation between disorder groups, with internalizing problems in particular showing a relatively poor detection accuracy.

In EBA, a decision to refer should follow administration of a screening instrument. This procedure allows for the common false positives of screening instruments to be corrected by clinician judgement, and acknowledges that screening often helps improve detection of less obvious problems such as internalising disorders, thereby improving adequate referrals and access to treatment. Regrettably, the use of screening instruments is infrequent, a problem often attributed to the limited time available for patient consultation.[11] Many of the current short screening questionnaires were specifically developed to address this problem. Unintendedly, development of these questionnaires may have further limited their implementation, because understanding the pros and cons of the wide array of current screening instruments, together with interpretation of outcomes, has become more challenging.[11-14] A recent review of accessible CAMH instruments identified 672 questionnaires, of which only four broad screening instruments qualified as brief, free, accessible, and with excellent psychometric characteristics.[14] One of these instruments is the strengths and difficulties questionnaire (SDQ), available in translation in over 70 languages.[15] The SDQ was found to be as reliable and feasible as the much lengthier Achenbach scales (the Youth self report (YSR), Child behaviour checklist (CBCL) and Teacher report form (TRF)) that are frequently used in many European countries.[16-18] The developers of the SDQ proposed using the instrument before a clinical appointment as a guide to decision-making.[19] However, regarding recognition of emotional problems, studies suggest that the SDQ might be insufficient, a problem likely related to the limited number of questions in the scale, to differences in study samples and to general difficulties in detecting internalizing disorders.[15 20]

The detection of mental health problems, including internalizing problems, often improves with the use of more extensive assessment methods. In EBA, more extensive assessment methods are in fact recommended in the case of individuals with high scores during screening. The Development and Well-Being Assessment (DAWBA) instrument combines the responses of various informants' (adolescents, parents and/or caregivers and teachers) to

closed-ended questions into so-called DAWBA band scores that indicate the likelihood of a child having any of 17 common mental health disorders. [19] [21] The DAWBA band scores were envisioned as a way to avoid the costly involvement of a clinician and to be a pragmatic solution for common issues at the point of care. Nonetheless, the value of DAWBA bands when accounted for the value of screening and clinical judgement in primary care is not yet investigated.

As part of the DAWBA, informants are also prompted to describe their problems and the context of their problems in their own words. These are then evaluated by a clinician who integrates the various factors to form a relatively nuanced image without the high cost of a full interview with a specialist clinician. DAWBA clinician ratings (CRs) were found to be conservative regarding the number of diagnoses made when compared with elaborate diagnostic interviews.[22] Studies of clinician rated DAWBA found that it was useful in reducing unnecessary referral for externalizing disorders, and that they highlighted internalizing disorders that would not have been detected otherwise.[23] [24] Nevertheless, the exact extent to which clinician ratings supplements information from a primary care clinician , screening results and automatised DAWBA probability band scores remains an important but unanswered question.

Aims

In summary, the feasibility and psychometric properties of the DAWBA and SDQ have been individually well-researched in community, clinical and research settings in various European countries. However, less information is available regarding the predictive value of instruments when taking into account the usual overlap of information gained during successive steps in EBA. The aim of the present study was to determine both the unique and incremental predictive values for four sources of information in predicting a medical record consensus diagnosis: referral letters, a screening questionnaire (SDQ[15]), a more elaborate structured assessment (DAWBA band scores[19]) and the remote evaluation of structured and unstructured responses by a clinician (the clinician rated DAWBA). We hypothesized that each instrument would show incremental value in predicting the classification of five disorder groups commonly treated in CAMH: anxiety, depression, autism spectrum disorders (ASD), attention-deficit hyperactivity disorder (ADHD), and behavioural disorders.

Methods

Data source and procedure

The starting point for the sample consisted of children and adolescents who were referred to Leiden University Medical Centre Curium (LUMC Curium). LUMC Curium is an inpatient and outpatient mental health clinic delivering specialized care to young people aged 3 to 18 years.

About 70% of the yearly caseload at the institution consists of outpatient referrals that follow a routine procedure including referral letters, the SDQ and DAWBA. The remainder consists of inpatient referrals that follow a referral intake procedure adapted to cases in need of urgent evaluation, in which case questionnaires are not completed at registration. We included young people who I) registered between January 2015 and December 2017, II) followed the routine procedure including the SDQ and DAWBA, and III) had an accessible referral letter in the medical record. The procedures used to extract and code referral letters are described in detail in our recent publication on referral letters general practice.[10] To briefly summarize, using an iterative process we created a manual to extract and code text in referral letters. The manual was then tested for inter-rater reliability by authors S.A., M.R.C., B.M.S. and P.M.W. (κ=.77 to κ=.90). We did not differentiate symptoms indicated in referral letters from suggested diagnoses. For instance, when a referral letter reported “treatment for anxiety disorders?” or “fearful”, both were coded as an indicator of the category anxiety disorders and related problems. Multiple indications were often found in referral letters and were thus coded. However, fewer than 20% of referral letters indicated more than four problems,[10] which was also the case in the current sample.

The LUMC Medical Ethical committee waived a need for informed consent because of the retrospective nature of the study (approval number G18.080). Furthermore, the data management plan was approved by the scientific committee of the LUMC Departments of Public Health and Primary Care, LUMC Curium Child and Adolescent Psychiatry, and the Institute of Psychology of Leiden University.

Measures

All measures were extracted from medical records. We extracted referral letters as they were scanned and filed in individual patient medical records. The SDQ, structured DAWBA

data and classifications that are also outcome measure were extracted simultaneously from the medical record system.[25]

In The Netherlands, only a healthcare professional can make a formal referral to youth and adolescent psychiatry, which then proceeds via either general practice, specialised healthcare (hospitals) or youth welfare offices (also called local youth teams). We did not include the type of professional as a covariate in the main analyses, as initial logistic regression analyses showed wide confidence intervals and no statistically significant differences between the various types of referrers.

Structured assessment: SDQ and DAWBA

During registration, families are provided with unique login codes for the online DAWBA package, which can be completed by up to two parents or caregivers, the young person themselves (if aged >11 years) and up to two teachers. The package always starts with the SDQ, and then moves on to the DAWBA instrument. Rules regarding skipping come into play when an informant shows low scores on a conceptually-related SDQ scale and provides negative answers to a gate-keeping question at the beginning of each DAWBA chapter.[19] In the DAWBA package, SDQ scale scores and DAWBA probability band scores are generated for each informant individually and subsequently integrated into an overall SDQ score for each scale (0, 1, 2) and a DAWBA probability band score for each chapter (0 to 5). The cut-off scores and rules concerning integration of informant's scores can be found at www.sdqinfo.org and www.dawba.net. If not otherwise specified, we used integrated scores for all analyses. To analyse whether each assessment method indicated the presence of a disorder group, we dichotomized scores by separating the upper two scores from the lower score(s).[21 24 26]

SDQ

The SDQ covers four problem areas (emotional, conduct, hyperactivity and peer problems scales) across 20 items, asks about children's strengths in five items (prosocial scale), and the impact and burden of problems in eight items. Informants rate items on a three-point Likert scale (0= not true, 1=somewhat true, 2=certainly true), with higher scores indicating more problems. Although the SDQ was not formally created to give indications of a probable autism spectrum disorder (ASD), in a later study Goodman et al[27] proposed use of a difference score by subtracting the total for the peer problems scale from the score for the

prosocial scale. We calculated this difference score solely based on parental scores, as the few studies available suggest that parents show the highest accuracy in detecting ASD.[20 28 29]

DAWBA probability band scores

The DAWBA[19] estimates the likelihood of the presence of 17 common mental health disorders. These so-called probability bands are automatically generated in the online DAWBA environment by integrating various informant responses to closed-ended questions.[21] The questions are linked to the DSM criteria and result in probability band scores of 0, 1, 2, 3, 4 and 5, corresponding to prevalences found in the original British epidemiologic sample and approximating likelihoods <0.1% , 0.5%, 3%, 15%, 50% and >70%.[21] Thus, a probability band score of 5 suggests that 70% or more of the cases with a similar response profile to the British reference sample were found to have that diagnostic outcome. When the DAWBA did not produce a score for a disorder group (e.g. behavioural disorders), we took the highest probability band score among the more specific disorders (i.e. the highest score among conduct and oppositional deviant disorder).[21]

Clinician rated DAWBA

Informants are also prompted to describe problems and their context in their own words. A senior clinical psychologist evaluated the open-ended questions, together with the SDQ and DAWBA probability band results, and scored the likelihood of a disorder on a three-point scale (absent, unsure, present). This final stage facilitates the incorporation of the diverse strands of information to develop a nuanced image without the accompanying cost of visiting a specialist clinician. The next step is to add a short report to a patient medical record in order to guide prioritization of appointments and to prevent tunnel vision during a face-to-face intake. In some study reports, clinician ratings are referred to as a DAWBA research diagnosis. In this paper, however, we use the term clinician rating to prevent confusion with the outcome classification.

Clinical Classification

The primary outcome measure was a patient's digital medical record classification according to the Longitudinal, Expert, and All Data (LEAD) procedure.[30] This is a product of all collected information and clinical judgement, including patient and family history, mental health treatment history, structured assessment and, if necessary, process diagnostics and additional assessment methods depending on suspected differential diagnoses.[31 32] Based on

these insights, a case conceptualization is formed as a basis for treatment initiation and a classification selected and entered into the patient's medical record. Up to five different classifications could be recorded per case and all were extracted for this study.

Missing data

SDQ scale scores were available for all cases and DAWBA band scores were available for 97.7-98.9% of cases, depending on disorder group, but clinician-rated DAWBA data were available for only 52.1% of cases, as DAWBAs were not evaluated by a clinician during the first half of the study period. As this was a result of management decisions and unrelated to our research question, we could assume the data to be missing at random. To reliably estimate missing data, we applied multiple imputation (MI, with m=100) using the *mice* package in the R environment.[33-37] Multiple imputation creates multiple sets with plausible values for missing cells by drawing values from the observed cases and by predicting from other associated variables in a dataset. Hence, it minimizes bias relative to complete case analysis. Generating multiple datasets enables estimation of the uncertainty in the imputation process as compared with, for example, simple mean imputation. In multiple imputation, it is necessary to balance the number of predictors and observed cases, as with regression analyses in general. Therefore, we limited the number of predictors during multiple imputation such that a minimum number of 15 cases had to be observed for each contributing predictor.

Statistical analysis

In the statistical analysis, we first computed diagnostic metrics such as sensitivity and specificity for each instrument. Next, we inspected youth diagnostic trajectories through the current sequence of four methods. To this end, we cross-tabulated frequencies of positive and negative indications in a four-layer table with each of the methods and the diagnostic outcome. To examine the effect of each added predictor on model fit, likelihood ratio tests[38] were performed with the *D3()* function in *mice*.[33] Multiple logistic regression analyses were performed, with each of the five diagnostic groups as the outcome and the assessment methods as the predictor, in order to quantify unique and corrected predictive values. Diagnostic odds ratios (ORs) of the instruments were extracted from the univariable and multivariable logistic regression models.

Results

The sample age ranged between 5 and 18 years (mean=11.08, s.d.=3.45) and 57.4% were boys (Table 1).

Univariable diagnostic metrics

The diagnostic metrics of the assessment methods as standalone measures are depicted in Table 2. The sensitivity and specificity of the successive assessment tools varied per mental health disorder. The value of referral letters in detecting patients with anxiety disorders was relatively low compared with the other disorder groups and to the other instruments: 46.9% of those eventually classified with an anxiety disorder had been indicated as such in referral letters. However, referral letters showed a relatively high specificity in excluding minors without the condition (85.9%). The highest sensitivity regarding anxiety disorders was found for the SDQ (95.1%), but was accompanied by a risk of being overinclusive (specificity 22.9%; false discovery rate 85.2%, supplementary material). The SDQ and referral letters showed the highest sensitivity or specificity, respectively, the DAWBA probability band and the clinician-rated DAWBA showed a more balanced profile.

We found that all instruments except the SDQ performed similarly in discriminating minors with or without depressive disorders (Table 2). In line with earlier studies, the SDQ frequently gave a positive indication in this clinical sample, yet often for the wrong persons (specificity 22.4%).

Upon inspecting the metrics for ASDs, the low number of positive indications by the DAWBA probability band was remarkable. Although the bands indicated ASDs infrequently, they did so for genuine cases, resulting in a high positive predictive value (78.3%, supplementary material) but low sensitivity (9.0%). The SDQ difference score (peer problems - prosocial score, see Methods) showed the highest sensitivity for ASDs as compared with other instruments. In contrast to high false positives for anxiety and depressive disorders, the SDQ showed a better specificity for ASDs (54.7%). Referral letters and clinician rated DAWBA scores showed a fairly even balance of sensitivity and specificity for ASD.

When considering ADHD, most instruments showed values similar to those for ASDs, with the DAWBA probability band showing the best performance in the detection of ADHD (sensitivity 59.3%).

Table 1. Sample Characteristics N=1259		
		n (%)
Age, years	5-9	474 (37.6)
	10-14	508 (40.4)
	15-18	277 (22.0)
Gender	Male	723 (57.4)
	Female	536 (42.6)
CGAS	20-40	83 (6.6)
	41-50	503 (40.0)
	51-60	514 (40.8)
	>61	96 (7.6)
	Missing	63 (5.0)
Medical conditions	None classified	958 (76.1)
	Singular	99 (7.9)
	Complex	52 (4.1)
	Missing	150 (11.9)
Number of clinical classifications (comorbidity)	0	175 (13.9)
	1	544 (43.2)
	2	368 (29.2)
	3	125 (9.9)
	4	35 (2.8)
	5-6	12 (1.0)
Type of clinical classifications		
Neurodevelopmental disorders		727 (57.7)
Schizophrenia spectrum and other psychotic disorders		2 (0.2)
Depressive disorders		134 (10.6)
Anxiety disorders		174 (13.8)
Obsessive-compulsive and related disorders		13 (1.0)
Trauma and stressor-related disorders		68 (5.4)
Somatic symptom and related disorders		42 (3.3)
Feeding and eating disorders		54 (4.3)
Elimination disorders		19 (1.5)
Gender dysphoria		7 (0.6)
Disruptive, impulse-control, and conduct disorders		71 (5.6)
Substance-related and addictive disorders		2 (0.2)
Personality disorders		49 (3.9)

Distributions of the clinical classifications in the sample are depicted based on the higher order chapters of the DSM-5 (e.g. Neurodevelopmental disorders). The number of clinical classifications is depicted on the level of the specific disorders (e.g. ADHD and ASD). CGAS= Children's Global Assessment Scale score.

Behavioural disorders were frequently indicated by all instruments yet seldom classified. This resulted in a very low predictive value. This frequent indication of behaviour problems resulted in relatively high sensitivity (86.4%).

After inspecting single descriptives, we explored frequencies of the instrument's successive positive and negative indications in order to gain insight into the potential of the sequence

Table 2. Two by two cross-tabulation of the instruments per disorder group

		Anxiety disorders		Depressive disorders		ASD		ADHD		Behavioural disorders	
		+	-	+	-	+	-	+	-	+	-
RL	+	38 (46.9)	81 (14.1)	39 (60.0)	73 (12.4)	108 (54.8)	89 (45.2)	114 (55.9)	99 (22.0)	26 (59.1)	156 (25.5)
	-	43 (53.1)	492 (85.9)	26 (40.0)	516 (87.6)	89 (45.2)	361 (80.2)	90 (44.1)	350 (78.0)	18 (40.9)	455 (74.5)
SDQ	+	77 (95.1)	442 (77.1)	62 (95.4)	457 (77.6)	140 (71.1)	204 (45.3)	181 (88.7)	230 (51.2)	38 (86.4)	328 (53.7)
	-	4 (4.9)	131 (22.9)	3 (4.6)	132 (22.4)	57 (28.9)	246 (54.7)	23 (11.3)	219 (48.8)	6 (13.6)	283 (46.3)
Band	+	57 (70.4)	185 (32.2)	45 (69.2)	94 (16.0)	18 (9.1)	5 (1.1)	121 (59.3)	78 (17.4)	16 (36.4)	225 (36.9)
	-	24 (29.6)	388 (67.7)	20 (30.8)	495 (84.0)	179 (90.9)	445 (98.9)	83 (40.7)	371 (82.6)	28 (63.6)	384 (63.1)
CR	+	62 (76.5)	194 (33.9)	49 (75.4)	104 (17.7)	151 (76.6)	154 (34.2)	170 (83.3)	158 (35.2)	26 (59.1)	200 (32.7)
	-	19 (23.5)	379 (66.1)	16 (24.6)	485 (82.3)	46 (23.4)	296 (65.8)	34 (16.7)	291 (64.8)	18 (40.9)	411 (67.3)

Frequency (%) of the positive and negative indications made per instrument and per disorder group, as a ratio of the total number of positive and negative cases. RL= referral letters, SDQ= strength and difficulties questionnaire, Band= DAWBA probability band score, CR= clinician rating in the DAWBA environment. Number of diagnoses and sample size were as follows: anxiety disorders n=81 and N=654; depressive disorder n=65 and N=654, autism spectrum disorders (ASD) n=197 and N=647; attention-deficit hyperactivity disorder (ADHD) n=204 and N=653; behavioural disorders n=44 and N=655. n= number of cases, N=total sample size.

for prognostic use. Of the youth with an anxiety disorder indicated by all four instruments, 48.8% were eventually classified with anxiety disorders (supplementary material). The classification rate was 54.9% for four successive indications of depressive disorders, 85.7% for ASDs, 70.0% for ADHDs and 10.7% for behavioural disorders.

When we considered the predictive value of successive negative indications, we found that 98.2% of those negative on all four instruments were not classified to anxiety disorders, 98.3% were not classified to depressive disorders, 90.5% to ASDs, 95.8% to ADHD and 99.1% were not classified to behavioural disorders.

Incremental and independent predictive values

When we examined the incremental value of the four assessment tools relative to each other, successive addition of a following instrument resulted in improvement in model fit for nearly all of the (4*5) models (Table 3). Only the fit for behavioural disorders did not improve with addition of the clinician-rated DAWBA scores to the model (p=0.82).

By controlling for the value of up to three other instruments, we explored independent associations of the four instruments with the outcome classifications (Figure 1). In these multivariable models most instruments showed predictive value. Only in the case of the SDQ did we see a failure to improve the prediction of depressive disorders and behavioural disorders (respectively OR=1.24, 95% CI 0.58-2.62; OR=1.85, 95% CI 0.82-4.16).

For most disorder groups and instruments, we found no differences in magnitude of the associations in the multivariable models as compared with the univariable prediction models. Similarly, no difference in patterns was observed when inspecting differences in the predictive value of the earlier instruments compared with the later instruments. The clinician rated DAWBA, for instance, did not show consistently higher predictive values as compared with the referral letters.

Table 3. Likelihood ratio test values comparing the effect of addition of instruments on model fit per disorder group

	RL	+SDQ	+DAWBA Band	+DAWBA CR
Anxiety disorders	92.74	33.47	41.81	15.1
Depressive disorders	136.81	8.28*	39.63	17.48
ASD	166.29	44.48	15.25	14.50
ADHD	203.53	79.52	42.23	39.58
Behavioural disorders	44.26	16.04	16.78	0.02**

Likelihood ratio test results depicting change in model fit by successive addition of the instruments, computed in the imputed dataset. All values are significant at the p<.001 level, except *p=.004 and **p=0.82. ASD= Auism spectrum disorders. ADHD= attention-deficit hyperactivity disorders. PPV= positive predictive value. NPV= negative predictive value. Note the low frequency of four successive positive indications for ASDs and ADHD, as it was uncommon for these minors to have positive scores on all four instruments.

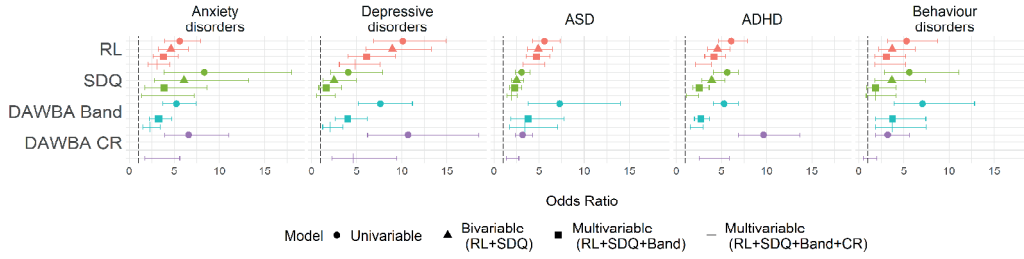


Figure 1. Univariable and multivariable odds ratios per instrument and per diagnostic outcome

Odds ratios per instrument and per disorder group for four models, computed in the imputed dataset. Each successive model contains one more instrument as predictor, presenting how the ORs change when controlling for overlap with more instruments. The vertical line presents OR=1.

Conclusion

To the best of our knowledge, this study is the first to compare the predictive value of referral letters, broad band screening, structured multi-informant assessment and a clinician's remote evaluation in predicting diagnostic outcome in a single population. We found that all four nodes of assessment generally showed a positive contribution to the prediction of common child adolescent mental health problems. Referral letters and SDQ scale scores showed either a high sensitivity or a high specificity, whereas DAWBA probability bands and clinician ratings were more balanced in terms of sensitivity and specificity. Referral letters performed especially well for depressive disorders, which might be related to an earlier observation made during the pilot phase of our previous study: professionals might focus on mood problems and associate it with risk of suicidal ideation.[39] For the other disorder groups, referral letters showed better performance in terms of specificity compared with sensitivity. The SDQ, by contrast, was overinclusive, particularly for emotional problems,[20] a finding in line with earlier conclusions that advised against complete reliance on the SDQ to guide referrals.[18] To determine whether this might be a result of our categorization of the SDQ indication as positive from the upper two indications on, we reanalysed the data categorising only the upper category as positive. This resulted in a sensitivity decrease of 15 percentage points (to 80.5 for anxiety and 78.5 for depressive disorders), whereas specificity doubled to around 50% false positives Nonetheless, in comparison with the other instruments, SDQ screening was still overinclusive, an issue inherent to a screening instrument's function (to detect problems), the clinic population, and, as underlined in the introduction, screening instruments should be accompanied by clinical judgement.

Although the SDQ does not officially have an ASD scale, we also included children and adolescents with ASD in the study to shed light on the issue of EBA in this clinically-widespread population. We used a difference score suggested by the SDQ developers[27] and found that youngsters with ASD were detected at a similar rate to other problem types on conceptually related SDQ scales. However, other studies have used other computational methods,[20 29 40 41]and the respective methods have not yet been compared.

We also inspected frequencies of successive positive and negative indications as a first approach to the question of outcomes for young people who show successive positive or negative scores on a sequence of assessment instruments. In this explorative inspection we found that four successive indications of anxiety or depressive disorders resulted in only a one in two chance of being classified to these outcomes. By contrast, when all instruments indicated ASD or ADHD, cases were indeed clinically classified as such. Regarding behaviour problems, we found that even four successive positive indications were not predictive of a classification to behavioural disorders. When considering the opposite situation, those with four successive negative indications, we found that about 1% was classified to anxiety, depressive or behavioural disorders, whereas around 5% or 10% were still classified to ADHD or ASD, respectively. It is unsurprising that rates were highest for ASD, because if initial instruments fail to suggest this relatively difficult diagnosis further clinician based investigations subsequently detect ASD. These results underline the need for elaborate diagnostics, the inclusion of clinicians when aiming for specialised treatment and the importance of future studies with a diverse sample for better generalisability.

We found added benefits with each successive node of assessment, with only one exception for one outcome: the clinician ratings showed no improvement in the prediction of behavioural disorders relative to the three previous instruments combined. This might be because of the already marginal prediction of behavioural disorders and to the relatively conservative properties of the clinician rated DAWBA.[22] With regards to the independent predictive value, we found that nearly all instruments remained individually associated with the outcome even when corrected for overlap with other instruments. Only the SDQ showed no independent value in predicting depressive and behavioural disorders when corrected for information provided by other nodes of assessment. In contrast to general literature suggesting that instruments applied later in a sequence might show stronger effects,[42] we observed no increase in effect. While each instrument certainly contributed new information in our study, there was no indication that the most extensive assessment method should have been used first. Therefore, the study results give no support for use of the most elaborate instrument first and only, and support a stepwise approach to assessment.[43]

Limitations

Although this study presented unique data on an important question, some limitations should be kept in mind. Firstly, people involved in classifying outcomes were not blinded to the instrument's results. To what extent results were viewed when formulating a diagnosis is not known. As regards the effect of the availability of DAWBA data, for instance, there are indications that it improves decision-making in the case of internalising problems but not in the case of externalizing problems.[24] In an effort to explore this type of potential effect, we split the sample between those with or without clinician ratings (see Methods section) but did not find differences in odds ratios between subsamples. Regardless, if disclosure had any effect it would likely result in the presented odds ratios overestimating associations. Looked at more positively, our research question concerned the relative predictive value of the instruments and, in principle, all instruments were accessible and have shown predictive value, also, in other studies with blinding.

Another limitation concerns discriminant ability of the instruments. If the aim is to predict the type and classification of a problem, insight into how scales relate to conceptually parallel classifications is not sufficient. Future studies could therefore focus on the discriminant ability of the tools and investigate cross relations between scales and types of problems. Furthermore, we focused only on the type of problems, whereas taking the staging and impact of symptoms into account could benefit clinical practice.[44]

Implications

The questions addressed in this study are directly relevant to clinical practice. Referral letters are by definition available for many cases yet are seldom incorporated into the diagnostic process. In this study we found that referral letters add value, even when corrected for overlap with structured assessment instruments. Similarly, the DAWBA package has the potential to ease the assessment process by capturing the SDQ as a short yet sensitive screening instrument, the DAWBA structured questions as a broad assessment tool to "cast a wide net regarding the presenting problem of a client",[45] and the clinician rated DAWBA to add some nuance regarding the fuller picture while not being overinclusive. When used within a sequential approach, the DAWBA package may help develop a shared language between primary care and specialised care professionals and parents, just as the DAWBA package also produces a report for parents when requested.[46] This in turn might stimulate

fruitful discussions within families and help ameliorate discrepancies between the problem perceptions of minors versus caregivers, the perceived focus of treatment and treatment outcomes.[1 43 47 48] Moreover, a harmonised sequential diagnostic approach might facilitate real integration and joint working in the primary-secondary care interface, a challenge that has not been overcome despite decades of research and dissemination of the importance of EBA. The idea of working within and towards a complete and reliable work-up might be more palatable as compared with choosing from a list of measures purely based on one's own familiarity and time limits, without any insight regarding subsequent steps.[6 48] Earlier studies found the DAWBA to be relatively conservative in terms of the number of diagnoses made and required administration time when compared with other elaborate diagnostic instruments.[22] This suggests that it might hold potential for use at the primary-secondary care interface, as a second step for those with high scores on screening instruments in primary care and to prioritise referrals and registration in secondary mental healthcare.

Conclusion

In conclusion, our results suggest that integrating referral letters, screening questionnaires and information obtained from assessment is likely to facilitate diagnosis in clinical practice. Prospective studies could further quantify the clinical and economic value of this type of multi-tiered approach in relation to the facilitation of psychometrically sound and feasible decision-making, timely recognition of problems, determination of required care intensities and treatment outcomes.

Reference list Study 3

1. Jensen-Doss, A. and Weisz, J. R., *Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics.* J Consult Clin Psychol, 2008. **76**(5): p. 711-22.
2. Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., et al., *Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment.* Clinical Psychology: Science and Practice, 2017. **24**(4): p. 331-363.
3. Yates, B. T. and Taub, J., *Assessing the costs, benefits, cost-effectiveness, and cost-benefit of psychological assessment: we should, we can, and here's how.* Psychol Assess, 2003. **15**(4): p. 478-95.
4. Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., and Ivanova, M. Y., *Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews.* Int J Methods Psychiatr Res, 2009. **18**(3): p. 169-84.
5. Kuhn, C., Aebi, M., Jakobsen, H., Banaschewski, T., Poustka, L., Grimmer, Y., et al., *Effective Mental Health Screening in Adolescents: Should We Collect Data from Youth, Parents or Both?* Child Psychiatry Hum Dev, 2017. **48**(3): p. 385-392.
6. Jensen-Doss, A. and Hawley, K. M., *Understanding barriers to evidence-based assessment: clinician attitudes toward standardized assessment tools.* J Clin Child Adolesc Psychol, 2010. **39**(6): p. 885-96.
7. Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., and Jensen-Doss, A., *Clinical Guide to the Evidence-Based Assessment Approach to Diagnosis and Treatment.* Cognitive and Behavioral Practice, 2015. **22**(1): p. 20-35.
8. Johnston, C. and Murray, C., *Incremental validity in the psychological assessment of children and adolescents.* Psychol Assess, 2003. **15**(4): p. 496-507.
9. Sayal, K., Tischler, V., Coope, C., Robotham, S., Ashworth, M., Day, C., et al., *Parental help-seeking in primary care for child and adolescent mental health concerns: qualitative study.* Br J Psychiatry, 2010. **197**(6): p. 476-81.
10. Aydin, S., Crone, M. R., Siebelink, B. M., Ginkel, J. v., Numans, M. E., Vermeiren, R. R. J. M., et al., *Informative value of referral letters from general practice for child and adolescent mental healthcare.* European Child & Adolescent Psychiatry, 2021. **(forthcoming)**.
11. Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., et al., *Free, brief, and validated: Standardized instruments for low-resource mental health settings.* Cogn Behav Pract, 2015. **22**(1): p. 5-19.
12. Doss, A. J., *Evidence-based diagnosis: incorporating diagnostic instruments into clinical practice.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(9): p. 947-52.
13. O'Brien, D., Harvey, K., Howse, J., Reardon, T., and Creswell, C., *Barriers to managing child and adolescent mental health problems: a systematic review of primary care practitioners' perceptions.* Br J Gen Pract, 2016. **66**(651): p. e693-707.
14. Becker-Haimes, E. M., Tabachnick, A. R., Last, B. S., Stewart, R. E., Hasan-Granier, A., and Beidas, R. S., *Evidence Base Update for Brief, Free, and Accessible Youth Mental Health Measures.* J Clin Child Adolesc Psychol, 2020. **49**(1): p. 1-17.

15. Goodman, R., *The Extended Version of the Strengths and Difficulties Questionnaire as a Guide to Child Psychiatric Caseness and Consequent Burden.* Journal of Child Psychology and Psychiatry, 1999. **40**(5): p. 791-799.
16. Becker, A., Woerner, W., Hasselhorn, M., Banaschewski, T., and Rothenberger, A., *Validation of the parent and teacher SDQ in a clinical sample.* Eur Child Adolesc Psychiatry, 2004. **13 Suppl 2**: p. II11-6.
17. van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D., and Goodman, R., *Dutch version of the Strengths and Difficulties Questionnaire (SDQ).* Eur Child Adolesc Psychiatry, 2003. **12**(6): p. 281-9.
18. Janssens, A. and Deboutte, D., *Screening for psychopathology in child welfare: the Strengths and Difficulties Questionnaire (SDQ) compared with the Achenbach System of Empirically Based Assessment (ASEBA).* Eur Child Adolesc Psychiatry, 2009. **18**(11): p. 691-700.
19. Goodman, R., Ford, T., Richards, H., Gatward, R., and Meltzer, H., *The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology.* Journal of Child Psychology and Psychiatry, 2000. **41**(5): p. 645-655.
20. Vugteveen, J., De Bildt, A., Hartman, C. A., and Timmerman, M. E., *Using the Dutch multi-informant Strengths and Difficulties Questionnaire (SDQ) to predict adolescent psychiatric diagnoses.* Eur Child Adolesc Psychiatry, 2018. **27**(10): p. 1347-1359.
21. Goodman, A., Heiervang, E., Collishaw, S., and Goodman, R., *The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples.* Soc Psychiatry Psychiatr Epidemiol, 2011. **46**(6): p. 521-32.
22. Angold, A., Erkanli, A., Copeland, W., Goodman, R., Fisher, P. W., and Costello, E. J., *Psychiatric diagnostic interviews for children and adolescents: a comparative study.* J Am Acad Child Adolesc Psychiatry, 2012. **51**(5): p. 506-17.
23. Citroen, A., Siebelink, B., and van Lang, N., *DAWBA: de nieuwe poortwachter van de jeugd-ggz?* Tijdschrift voor gezondheidswetenschappen, 2017. **95**(6): p. 246-249.
24. Aebi, M., Kuhn, C., Metzke, C. W., Stringaris, A., Goodman, R., and Steinhausen, H. C., *The use of the development and well-being assessment (DAWBA) in clinical practice: a randomized trial.* Eur Child Adolesc Psychiatry, 2012. **21**(10): p. 559-67.
25. Nesvag, R., Jonsson, E. G., Bakken, I. J., Knudsen, G. P., Bjella, T. D., Reichborn-Kjennerud, T., et al., *The quality of severe mental disorder diagnoses in a national health registry as compared to research diagnoses based on structured interview.* BMC Psychiatry, 2017. **17**(1): p. 93.
26. Crone, M. R., Zeijl, E., and Reijneveld, S. A., *When do parents and child health professionals agree on child's psychosocial problems? Cross-sectional study on parent-child health professional dyads.* BMC Psychiatry, 2016. **16**: p. 151.
27. Goodman, A., Lamping, D. L., and Ploubidis, G. B., *When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children.* J Abnorm Child Psychol, 2010. **38**(8): p. 1179-91.
28. Iizuka, C., Yamashita, Y., Nagamitsu, S., Yamashita, T., Araki, Y., Ohya, T., et al., *Comparison of the strengths and difficulties questionnaire (SDQ) scores between children with high-functioning autism spectrum disorder (HFASD) and attention-deficit/hyperactivity disorder (AD/HD).* Brain Dev, 2010. **32**(8): p. 609-12.

29. Vugteveen, J., de Bildt, A., Theunissen, M., Reijneveld, S. A., and Timmerman, M., *Validity Aspects of the Strengths and Difficulties Questionnaire (SDQ) Adolescent Self-Report and Parent-Report Versions Among Dutch Adolescents.* Assessment, 2021. **28**(2): p. 601-616.
30. Spitzer, R. L., *Psychiatric diagnosis: Are clinicians still necessary?* Compr Psychiatry, 1983. **24**: p. 399-411.
31. Jensen-Doss, A., Youngstrom, E. A., Youngstrom, J. K., Feeny, N. C., and Findling, R. L., *Predictors and moderators of agreement between clinical and research diagnoses for children and adolescents.* J Consult Clin Psychol, 2014. **82**(6): p. 1151-62.
32. Jensen, A. L. and Weisz, J. R., *Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents.* Journal of Consulting and Clinical Psychology, 2002. **70**(1): p. 158-168.
33. van Buuren, S. and Groothuis-Oudshoorn, K., *Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software, 2011. **45**: p. 1-67.
34. van Ginkel, J. R., Linting, M., Rippe, R. C. A., and van der Voort, A., *Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data.* J Pers Assess, 2020. **102**(3): p. 297-308.
35. Rubin, D. B., *Multiple Imputation for Nonresponse in Surveys.* 2004, New York: John Wiley & Sons Inc.
36. van Buuren, S., *Flexible imputation of missing data"*. 2018, Boca Raton: Chapman & Hall/CRC Interdisciplinary Statistics.
37. R Core Team, *R: a Language and Environment for Statistical Computing*. 2021, R Foundation for Statistical Computing,: Vienna, Austria.
38. Meng, X.-L. and Rubin, D. B., *Performing Likelihood Ratio Tests with Multiply-Imputed Data Sets.* Biometrika, 1992. **79**(1).
39. Aydin, S., Crone, M. R., Siebelink, B. M., Vermeiren, R., Numans, M. E., and Westenberg, P. M., *Recognition of anxiety disorders in children: a cross-sectional vignette-based survey among general practitioners.* BMJ Open, 2020. **10**(4): p. e035799.
40. Russell, G., Rodgers, L. R., and Ford, T., *The strengths and difficulties questionnaire as a predictor of parent-reported diagnosis of autism spectrum disorder and attention deficit hyperactivity disorder.* PLoS One, 2013. **8**(12): p. e80247.
41. Salayev, K. A. and Sanne, B., *The strengths and difficulties questionnaire (SDQ) in autism spectrum disorders.* International Journal on Disability and Human Development, 2017. **16**(3).
42. Moons, K. G., de Groot, J. A., Linnet, K., Reitsma, J. B., and Bossuyt, P. M., *Quantifying the added value of a diagnostic test or marker.* Clin Chem, 2012. **58**(10): p. 1408-17.
43. Whitmyre, E. D., Adams, L. M., Defayette, A. B., Williams, C. A., and Esposito-Smythers, C., *Is the focus of community-based mental health treatment consistent with adolescent psychiatric diagnoses?* Child Youth Serv Rev, 2019. **103**: p. 247-254.
44. Scott, J., Leboyer, M., Hickie, I., Berk, M., Kapczinski, F., Frank, E., et al., *Clinical staging in psychiatry: a cross-cutting model of diagnosis with heuristic and practical value.* Br J Psychiatry, 2013. **202**(4): p. 243-5.
45. Mash, E. J. and Hunsley, J., *Evidence-based assessment of child and adolescent disorders: issues and challenges.* J Clin Child Adolesc Psychol, 2005. **34**(3): p. 362-79.
46. Bennett, S. D., Heyman, I., Coughtrey, A. E., Buszewicz, M., Byford, S., Dore, C. J., et al., *Assessing feasibility of routine identification tools for mental health disorder in neurology clinics.* Arch Dis Child, 2019. **104**(12): p. 1161-1166.

47. Priebe, S., McCabe, R., Bullenkamp, J., Hansson, L., Lauber, C., Martinez-Leal, R., et al., *Structured patient-clinician communication and 1-year outcome in community mental healthcare: cluster randomised controlled trial.* Br J Psychiatry, 2007. **191**: p. 420-6.
48. Kazdin, A. E., *Evidence-based assessment for children and adolescents: issues in measurement development and clinical application.* J Clin Child Adolesc Psychol, 2005. **34**(3): p. 548-58.

"Knowing when to say 'I don't know' is half of knowledge."

Malik ibn Anas

Chapter 5

General discussion

In this dissertation we aimed to contribute to gaps in the evidence on timely recognition, referral and assessment of children and adolescents with mental health disorders by addressing three main questions. The first main question was: to what extent do professionals recognize typical symptoms of mental health problems. As a major example of a prevalent but overlooked mental health problem, we experimentally focused on the early recognition of anxiety disorders. As a first in literature, this study focused on professionals' very first diagnostic considerations (chapter 2). We presented both general practitioners and mental health professionals with examples of how problems could be expressed by minors with anxiety disorders and their families during a typical consult and asked for professionals first diagnostic interpretation. The second main objective was to quantify the predictive value of reasons for referral once the need for mental health treatment is recognized in minors. To this end we retrospectively extracted and coded referral letters from general practice to child and adolescent psychiatry. A coding scheme was developed and suggestions for meaningful use of referral letters are presented (chapter 3). Thirdly, we investigated whether an integrated use of various assessment forms can potentially improve the prediction of the type of mental health disorder. Many studies on evidence-based assessment instruments exist, yet in practice the question remained how we could increase their meaningful use. To add to this end, we investigated the incremental value of structured screening and elaborated assessment together with referral letters and clinicians remote evaluation of available information (chapter 4). In the current chapter, the findings of these studies are summarized and discussed. Subsequently general implications for future studies and clinical practice are elaborated on.

Summary of findings

Recognition

Studies on the prevalence and underrecognition of mental health disorders are common. In most of these studies recognition is quantified as a diagnostic outcome. Such studies examine whether professionals recognize the presence of a demarcated mental health disorder. In clinical practice, however, it is rather the rule than exception that early mental health problems are not defined clearly. This is of major importance when considering timely recognition. In the face of undefined problems, professionals' familiarity with the frequent

occurrence of anxiety disorders is critical. Therefore, we questioned whether professionals sufficiently consider the likelihood of the presence of an anxiety disorder in their early decision-making. As anxiety disorders are a relatively silent problem associated with a broad range of symptoms, we hypothesized that the underrecognition of anxiety disorders could be related to professionals' unfamiliarity with the wide variety of symptoms and how widespread these problems are –and not only related to for example patients avoidance of disclosure during consult situations. Nonetheless, to the best of our knowledge, there were no empirical studies on the early diagnostic interpretation of professionals as a factor of influence on early recognition. To contribute to this question, we presented general practitioners and mental health professionals with hypothetical cases, so called vignettes, in which mixed anxiety symptoms were described (chapter 2). The professionals were asked what their first interpretation of the type of problem was. Both general practitioners and mental health professionals selected anxiety related problems less than what could be expected based on the content of the vignettes and compared to their selection rate of the other type of mental health problems. General practitioners recognized anxiety related problems also less than what could be expected based on the prevalence of anxiety disorders in community and primary care samples. Notably, general practitioners and mental health professionals did not differ significantly with regard to their recognition rate of the other disorder groups in these mixed anxiety vignettes. The study results (chapter 2) suggest that already at the very beginning of the diagnostic process professionals overlook anxiety disorders. In view of early recognition and prevention, the question remains what the recognition rates would be if they were not made aware of the focus of the study on psychosocial problems during the informed consent.

Referral

Anxiety disorders tend to be evaluated as a relatively mild problem.[1] In line therewith, in the vignette study (chapter 2) we hypothesized that general practitioners would report that it is usually suitable and sufficient to treat anxiety disorders in primary care.[1] This would partly explain why most anxiety disorders are less often referred for specialised mental health treatment.[2] Contrary to expectation, general practitioners reported preferring a treatment in mental healthcare for anxiety disorders when asked explicitly. This finding suggests that general practitioners do not per se underprioritise anxiety disorders, and that the low

referral rates in clinical practice might be more of an issue of underrecognition rather than underprioritisation and trivializing of anxiety disorders. Also, in line with this conclusion, the referral letters investigated in the second study (chapter 3) showed the lowest agreement rate for anxiety disorders when inspecting the agreement between the reason for referral and the classifications of disorders that were made in child and adolescent psychiatry. Even in referred samples, the recognition of anxiety disorders by general practitioners, falls behind when compared to other common mental health disorders.[3]

Although there are several studies on the substantial value of reasons for referral as indicated on referral letters for the adult mental health process, only two earlier studies presented the predictive value of referral letters to child and adolescent mental healthcare.[4] [5] One of these studies is on autism spectrum disorders and one on non-obsessive compulsive anxiety disorders. In chapter 3, we presented agreement metrics between the reason for referral mentioned in referral letters and the clinical classification made in child and adolescent mental healthcare for all commonly treated mental health disorders obtained from a large sample referred by general practitioners. To this end we coded the reasons for referral by including both tentative diagnoses and symptoms and problems mentioned in referral letters. Over half of the referral letters indicated one or more of the core symptoms or the tentative diagnoses of the clinically established classifications. When indications of internalizing and developmental/ externalizing problems were grouped, we found that respectively half and two-thirds of the outcomes were in line with the suggested reason for referral. Variations between specific disorder groups were observed with the lowest sensitivity for anxiety disorders and the highest for eating disorders. Referrers often indicated the child's context, such as problems between parents, difficulties studying or being bullied. We found no effects of gender, age, the severity of the problems as estimated by the specialist or whether the child has a mental healthcare history or not, except for age and attention-deficit hyperactivity disorders (ADHD). Referral letters better predicted ADHD with increasing age. An indication of urgency or a severe status in the referral letters was significantly associated with a lower functional impairment score (CGAS-score, as estimated in child and adolescent mental healthcare). All in all, the study results imply that there is more opportunity to draw from the contents of referral letters than expected based on anecdotal evidence and clinicians' stances, as well as room to improve the value and use of

referral letters, particularly on the level of the specific disorders (e.g., anxiety disorders versus depressive disorder).

Assessment

Adequate information is critical to facilitate diagnosis making and allocating the needed care.[6] This holds both for decision-making in primary care and secondary care. However, as the advised assessment methods differ at various stages of care (e.g., primary care versus specialised mental healthcare) yet also overlap, insight is needed in the incremental value of various nodes of information. This is what we examined in chapter 4 of this dissertation. Here we linked the I) referral letters that form a proxy for the tentative diagnosis made by referring clinicians, II) to results of the broadband screening questionnaire with potential for use in primary care, III) to results of elaborate structured assessment with potential for use at registration and diagnosis-making at secondary mental healthcare, and IV) to the tentative remote diagnosis made by a clinician with online access to the results of the previous instruments. The value of these four nodes of assessment in predicting the best estimate clinical consensus diagnoses was investigated (also called LEAD diagnoses; longitudinal, expert, and all data[7]). Nearly all instruments showed statistically significant independent predictive value in predicting the classification of commonly treated mental health disorders (chapter 4). This suggests that structured acquisition and integrated use of information obtained at various stages of the care landscape might add to the diagnostic procedure. Although each instrument certainly contributed additional information in our study, there was no indication that the most extensive assessment method (the remote evaluation of a clinician of the DAWBA open and close ended questions) holds the highest value and should have been used first and only.[8] More specifically, we found that the different instruments showed different strengths and weaknesses. In general, whereas the referral letters were conservative about indicating specific mental health disorders (high specificity, lower sensitivity), the screening questionnaire was over-inclusive (high sensitivity, low specificity). This exemplifies the importance of integrated use of clinicians judgement with a screening instrument (the SDQ) in primary care. The more extensive method (i.e. the automatised DAWBA score and the remote clinical rating within the DAWBA environment) showed a more balanced profile considering sensitivity and specificity. The study results illustrate that these nodes of information all have a unique value.

Clinical implications and practical value

The studies in this dissertation highlight potential to improve the diagnostic process for minors with mental health problems. The first study implicates the necessity to focus on professionals' awareness of the likelihood of a mental health problem (chapter 2). To ameliorate the likelihood of recognition and adequate management, professionals' familiarity with characteristic features, prevalence, and impact of mental health problems needs to be improved. Although precise diagnosis-making is not the task of a general practitioner, the finding that they recognized anxiety even less than the prevalence of anxiety disorders in the population highlights the importance of increasing knowledge to improve their ability to identify or exclude problems. As portrayed in the case described in Box 1, awareness is a step towards enquiring further, and without further enquiring recognition is unlikely or delayed.[9 10] Knowledge of the epidemiology of a disorder could increase awareness on the presence of caseness during consultations. Especially in practices where structured assessment is not implemented, readily available knowledge by the person of the clinician is crucial.[11]

Results of the second study imply that information transferred from general practice holds substantial value for specialised mental healthcare (chapter 3). When viewed as more than a bureaucratic piece of paper by both primary and secondary care professionals, referral letters cosuld be incorporated explicitly in the referral-intake process in mental healthcare and elaborated on during the interview with patients. Then, next to improving communication between primary and specialised care,[6] referral letters might contribute to diminishing families experience of fragmented care. This potential of referral letters is not a capacity to depreciate as many families experience a referral as being in limbo which in turn negatively impacts their reach for support and clinical outcomes.[12-14] To reach this potential however, a clear division of the responsibilities is needed at the referring and receiving end of healthcare as well as a mutual understanding of what referral letters stand for.[6] For the latter, an important consideration is what the reason for referral reveals: do the reasons for referral reflect referrers' true diagnostic opinion or is it what they believe needs to be written to access mental healthcare? Nonetheless, there is ground to assume that referral letters mirror the most outstanding symptoms and complaints as captured by its writer.[15]

Box 1

Imagine again a consultation in general practice. This time, Ann, an-eleven-year-old girl, visits for her recurring earache. The GP, Dr. Hartveld, knows Ann as a shy child who cooperates well with no other particularities. While she is examining the child's ears, the mother mentions that although Ann looks timid, she can certainly be temperamental. Usually, this occurs when they are in a hurry.

The GP considers whether it would be wise to ask anything about what is told here, in between the lines. Even though there is no explicit request for help considering these problems, she decides to enquire further. It turns out that the mother means that although Ann can sit so quietly, she can also get pretty upset. About three times a week. But "Luckily Ann also has many strong points". She goes to school as usual even when she has not slept well because of the earaches. Ann doesn't want to miss any school and wants to finish all her homework properly. Her study skills are good. The teacher thinks Ann is a smart child. She does say, however, that "Ann can respond rather impatiently if she can not handle certain situations". The initial interpretation of the GP is that this might be an example of typical development in a child with a strong will or might point towards characteristics on the autism spectrum. She also remembers that recently she had learned about how common anxiety disorders are and that they might manifest as temper tantrums.

The GP concludes that it is likely that Ann shows subclinical levels of anxiety symptoms as her earaches happen to strike before spelling tests. She has a tendency to worry a lot which might be why she shows temperamental outbursts when they are in a hurry or doing something new. Ann is also able to develop an adequate relation with the GP and mother does not report other typicalities pointing towards autism. Dr. Hartveld decides to hand parents a screening questionnaire which they later discuss with the general practice-based mental health nurse. The mental health nurse shares some tips regarding coping with her worries and temper tantrums. They decide to monitor her anxious behaviour; especially as she will go to secondary school next year and these transition periods often mark whether a child will adapt or experience more challenges.

A few months later mother revisits and explains that they have read more about emotional and behavioural problems in families. She explains that she recognized several other challenges related to the anxious attitude of Ann and also experienced difficulties in how they, as parents, should respond to her needs. The GP searches in the medical record for the notes of the general practice-based mental health nurse. He had reported that the family could consult the local youth teams for parenting support or be referred to mental healthcare if Ann is particularly hindered by her fears. Parents agree with a referral and the GP writes out a referral letter in which she shares the consult notes and the results of the screening questionnaire. Once the family registers with Ann at mental healthcare an intake is planned. Since they have no other request and the screening results that were shared pointed towards emotional problems only, the professional starts with a more elaborate assessment and clinical interview to gain and create a more in depth understanding of the context of their challenges, strengths and possibilities.

The transferred information is what the referrer recognized or acknowledged as a possible need and substantiated its likeliness to reach mental health treatment.[16] From this view, referral letters could be seen as a product of the referrer's decision-making and form an approximation of their repertoire, beliefs and preferences for diagnosis and treatment. In this perspective, referral letters are no more to be viewed as the result of a simple discretionary activity but to be acknowledged as valuable for clinical and educational purposes. When viewed as such, referral letters could also inform policy-making and curricula development for general practitioners. In some countries, referral letters are already included in postgraduate training.[17 18] As concluded by a review study on communication in healthcare, further efforts are required to feature communication between professionals as an essential skill of caregivers.[15]

The third study is a first infer on the potential of a sequential approach to assessment investigated in a clinical dataset (chapter 4). Within data obtained from a best practice, we explored the incremental value of four structured assessment instruments applicable for different purposes in and between primary and secondary mental healthcare. The results suggest that they all add, but none are sufficient to determine the diagnostic classifications. This implies that it is worthwhile to do further research on the integrated use of these instruments across the board of primary and secondary care and how they impact the decision-making. Studies on integrated care found that professionals do report a need for sound methods to evaluate and monitor patients' needs, however, the selection and interpretation of such methods forms an initial barrier and constraints in time and resources form further barriers.[19 20] The evidence-based assessment implementation is called "the thorniest challenge faced in the mental health field".[21] An explicit sequential procedure that efficiently brings together purposes as initial detection and later prioritisation of registration (i.e., from screening, up to elaborate assessment and evaluation by clinicians), and helps to combine all information from various informants, might be a method to bridge needs in terms of assessment at the interface of primary and secondary care and add to implementing stepped-care and matched-care approaches. A recent systematic review concluded that there is a need for readily available systems and methods that target primary care physicians as they work in an environment with fewer resources.[22] Unfortunately, there are not much well disseminated lists of psychometrically sound instruments that suffice to

feasibility criteria and are adaptable to the needs in and between primary and secondary care are scarce.[23] The instruments presented in chapter four form an example to this end.

Although there are proven benefits of using sound assessment methods in the clinical process,[24-27] we do not aim to suggest their dogmatic use for the sake of evidence-based assessment in itself. Evidence-based assessment captures both the process and the methods of assessment (chapter 1), and aiming to increase the use of structured assessment methods should not oppose the process - or vice versa. The aim of evidence-based assessment is to improve the delivery of high quality care through "integrating individual clinical expertise with the best available external clinical evidence from systematic research" (chapter 1[28]). Thus, clinicians should be able to freely move between procedures and omit components when for the benefit of the individual patient. In our view, this is a practically self-evident side note, as is implicitly portrayed in the methods section of chapter four by the many missing data points due to circumstances that are the reality of day-to-day clinical practice (such as emergency referrals). However, an unsubstanstiated aversion towards structured assessment in the name of flexibility and patient satisfaction will not benefit the quality of care. When asked to clinicians in the field, the focus is yet too often placed on the perceived difficulties of structured assessment (such as that it might impede patient satisfaction or time constraints), whereas patient surveys do not subscribe to this concerns.[29] From a policy point of view, the inclusion of assessment instruments in primary care might also be economically beneficial. A prospective pilot study on the effects of the SDQ[30] and the DAWBA[27] at the interface of primary care found that it improved the detection of internalizing problems and decreased unnecessary referral for externalizing problems. The latter type of problems often forms pressure on referrals whereas young people with internalizing problems are not detected sufficiently without elaborate diagnosis making.[10 11 31]

In a recent evaluation of the changes since the youth health act, a shared element in the problematic cases was found to be the absence of a timely assessment and qualitative diagnostics.[32 33] Considering the research agenda on evidence-based assessment, three consecutive steps are described:[24] evaluating the evidence on the accuracy of methods and for the population at hand, evaluating feasibility in terms of costs for institutions, clinicians and patients, and implementation. The many efforts and reports on the first two steps now

give passage to the last but not easiest step concerning implementation. Since we now do have a massive list of structured assessment instruments,[34] it is time to focus on the how and when of assessment in daily practice.[24] One practical port of implementation for these methods in the Netherlands, might be the mental health nurses that are commonly appointed in general practices since 2008. The number of practices with a mental health nurse increased from 27% in 2010 to 87% in 2016 and their workflow is crystallized increasingly considering help for adults.[35 36] Nevertheless, there are major variations between practices and mental health nurses and no guidelines for most child and adolescent mental health problems. Compared to general practitioners, most mental health nurses see fewer children and adolescents as most mental health nurses are educated in adult psychiatry.[36] This further emphasizes the importance of clear guidelines and integrated approaches for detection and diagnosis in primary care.

Methodological considerations and suggestions for future studies

The studies in this dissertation fill gaps in the evidence base on recognition, referral, and assessment of child and adolescent mental health problems. However, future studies extending the study-methods are needed to better inform daily practice and policy-making. In the following section, starting points for such future studies are introduced and elaborated on. Specific limitations of each empirical study have already been mentioned in the previous chapters. Here we will reflect on more general overarching methodological issues and starting points for future studies.

A limitation of the study results, in general, is the focus on symptoms and signs as defined in the diagnostic and statistical manual of mental disorders (DSM[37 38]), sometimes at the expense of inclusion of the role of personal and social stressors.[39] Patients surrounding and contextual factors are critical as they affect mental health and the decision-making.[40-42] In the vignette study described in chapter 2 for instance, even though the vignettes were created with some context, the influence of factors such as family and school context on decision-making were not examined in specific.[43] This vignette study was the first to use descriptions specifically developed to mimic real-life consultations and early presentation of

yet undefined symptoms. This was sufficient to investigate average recognition. However, it was not sufficient to explore the specific effects of the described symptoms and expressions on the professionals' decision-making. Future studies using a factorial vignette design[44] could systematically vary potential determinants of decision-making such as the level of impairment or the described content of the symptoms and signs.[40 45 46] Also, to enhance our understanding of moderators of recognition, factors related to the person of the professional might be added. Studies implementing a qualitative approach might help us gain further understanding of the professionals' internal recognition process.[45]

Considering the study described in chapter 3, we examined the frequency of contextual factors mentioned in the referral letters. An improvement would be to include the interaction between context, symptoms and level of impairment in future studies. Also, for generalisable conclusions on the value of referral letters for the diagnostic process, a multi-center study is necessary as diagnostic metrics are dependent on the prevalence of the type of mental health disorder, which could differ between institutions.[47] Another methodological consideration for the studies on referral letters is the question of "value". As a first infer, we mainly focused on value as being an agreement between referral letters and the final classifications made. This reflects a top-down approach starting from the perspective of mental healthcare. Such a top-down approach is not likely to result in lasting improvements, especially when viewed in light of previous studies in various areas of medicine that found that there are discrepancies in how primary care physicians perceive their referral letters and how specialised care specialists evaluate them.[6 12 16] These studies reveal that both parties evaluate their own communication and referral letters as qualitatively better, and luckily, that they both wish to receive more feedback from each other.[6] To stimulate sincere communication and collaboration between primary and secondary healthcare, future studies could take a bottom-up approach by, for example, starting with a qualitative approach to the topic of referral letters and aim to understand the end-product of a referral letter by focusing on professionals wordings, associations, internal working schemes and implicit algorithms. Also, other elements that make a referral letter valuable should be explored and investigated to facilitate efficacy in patients' care journey. Such elements are for instance, how the referral came into being or the patient's own help request.[6 48]

A related limitation applies to the third study described in chapter 4. In that study, we had included referral letters from all referrers including for instance, local youth teams and paediatricians. A pattern seemed to emerge when we analyzed the possible effects of the pathway to care. In general the associations between what the referrer mentioned and the outcome were higher for specialist referrals than for general practitioner referrals, which were higher than that of local youth teams. Nonetheless, we found very wide confidence intervals and no significant effects. This might be a result of the categories we had created by combining various types of specialists (for example: psychiatrists and paediatricians) and the major variations between local youth teams as they were still developing as a newly created function group. Future studies in diverse samples could implement a multilevel analysis to account for potential effects of the pathway to care. With such a study, the strengths and weaknesses of each type of referrer could be found and build upon.

In chapter 4, we aimed to shed light on the issue of incremental value of evidence-based assessment instruments. The results of this study should be viewed with the realization that we inspected incremental value in the chronological order in which the data was obtained. As described in the introduction of chapter 4, ideally, the decision to refer should result from both clinical judgement and screening. Thus the screening should be followed by the decision to refer, rather than being obtained after referral (Figure 3). Future studies could inspect incremental value in this order. Also, rather than investigating a selected sample that reached secondary care, future studies could investigate the value of a sequential approach in a prospective design by starting with a sample that is just visiting primary care professionals. This will enable better quantification of the false negatives and true negatives. Furthermore, studies that quantify the broad impact and societal benefits of an integrated sequential approach are needed, as cost-utility is a major topic in today's healthcare landscape and we have a sufficient base to move from descriptive studies to comparative and validation studies.[22 49] An earlier randomized controlled trial in specialised care found that disclosure of the results of the DAWBA had a statistically significant effect on the classifications made for some anxiety disorders, but not other disorders.[50] The authors conclude that the DAWBA should be examined as a referral tool rather than an assessment instrument. No comparative studies, however, measured effects at the interface of primary and secondary care. Last but not least, the study measures could be enriched with readily

available data such as consult notes of clinicians from various stages of healthcare (Figure 3). This could facilitate analysis of different stages of signaling and triage from the build-up to the initial inquiry to the initiation of treatment. By definition, medical records and consult notes capture broad and longitudinal health information with the potential to enable highly generalisable study results, the exploration of changes over time and integration of environmental (i.e., family) data. If proven as a sufficient prediction or classification tool, studies on the automatised use of readily available data might provide an efficient decision support system and contribute to timely detection and reducing the workload of professionals.

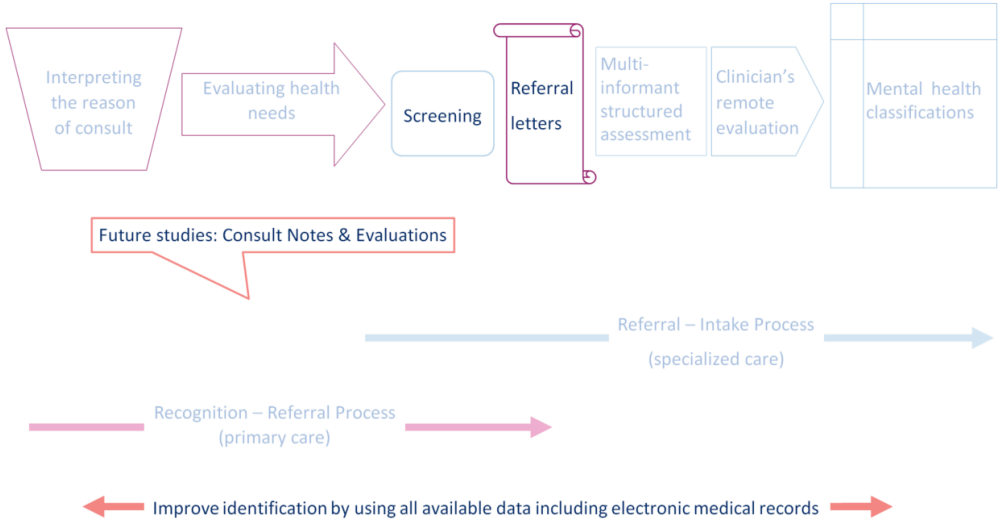


Figure 3. Future studies

In future prospective studies, the order of the instruments could better resemble the reality at the interface of primary and secondary care, by completing the screening in primary care, before initiating a management plan or referral. Future studies could also investigate identification, referral and treatment by integrating other readily available data.

To note on the methodological paradigm of the studies, the starting point for this dissertation is the idea that drawing up specific symptoms is beneficial. This relies on several assumptions. To begin with, it relies on the assumption that some sort of diagnosis-making and assessment is necessary to guide towards focused and effective treatment. Supporting this premise, several studies suggest that disorder-specific treatment methods are more effective than general interventions.[51-54] Used as a label to communicate a cluster of problems that often occur together, classifications influence health outcomes by facilitating

the selection of appropriate treatment techniques.[26 55 56] However, in practice, with scarcity and efficiency as bottlenecks, a relevant question is: “eliciting symptoms, then what?”,[57] or, assessing, then what? Especially given the status quo considering waiting lists for mental healthcare, we would like to note that recognition is insufficient as long as referring professionals have to decide when and how one can reach support or treatment.[1 16 58 59] Efforts to improve access and efficacy of treatment should continue, while increasing efforts to improve timely recognition, and prevention.[60]

Moreover, and relatedly, is the concept of diagnosis and gold standard in psychiatry. In each of the studies, we used the best estimate-all data classifications that were made by drawing from both structured assessment and clinical judgment.[61] Nonetheless, in the pure sense of the word, first the outcome should be operationalised perfectly to quantify predictive value. This means that a gold standard should have perfect diagnostic metrics (100% sensitivity and 100% specificity).[62] Given that this is not the reality in medicine and even less in psychiatry, future studies on prediction and assessment could account for and calculate the uncertainty in the outcome measure.[63]

Irrespective of the reliability of study methods and statistics used, the topical question of the validity of diagnosis-making in psychiatry remains. An in-depth review and discussion of this topic is beyond the scope of the empirical studies presented here. However, independent from the question of what we should measure, we hold the stance that a form of diagnosis - and thus assessment- is needed to guide the mental health treatment. To inform on the focus of the treatment, it is necessary to understand where a patient and his or her system is standing in terms of strengths and challenges in various fronts such as cognitions, interpersonal characteristics, coping and social support. Thus regardless of what the reference standard concerns (the well-known and often used DSM classifications,[37] or alternatives such as the general psychopathology factor,[64] transdiagnostic factors,[65] the context of the challenges as in the power threat meaning framework,[66 67] or syndemics[41]), a method to bridge gaps in assessment in and between primary and secondary mental healthcare is needed. The methods described in this study might act as such examples in designing and testing acceptable methods that meet the needs of primary care (meeting the needs of detection, management and referral) and secondary care (facilitating prioritisation

of registrations, implementation of reliable assessment and monitoring of treatment effects).[68]

Conclusion

The societal toll and human misery associated with mental health disorders is well established. Nonetheless, only about one in five minors with mental health difficulties access adequate professional support. Care pathways and procedures have to facilitate timely recognition and adequate evaluation of patients' needs to navigate those who can benefit towards the frequently cited and meaningful goal of the right service in the right place, at the right time, and delivered by the right person. In the ideal situation, as in the example described above (Box 1), professionals can 'look' at a patient and 'see' patients' needs by relying on their sufficient knowledge to recognize a probable mental health need, their skills to enquire further, known methods to reliably assess strengths and weaknesses, and readily available resources to translate what they see into an adequate support or management plan.

Reference list General discussion

1. O'Brien, D., Harvey, K., Young, B., Reardon, T., and Creswell, C., *GPs' experiences of children with anxiety disorders in primary care: a qualitative study.* Br J Gen Pract, 2017. **67**(665): p. e888-e898.
2. Chavira, D. A., Stein, M. B., Bailey, K., and Stein, M. T., *Child anxiety in primary care: prevalent but untreated.* Depress Anxiety, 2004. **20**(4): p. 155-64.
3. Asp, M., Lindqvist, D., Fernstrom, J., Ambrus, L., Tuninger, E., Reis, M., et al., *Recognition of personality disorder and anxiety disorder comorbidity in patients treated for depression in secondary psychiatric care.* PLoS One, 2020. **15**(1): p. e0227364.
4. Hansen, B. H., Oerbeck, B., Skirbekk, B., and Kristensen, H., *Non-obsessive-compulsive anxiety disorders in child and adolescent mental health services--Are they underdiagnosed, and how accurate is referral information?* Nord J Psychiatry, 2016. **70**(2): p. 133-9.
5. Ford, T., Kenchington, R., Norman, S., Hancock, J., Smalley, A., Henley, W., et al., *The agreement between the referrer, practitioner and research diagnosis of autistic spectrum conditions among children attending child and adolescent mental health services.* Eur Child Adolesc Psychiatry, 2019. **28**(9): p. 1253-1264.
6. Vermeir, P., Vandijck, D., Degroote, S., Peleman, R., Verhaeghe, R., Mortier, E., et al., *Communication in healthcare: a narrative review of the literature and practical recommendations.* Int J Clin Pract, 2015. **69**(11): p. 1257-67.
7. Spitzer, R. L., *Psychiatric diagnosis: Are clinicians still necessary?* Compr Psychiatry, 1983. **24**: p. 399-411.
8. Moons, K. G., de Groot, J. A., Linnet, K., Reitsma, J. B., and Bossuyt, P. M., *Quantifying the added value of a diagnostic test or marker.* Clin Chem, 2012. **58**(10): p. 1408-17.
9. Elstein AS, S. A. M., *Clinical problem solving and diagnostic decision making: selective review of the cognitive literature.* BMJ, 2002. **324**(7339): p. 729-32.
10. Lykke, K., Christensen, P., and Reventlow, S., *"This is not normal ... "--signs that make the GP question the child's well-being.* Fam Pract, 2008. **25**(3): p. 146-53.
11. Olariu, E., Forero, C. G., Castro-Rodriguez, J. I., Rodrigo-Calvo, M. T., Alvarez, P., Martin-Lopez, L. M., et al., *Detection of Anxiety Disorders in Primary Care: A Meta-Analysis of Assisted and Unassisted Diagnoses.* Depress Anxiety, 2015. **32**(7): p. 471-84.
12. Durbin, J., Barnsley, J., Finlayson, B., Jaakkimainen, L., Lin, E., Berta, W., et al., *Quality of communication between primary health care and mental health care: an examination of referral and discharge letters.* J Behav Health Serv Res, 2012. **39**(4): p. 445-61.
13. Abdinasir, K. and Pona, I., *Access denied: A teenager's pathway through the mental health system*. 2015, The Children's Society: Londen, United Kingdom.
14. Scottish Association for Mental Health and Information Services Division Scotland and NIHS, *Rejected Referrals Child and Adolescent Mental Health Services (CAMHS): A qualitative and quantitative audit*, Group Scotland, T. S., Edinburgh EH6 5NA PPDAS433246 (06/18), Editor. 2018, The Scottish Government.
15. Shaw, I., Smith, K. M., Middleton, H., and Woodward, L., *A letter of consequence: referral letters from general practitioners to secondary mental health services.* Qual Health Res, 2005. **15**(1): p. 116-28.

16. O'Brien, D., Harvey, K., Howse, J., Reardon, T., and Creswell, C., *Barriers to managing child and adolescent mental health problems: a systematic review of primary care practitioners' perceptions.* Br J Gen Pract, 2016. **66**(651): p. e693-707.
17. Hamburger, E. K., Lane, J. L., Agrawal, D., Boogaard, C., Hanson, J. L., Weisz, J., et al., *The referral and consultation entrustable professional activity: defining the components in order to develop a curriculum for pediatric residents.* Acad Pediatr, 2015. **15**(1): p. 5-8.
18. Royal College of General Practitioners. *Guide to supporting information for appraisal and revalidation. Presenting your evidence: Referral letters and responses*. [cited 2020 May 19]; Available from: https://www.rcgp.org.uk/training-exams/practice/revalidation/guide-to-supporting-information-for-appraisal-and-revalidation/reflection-and-professionalism.aspx.
19. Nooteboom, L. A., van den Driesschen, S. I., Kuiper, C. H. Z., Vermeiren, R., and Mulder, E. A., *An integrated approach to meet the needs of high-vulnerable families: a qualitative study on integrated care from a professional perspective.* Child Adolesc Psychiatry Ment Health, 2020. **14**: p. 18.
20. Gilbody, S., Bower, P., and Whitty, P., *Costs and consequences of enhanced primary care for depression: systematic review of randomised economic evaluations.* Br J Psychiatry, 2006. **189**: p. 297-308.
21. McHugh, R. K. and Barlow, D. H., *The dissemination and implementation of evidence-based psychological treatments. A review of current efforts.* Am Psychol, 2010. **65**(2): p. 73-84.
22. Venrooij, *Clinical decision support methods for children and youths with mental health disorders in primary care.* Family practice, 2022.
23. Doss, A. J., *Evidence-based diagnosis: incorporating diagnostic instruments into clinical practice.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(9): p. 947-52.
24. Hunsley, J., *Translating Evidence-Based Assessment Principles and Components Into Clinical Practice Settings.* Cognitive and Behavioral Practice, 2015. **22**(1): p. 101-109.
25. Mash, E. J. and Hunsley, J., *Evidence-based assessment of child and adolescent disorders: issues and challenges.* J Clin Child Adolesc Psychol, 2005. **34**(3): p. 362-79.
26. Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., et al., *Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment.* Clinical Psychology: Science and Practice, 2017. **24**(4): p. 331-363.
27. Citroen, A., Siebelink, B., and van Lang, N., *DAWBA: de nieuwe poortwachter van de jeugd-ggz?* Tijdschrift voor gezondheidswetenschappen, 2017. **95**(6): p. 246-249.
28. Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., and Richardson, W. S., *Evidence based medicine: what it is and what it isn't.* BMJ, 1996. **312**(7023): p. 71-2.
29. Jensen-Doss, A. and Hawley, K. M., *Understanding barriers to evidence-based assessment: clinician attitudes toward standardized assessment tools.* J Clin Child Adolesc Psychol, 2010. **39**(6): p. 885-96.
30. Lambregtse, C., *Kinderproject 'Alle hens aan dek' in Katwijk: Vier vragen bij buikpijn*, in *De dokter.* 2012, Landelijke Huisartsen vereningen.
31. Fleury, M. J., Imboua, A., Aube, D., Farand, L., and Lambert, Y., *General practitioners' management of mental disorders: a rewarding practice with considerable obstacles.* BMC Fam Pract, 2012. **13**: p. 19.
32. Netherlands Youth Institute, *Reform of the Dutch system for child and youth care. 4 years later.* 2019: Utrecht.

33. Friele, R. D., Hageraats, R., Fermin, A., Bouwman, R., Zwaan J. van der *De jeugd-GGZ na de Jeugdwet: een onderzoek naar knelpunten en kansen*. 2019.
34. Becker-Haimes, E. M., Tabachnick, A. R., Last, B. S., Stewart, R. E., Hasan-Granier, A., and Beidas, R. S., *Evidence Base Update for Brief, Free, and Accessible Youth Mental Health Measures.* J Clin Child Adolesc Psychol, 2020. **49**(1): p. 1-17.
35. Magnée T, S. H., Van Weelderen G, Nuijen J, Kenter A., *Hoe denken poh's-ggz over hun werk?* Huisarts en Wetenschap, 2020. **63**.
36. Beurs, D. d., Magnée, T., Bakker, D. de, Verhaak, P. , *Analyse van de inzet van de POH-GGZ in de huisartsenpraktijk over de periode 2010-2015*. 2016: Utrecht: NIVEL.
37. American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (5th ed.)*. 2013, Washington, DC.
38. Andreasen, N. C., *DSM and the death of phenomenology in america: an example of unintended consequences.* Schizophr Bull, 2007. **33**(1): p. 108-12.
39. Sampson, R., Cooper, J., Barbour, R., Polson, R., and Wilson, P., *Patients' perspectives on the medical primary-secondary care interface: systematic review and synthesis of qualitative research.* BMJ Open, 2015. **5**(10): p. e008708.
40. Wissow, L. S., Zafar, W., Fothergill, K., Ruble, A., and Slade, E., *Using vignettes to assess contributions to the work of addressing child mental health problems in primary care.* BMC Health Serv Res, 2016. **15**: p. 584.
41. The, L., *Syndemics: health in context.* The Lancet, 2017. **389**(10072).
42. Lovasi, G. S., Mooney, S. J., Muennig, P., and DiMaggio, C., *Cause and context: place-based approaches to investigate how environments affect mental health.* Soc Psychiatry Psychiatr Epidemiol, 2016. **51**(12): p. 1571-1579.
43. Aydin, S., Crone, M. R., Siebelink, B. M., Vermeiren, R., Numans, M. E., and Westenberg, P. M., *Recognition of anxiety disorders in children: a cross-sectional vignette-based survey among general practitioners.* BMJ Open, 2020. **10**(4): p. e035799.
44. Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., et al., *Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries.* Depress Anxiety, 2018. **35**(3): p. 195-208.
45. Erfanian, F., Latifnejad Roudsari, R., Heydari, A., Noghani Dokht Bahmani, M. A *Narrative on Using Vignettes: Its Advantages and Drawbacks.* Journal of Midwifery and Reproductive Health, 2020. **8**: p. 2134-2145.
46. Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., et al., *Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies.* Int J Clin Health Psychol, 2015. **15**(2): p. 160-170.
47. Abu-Akel, A., Bousman, C., Skafidas, E., and Pantelis, C., *Mind the prevalence rate: overestimating the clinical utility of psychiatric diagnostic classifiers.* Psychol Med, 2018. **48**(8): p. 1225-1227.
48. Hartveit, M., Thorsen, O., Biringer, E., Vanhaecht, K., Carlsen, B., and Aslaksen, A., *Recommended content of referral letters from general practitioners to specialised mental health care: a qualitative multi-perspective study.* BMC Health Serv Res, 2013. **13**: p. 329.
49. Bettencourt, A. F. and Plesko, C. M., *A Systematic Review of the Methods Used to Evaluate Child Psychiatry Access Programs.* Acad Pediatr, 2020. **20**(8): p. 1071-1082.
50. Ford, T., Last, A., Henley, W., Norman, S., Guglani, S., Kelesidi, K., et al., *Can standardized diagnostic assessment be a useful adjunct to clinical assessment in child mental health*

services? A randomized controlled trial of disclosure of the Development and Well-Being Assessment to practitioners. Soc Psychiatry Psychiatr Epidemiol, 2013. **48**(4): p. 583-93.
51. Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., and Christensen, H., *School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis.* Clin Psychol Rev, 2017. **51**: p. 30-47.
52. Fisak, B. J., Jr., Richard, D., and Mann, A., *The prevention of child and adolescent anxiety: a meta-analytic reviews.* Prev Sci, 2011. **12**(3): p. 255-68.
53. Teubert, D. and Pinquart, M., *A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents.* Journal of Anxiety Disorders, 2011. **25**(8): p. 1046-1059.
54. Skeen, S., Laurenzi, C. A., Gordon, S. L., du Toit, S., Tomlinson, M., Dua, T., et al., *Adolescent Mental Health Program Components and Behavior Risk Reduction: A Meta-analysis.* Pediatrics, 2019. **144**(2).
55. Jensen-Doss, A. and Weisz, J. R., *Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics.* J Consult Clin Psychol, 2008. **76**(5): p. 711-22.
56. Yates, B. T. and Taub, J., *Assessing the costs, benefits, cost-effectiveness, and cost-benefit of psychological assessment: we should, we can, and here's how.* Psychol Assess, 2003. **15**(4): p. 478-95.
57. Sharp, L., Pantell, R. H., Murphy, L. O., and Lewis, C. C., *Psychosocial Problems During Child Health Supervision Visits: Eliciting,Then What?* Pediatrics, 1992. **89**(4): p. 619-623.
58. Westin, A. M., Barksdale, C. L., and Stephan, S. H., *The effect of waiting time on youth engagement to evidence based treatments.* Community Ment Health J, 2014. **50**(2): p. 221-8.
59. Fickel, J. J., Parker, L. E., Yano, E. M., and Kirchner, J. E., *Primary care - mental health collaboration: an example of assessing usual practice and potential barriers.* J Interprof Care, 2007. **21**(2): p. 207-16.
60. Vallance, A. K., Kramer, T., Churchill, D., and Garralda, M. E., *Managing child and adolescent mental health problems in primary care: taking the leap from knowledge to practice.* Prim Health Care Res Dev, 2011. **12**(4): p. 301-9.
61. Versi, E., *"Gold standard" is an appropriate term.* BMJ, 1992. **305**(6846): p. 187.
62. Claassen, J. A. H. R., *The gold standard: not a golden standard.* BMJ, 2005. **330**(1121).
63. Lewis, F. I. and Torgerson, P. R., *A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic.* Emerg Themes Epidemiol, 2012. **9**(1): p. 9.
64. Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., et al., *The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders?* Clin Psychol Sci, 2014. **2**(2): p. 119-137.
65. Krueger, R. F. and Eaton, N. R., *Transdiagnostic factors of mental disorders.* World Psychiatry, 2015. **14**(1): p. 27-9.
66. Johnstone, L. B., M. with Cromby, J., Dillon, J., Harper, D., Kinderman, P., Longden, E., Pilgrim, D. & Read, J. , *The Power Threat Meaning Framework Towards the identification of patterns in emotional distress, unusual experiences and troubled or troubling behaviour, as an alternative to functional psychiatric diagnosis*. 2018: Leicester: British Psychological Society.
67. Johnstone, L. and Boyle, M., *The Power Threat Meaning Framework: An Alternative Nondiagnostic Conceptual System.* Journal of Humanistic Psychology, 2018.

68. Barnes, M. C., Kessler, D., Archer, C., and Wiles, N., *Prioritising physical and psychological symptoms: what are the barriers and facilitators to the discussion of anxiety in the primary care consultation?* BMC Fam Pract, 2019. **20**(1): p. 106.

Appendices

Supplementary material belonging to Study 1

Translation of the vignettes

The vignettes were created, pilot-tested and surveyed in the Dutch language. Since only Dutch speaking participants were surveyed, the vignette script has not been translated back and forth, and thus, differences in connotation of the translated symptom expressions and idioms exist.

Vignette 1-Somatic: Rene is the first of two children, ten years old and consulting you because of recurring earaches. You know Rene as a shy child who cooperates well. While you are examining the child's ears, the mother mentions that although Rene looks timid, Rene can certainly be temperamental, usually when they are in a hurry. When you enquire further, it turns out that the mother means that although Rene can be so calm and sit so quietly, Rene can also get pretty upset. About three times a week, or so. But "Luckily Rene also has many strong points". Even when Rene has not slept well because of the earaches, in the morning Rene goes to school as usual. Rene doesn't want to miss any school and wants to finish all the homework properly. Rene's study skills are good. The teacher thinks Rene is a smart child. She does say, however, that "Rene can respond rather impatiently if Rene does not have a handle on certain situations". In addition, tiredness and overweight are also issues. Rene apparently did not have a healthy diet, but now they are watching what Rene eats.

Vignette 2-Behavioural: Finn is eleven years old and received mental healthcare as an eight year old. This was after their home was broken into. Finn has a fear of pain, small spaces, and burglars. Finn's mother confirms that Finn is still troubled by this, has concentration problems, and she adds that dyslexia was also noted at that time. This has an impact on Finn's performance at school. When you ask how things are going at home, Finn's mother responds that Finn "is impulsive, pushes the boundaries, can be rude, and goes on and on about things". She says that Finn fell down the stairs at the age of two and she wonders whether this might have influenced the current behaviour. It appears that the parents often have rows and disagreements. When you ask Finn how Finn usually feels, the answer is "cheeky and cool". Further, there are issues with sleeping; Finn has difficulty letting go. With the exception of minor respiratory symptoms there are no other somatic issues.

Vignette 3-Mood: Alin is twelve years old, is starting 8th grade, and lives with parents and twin brothers aged 14 who are known to have bronchial hyper responsiveness. Except for some bowel problems at the age of four, Alin is not known to you with other health issues. Alin's mother, who attends the consultation with Alin, tells you she has become increasingly worried over the past few years. Alin has difficulties concentrating at school, has few friends, and can be difficult at home. Alin expects everything to be worse than it actually is and is apprehensive – about all kinds of things. Alin increasingly prefers being alone. If there are tensions, Alin seems unreachable. Alin's mother wonders whether these problems will pass. She says she has never experienced problems of this kind with her other two children. Alin seems not to let on to others that things aren't going well and is convinced that there are a lot of things Alin is not good at. When you ask, Alin does report feeling tired, restless, and unhappy.

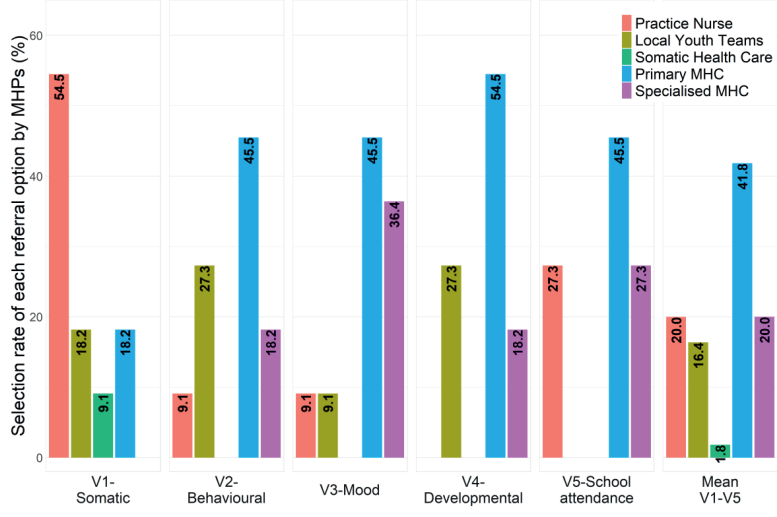
Vignette 4- Developmental: Lux is ten years old. Lux's mother turns to you for help because of increasing frustrations at home. She tells you that Lux has difficulty sharing with younger brother and seems to be controlling in friendships. With friends too this has sometimes led to problems interacting. Nevertheless, mother says that Lux is also concerned about little brother. Lux asks the parents questions about the brother's and other relation's health. The mother sees Lux as a child of a sensitive nature and with a strong will. Lux also had difficulty parting from parents when Lux went to

kindergarten. In the interaction with you, Lux initially comes across as rather passive, yet is able to develop a relation. You observe a somewhat restless, agitated attitude. When you ask how things are currently going at school, the mother says that Lux receives educational assistance for mild learning problems. Lux gets very stressed when things do not succeed. Lux tries to work neatly and carefully at school, though Lux is clearly easily distracted.

Vignette 5-Absenteeism: Jody is eleven years old and comes for a consultation because increasingly often, Jody does not manage to do a full day at school. Jody's parents are worried about this behaviour. The parents say that in recent months Jody has been crying a lot and has not been eating well. According to the teacher Jody is increasingly withdrawn and often seems not to finish the homework, but the parents say that Jody has never liked being the centre of attention. Otherwise there are no problems at school. They can't point to a specific incident that might have triggered these changes. Although two years ago Jody's family moved from the north of the country rather suddenly. The mother wonders whether Jody may have experienced this as traumatic and may still be troubled by it. Around that time Jody once fainted at home. Further there was nothing worrying up until now. Jody has always been a child who does well and wants to do well. It took some time after the relocation, but Jody is now best friends with two other children at school.

Demographical Characteristics of the Mental Health Professionals (MHPs)

Although all clinical psychologist and psychiatrist undergo a general training in the Netherlands, we asked MHPs what department they work for in order to have some insight in what problems they are particularly concerned with at the moment. A participation similar to what we see generally in psychiatry was found with three MHPs reporting affective disorders as their main focus, one developmental disorders, four comorbidity between psychiatric and somatic disorders, and one MHP employed at the department focusing on "comorbidity and complex problems within the primary support group". One MHP was affiliated to both latter departments. One MHP reported to have between 6-9 years of experience in this profession. Two MHPs reported to have 10 to 14 years of experience, four MHPs reported to have 15 to 19 years of experience, and another four to have more than 20 years of experience. All MHP in this validation sample were women. The gender distribution in the national MHP population is reported to be 5:1.



Supplementary Figure 1.MHPs' responses for the referral of each vignette

Supplementary Table 1. MHPs' selection rate of each disorder group per vignette

	V1-Somatic		V2-Behavioural		V3-Mood		V4-Developmental		V5-School attendance	
	First	Second	First	Second	First	Second	First	Second	First	Second
Anxiety	9.1	45.5	36.4	10.0	-	18.2	9.1	30.0	27.3	18.2
Trauma	-	-	-	20.0	-	-	-	-	-	63.6
Mood	9.1	-	-	-	63.6	18.2	-	10.0	63.6	9.1
Somatic	63.6	9.1	-	-	9.1	9.1	-	-	-	-
Eating	-	27.3	-	-	-	-	-	-	9.1	9.1
Autism	18.2	-	-	-	27.3	18.2	18.2	-	-	-
Attention-hyperactivity	-	-	27.3	10.0		9.1	54.5	30.0	-	-
Difficult behaviour	-	9.1	36.4	50.0	-	-	18.2	-	-	-
Typical development	-	-	-	-	-	-	-	-	-	-
No second complaint		9.1		10.0		27.3		30.0		-

Mental health professionals' selection rate of each disorder group per identification question of each vignette, shown as percentages. Responses were obtained from eleven MHPs on each question, except for ten responses on the second identification question of V2-Behavioural and V4-Developmental.

Supplementary material belonging to section "Results"

Tests of association between sample characteristics and the outcome measures

GPs' gender and years of experience did not show a relation with recognition and referral of anxiety. The tests were carried initially with the same number of levels as resulting from the survey (six levels for years of experience: 0-2, 3-5, 6-9, 10-14, 15-19, >20; six levels for recognition: 0 to 5). Since multiple cells had counts smaller than five analyses were recarried and depicted below with a median split for both variables. The median of the years of experience variable was >20 years of experience, and median recognition of anxiety was one; resulting in a binary median split. Analysis with median split did not change statistical significance of the results. Depicted below are the results of the tests for independence of gender and experience with (I) whether anxiety was recognized in any of the vignettes, (II) the referral options GPs chose in each of the vignettes, and (III) their explicit referral preferences for each of the disorder groups.

(I) GPs recognition of anxiety the five vignettes:
Gender: $\chi 2_{(1, n=211)} = 1.05$, $p = .31$; Experience: $\chi 2_{(1, n=228)} = .675$, $p = .411$.

(II) GPs answers for the referral of each vignette:
V_1 gender: $\chi 2_{(4, n=196)} = 2.25$, $p = .69$; experience: $\chi 2_{(4, n=212)} = 0.56$, $p = .97$
V_2 gender: $\chi 2_{(4, n=202)} = 2.63$, $p = .62$; experience: $\chi 2_{(4, n=219)} = 8.82$, $p = .07$
V_3 gender: $\chi 2_{(4, n=206)} = 2.40$ $p = .49$; experience: $\chi 2_{(4, n=223)} = 1.29$, $p = .86$
V_4 gender: $\chi 2_{(4, n=198)} = 2.17$, $p = .71$; experience: $\chi 2_{(4, n=214)} = 3.19$, $p = .53$
V_5 gender: $\chi 2_{(4, n=194)} = 7.85$, $p = .01$; experience: $\chi 2_{(4, n=210)} = 1.72$, $p = .79$

(III) GPs' referral preferences when they suspect:
Anxiety: gender: $\chi 2_{(4, n=206)} = 8.13$, $p = .09$, experience: $\chi 2_{(4, n=223)} = 7.32$, $p = .12$
Trauma: gender: $\chi 2_{(4, n=200)} = 1.18$, $p = .88$, experience: $\chi 2_{(4, n=216)} = 10.69$, $p = .03$
Mood: gender: $\chi 2_{(4, n=202)} = 0.92$, $p = .92$, experience: $\chi 2_{(4, n=219)} = 1.11$, $p = .89$
Somatic: gender: $\chi 2_{(4, n=194)} = 3.66$, $p = .45$, experience: $\chi 2_{(4, n=211)} = 3.79$, $p = .44$
Eating problems: gender: $\chi 2_{(4, n=202)} = 5.79$, $p = .22$, experience: $\chi 2_{(4, n=218)} = 5.18$, $p = .27$
Autism: gender: $\chi 2_{(4, n=197)} = 8.76$, $p = .07$, experience: $\chi 2_{(4, n=213)} = 13.58$, $p = .01$
Attention-hyperactivity: gender: $\chi 2_{(4, n=198)} = 7.67$, $p = .10$, experience: $\chi 2_{(4, n=213)} = 2.80$, $p = .59$
Difficult behaviour: gender: $\chi 2_{(4, n=198)} = 2.98$, $p = .56$, experience: $\chi 2_{(4, n=215)} = 3.69$, $p = .45$

Supplementary Table 2. Differences between GPs' and MHPs' recognition of anxiety in each of the five vignettes

	GPs' recognition	MHPs' recognition	Chi-square df=1	*P-value*
V1-Somatic	13 (5.7%)	6 (54.5%)	34.21	< .0001
V2-Behavioural	76 (33.8%)	5 (45.5%)	0.63	.426
V3-Mood	31 (13.7%)	2 (18.2%)	0.18	.653
V4-Developmental	20 (9.0%)	4 (36.4%)	8.49	.018
V5-School Attendance	27 (11.9%)	5 (45.5%)	10.08	.001

Selection frequency (%) of anxiety over the two identification questions shown in percentages per vignette. Excluding cases with a missing response on both identification questions of a vignette resulted in the following sample sizes for the GPs: V1-Somatic= 228, V2-Behavioural= 225, V3-Mood = 227, V4-Developmental= 222, V5-School Attendance= 226. Sample size for the MHPs=11. Fisher exact values are shared for V3-Mood and V4-Developmental as a result of small cell sizes (n<5).

Differences in the selection rate of the disorder groups

With a total of 229 respondents and 17 times a missing response on both identification questions of a vignette a total of 1128 possibilities (229*5-17) to label a vignette with each disorder group emerged for GPs. The total number of times MHPs could opt for each disorder group was 55 (11*5 vignettes). To investigate the recognition of each disorder group in the mixed vignettes within and between both groups of professionals, these totals (1128 and 55) were set against the number of times each disorder group was chosen. Given the small sample size and the multiple comparisons made, Bonferroni corrections were applied and the p-value was set at 0.05/24= .002. Analysis revealed that the GPs recognized anxiety less often than behavioural, mood, developmental and trauma related problems. Analysis of MHPs responses revealed that they recognized most frequently anxiety and mood problems. Comparison of GPs' and MHPs' recognition rate of each disorder group revealed that they differed only regarding anxiety (supplementary Table 3).

Supplementary Table 3. Differences between GPs and MHPs selection frequency of the disorder groups

	GPs' recognition	MHPs' recognition	OR (95% CI)	Chi-square value df=1	*P-value*
Anxiety	167 (14.8%)	22 (40.0%)	0.26 (0.15 to 0.46)	24.80	<.000001
Trauma	238 (21.1%)	9 (16.4%)	1.37 (0.66 to 2.83)	0.71	.3988
Mood	290 (25.7%)	19 (34.5%)	0.66 (0.37 to 1.16)	2.12	.1452
Somatic	87 (7.7%)	10 (18.2%)	0.38 (0.18 to 0.77)	7.64	.0057
Eating	60 (5.3%)	5 (9.1%)	0.56 (0.22 to 1.46)	1.44	.2307
Autism	253 (22.4%)	9 (16.4%)	1.35 (0.65 to 2.79)	1.119	.2901
Attention-hyperactivity	279 (24.7%)	14 (25.5%)	0.96 (0.52 to 1.79)	0.015	.9038
Difficult behaviour	332 (29.4%)	12(21.8%)	1.49 (0.78 to 2.87)	1.47	.2247
Typical development	177 (15.7%)	0 (0%)			
No second complaint group	184 (16.3%)	8 (14.5%)			

Supplementary Table 4. Referral of the vignettes by GPs who recognized anxiety and who did not

		Practice Nurse	Local Youth Teams	Somatic Healthcare/ Hospital	Primary Mental Healthcare	Specialised Mental Healthcare		Referral to mental health-care OR (95% CI), *P-value*
V_1	Selected	9 (69.2%)	3 (23.1%)	1 (7.7%)	0 (0.0%)	0 (0.0%)	**V_1**	0.03 (0.01 to 0.07), < .0001
	Not selected	117 (58.5%)	62 (31.0%)	13 (6.5%)	3 (1.5%)	5 (2.5%)		
V_2	Selected	13 (17.3%)	17 (22.7%)	29 (38.7%)	16 (21.3%)	0 (0%)	**V_2**	10.12 (4.53 to 22.59), < .0001
	Not selected	31 (21.8%)	37 (26.1%)	37 (26.1%)	36 (25.4%)	1 (0.7%)		
V_3	Selected	12 (40.0%)	4 (13.3%)	8 (26.7%)	6 (20.0%)	0 (0.0%)	**V_3**	11.37 (5.18 to 24.94), < .0001
	Not selected	48 (24.9%)	24 (12.4%)	65 (33.7%)	54 (28.0%)	2 (1.0%)		
V_4	Selected	8 (42.1%)	4 (21.1%)	5 (26.3%)	2 (10.5%)	0 (0.0%)	**V_4**	3.95 (1.73 to 9.02), = .001
	Not selected	67 (35.1%)	53 (27.7%)	46 (24.1%)	23 (12.0%)	2 (1.0%)		
V_5	Selected	14 (58.3%)	5 (20.8%)	3 (12.5%)	1 (4.2%)	1 (4.2%)	**V_5**	2.35 (0.98 to 5.63),= .055
	Not selected	105 (56.8%)	42 (22.7%)	23 (12.4%)	4 (2.2%)	11 (5.9%)		
Mean V_1-V_5	Selected	45.4%	20.2%	22.4%	11.2%	0.8%	**R**	0.70 (0.42 to 1.18), = .185
	Not selected	39.4%	24.0%	20.6%	13.8%	2.3%		

Selection frequency (%) of each referral option per vignette partitioned by general practitioners who selected anxiety and who did not. OR= odds ratio as obtained from the logistic multilevel analysis with the outcome measure whether a referral to mental healthcare was made, and the predictors: vignette (V_1 to V_5) and whether anxiety was recognized (R).

Supplementary Table 5. Results of the multilevel analysis on referral to mental healthcare

Mental health disorder	OR	95% CI	*P-value*
Anxiety	1.86	1.38 to 2.50	<.0001
Trauma	2.46	1.57 to 3.85	<.0001
Mood	1.03	0.69 to 1.55	.881
Somatic	0.27	0.18 to 0.40	<.0001
Eating	2.00	1.30 to 3.08	<.001
Autism	0.42	0.28 to 0.62	<.0001
Attention-hyperactivity	1.57	1.03 to 2.39	<.037
Difficult behaviour	0.07	0.04 to 0.11	<.0001

Supplementary Table 5 depicts the odds ratios (ORs) of referral to mental healthcare (yes/no) for each mental health disorder separately as reported by GPs when asked how they generally tend to manage children they suspect to have a mental health disorders. GPs responses for anxiety were included as the baseline. GPs chose between five options, which were combined into yes: primary mental healthcare and specialised mental healthcare, versus no: watchful waiting, practice nurse and local youth teams.

Supplementary material belonging to Study 2

Supplementary Table 1. Origin of the referral letters N=1259

	n (%)
General Practitioner (GP)	689 (54.7)
GP and another referrer	34 (2.7)
Specialists	
Psychiatry	61 (4.8)
Pediatrics	172 (13.7)
Rehabilitation doctors	13 (1.0)
Others	30 (2.4)
Local youth teams	
Youth and family centres	217 (17.2)
Municipal Health services	48 (3.8)
Juvenile probation officer	29 (2.3)

In the Netherlands, a formal referral to child and adolescent psychiatry proceeds either via general practice, specialized health centers (hospitals), the local youth welfare offices, or via youth protection boards. A total of 723 (57.4%) RLs were from general practice, and 34 of these cases had a RL from a GP and another referrer. For these cases we extracted only information from the RL originating from general practice.

Supplementary Table 2. Chance corrected agreement

	κ (95% CI)
Anxiety disorders	.81 (.73 - .86)
Depressive disorder	.82 (.71 - .94)
PTSD	.77 (.57 - .96)
Eating disorders	.87(.69 - 1.00)
ASD	.90 (.82 - .98)
ADHD	.90 (.83 - .97)
Behaviour	.77 (.63 - .66)

Chance corrected agreement, Kappa values, per disorder group computed over a random selection of 150 RLs that were coded by the author who coded all RLs and the three second coders who each coded a set of 50 letters.

Supplementary Table 4. How often and which ICPC-codes were written in referral letters?

	First	Second	Third	Fourth	Fifth
A Unspecified	23 (3.2)	20 (2.8)	9 (1.2)	6 (0.8)	4 (0.6)
B Blood and Immune mechanism	1 (0.1)	1 (0.1)	-	-	1 (0.1)
D Digestive	9 (1.2)	7 (1.0)	7 (1)	-	2 (0.3)
F Eye	6 (0.8)	3 (0.4)	2 (0.3%	3 (0.4)	-
H Ear	5 (0.7)	3 (0.4)	2 (0.3)	4 (0.6)	1 (0.1)
K Cardiovascular	4 (0.5)	5 (0.7)	1 (0.1)	1 (0.1%	-
L Musculoskeletal	8 (1.1)	13 (1.8)	8 (1.1)	1 (0.1)	4 (0.6)
N Neurological	10 (1.4)	5 (0.7)	5 (0.7)	3 (0.4)	-
P Psychological	176 (24.3)	65 (9.0)	29 (4.0)	11 (1.5)	5 (0.7)
R Respiratory	34 (4.7)	29 (4.0)	20 (2.8)	13 (1.8)	2 (0.3)
S Skin	20 (2.8)	21 (2.9)	11 (1.5)	5 (0.7)	5 (0.7)
T Endocrine/ metabolic	16 (2.2)	9 (1.2)	3 (0.4)	1 (0.1)	1 (0.1)
U/W Urological/ Pregnancy	2 (0.2)	2 (0.3)	7 (1.0)	-	1 (0.1)
X/Y Female/ Male genital	-	5 (0.7)	2 (0.3)	2 (0.3)	1 (0.1)
Z Social problems	17 (2.4)	6 (0.8)	2 (0.3)	1 (0.1)	2 (0.3)
No ICPC-code written in RL	392 (54.2)	529 (73.2)	617 (85.3)	672 (92.9)	695 (96.1)

Table 4 depicts the number (%) of ICPC-codes written in RLs. Frequently issued specific codes by the referrer were A12-Allergy (n=18), R96-Asthma (n=46), and S87-Eczema (n=37)[48]. Frequencies are not depicted on the level of the specific codes as a result of low frequencies and the differences we observed between these registered ICPC-codes and their accompanying short textual description (also written in RLs). Latter suggests probable unreliable use of the ICPC-code at that level (as discussed in the discussion section in the main manuscript).

Supplementary Table 5. Odds of classification per disorder group

		OR	95% CI
Anxiety disorders	RL	5.93	3.73 – 9.43
	RL+	4.76	2.90 – 7.83
	Age	1.15	1.07 – 1.24
	Gender	1.23	0.76 – 1.97
	History	1.28	0.72 – 2.26
	CGAS	1.00	0.97 – 1.04
Depression	RL	10.89	6.73 – 17.62
	RL+	4.79	2.83 – 8.10
	Age	1.42	1.26– 1.59
	Gender	1.70	0.98 – 2.96
	History	1.07	0.56 – 2.03
	CGAS	0.98	0.95 – 1.02
PTSD	RL	29.79	11.58 – 76.63
	RL+	45.47	15.31 – 135.06
	Age	1.10	0.95 – 1.28
	Gender	2.63	0.89 –7.78
	History	2.41	0.49 – 11.73
	CGAS	0.93	0.86 – 1.00
Eating disorders	RL	808. 36	170.43 – 3834.19
	RL+	681.78	95.95 – 4844.36
	Age	1.55	0.97 – 2.50
	Gender	3.10	0.26 – 36.67
	History	1.10	0.18 – 6.84
	CGAS	0.90	0.81 – 1.01
Autism spectrum disorders	RL	5.06	3.57 – 7.16
	RL+	5.17	3.48– 7.68
	Age	0.94	0.89 – 0.99
	Gender	0.47	0.31 – 0.71
	History	1.46	0.94 – 2.25
	CGAS	0.93	0.91 – 0.96
ADHD	RL	6.11	4.36 – 8.56
	RL+	7.39	5.09 – 10.74
	Age	0.91	0.86 – 0.96
	Gender	0.59	0.40 – 0.88
	History	0.89	0.58 – 1.35
	CGAS	1.00	0.97 – 1.03
Behavioural disorders	RL	6.02	3.11 – 11.66
	RL+	7.14	3.45 – 14.77
	Age	1.04	0.94 – 1.16
	Gender	1.08	0.54 – 2.17
	History	1.86	0.74 – 4.68
	CGAS	0.95	0.90 – 0.99

Values Table 5 depict the odds ratios (OR) resulting from logistic regression analyses with disorder group as outcome. PTSD= Post traumatic stress disorder. ADHD= attention-deficit hyperactivity disorders. Each upper row (RL) depicts the univariate analysis with only referral letter as predictor. Second to fifth rows (RL+) depict odds corrected for the main effects of age, gender, treatment history, and CGAS score. Age and CGAS were included as continuous variables. The reference for gender are boys. Psychiatric treatment history is included with '0 no treatment history' being the reference. In a third block the interaction terms RL*age, RL*gender, RL*history and RL*CGAS were added to test for the possibility that some disorder groups are mentioned more often in RLs depending on these factors. No significant interaction effects were found, except for an indication of ADHD*age (OR=1.14, 95% CI 1.03-1.27, *p*=.026).

Supplementary Table 6. Accuracy metrics for the most prevalent psychiatric disorder groups

		Anxiety disorders n=105	Mood disorders n=92	PTSD n=21	Eating disorders n=27	ASD n=214	ADHD n=243	Behavioural Disorders n=43	Personality disorder n=34	Psychosomatic disorders n=17	Re. attachm. & Disin. s. eng. d. n=16
Anxiety ICPC n=111	TP (Se, PPV)	**44 (41.9, 39.6)**	17 (18.5, 15.3)	3 (14.3, 2.7)	3 (11.1, 2.7)	28 (13.1, 25.2)	28 (11.5, 25.2)	5 (11.6, 4.5)	7 (20.6, 6.3)	4 (23.5, 3.6)	3 (18.8, 2.7)
	st.adj.res.	**8.2**	0.9	-0.1	-0.6	-1.1	-2.0	-0.7	0.9	0.9	0.4
	TN (Sp)	**551 (89.2)**	537 (85.1)	594 (84.6)	588 (84.5)	426 (83.7)	397 (82.7)	574 (84.4)	585 (84.9)	599 (84.8)	599 (84.7)
Mood ICPC n=148	TP (Se, PPV)	**36 (34.3, 24.3)**	**59 (64.1, 39.9)**	6 (28.6, 4.1)	7 (25.9, 4.7)	35 (16.4, 23.6)	30 (12.3, 20.3)	5 (11.6, 3.4)	**15 (44.1, 10.1)**	7 (41.2, 4.7)	2 (12.5, 1.4)
	st.adj.res.	**3.8**	**11.1**	0.9	0.7	-1.8	-3.9	-1.5	**3.5**	2.1	-0.8
	TN (Sp)	**506 (81.9)**	**542 (85.9)**	560 (79.8)	555 (79.7)	396 (77.8)	362 (75.4)	537 (79.0)	**556 (80.7)**	565 (80.0)	561 (79.3)
Trauma ICPC n=36	TP (Se, PPV)	7 (6.7, 19.4)	2 (2.2, 5.6)	**11 (52.4, 30.6)**	1 (3.7, 2.8)	3 (1.4, 8.3)	8 (3.3, 22.2)	5 (11.6, 13.9)	3 (8.8, 8.3)	1 (5.9, 2.8)	1 (6.3, 2.8)
	st.adj.res.	0.9	-1.3	**10.1**	-0.3	-2.9	-1.5	2.1	1.1	0.2	0.2
	TN (Sp)	589 (95.3)	597 (94.6)	**677 (96.4)**	661 (95.0)	476 (93.5)	452 (94.2)	649 (95.4)	656 (95.2)	671 (95.0)	672 (95.0)
Eating ICPC n=37	TP (Se, PPV)	7 (6.7, 18.9)	6 (6.5, 16.2)	1 (4.8, 2.7)	**25 (92.6, 67.6)**	4 (1.9, 10.8)	1 (0.4, 2.7)	1 (2.3, 2.7)	4 (11.8, 10.8)	0 (-)	0 (-)
	st.adj.res.	0.8	0.7	-0.1	**21.0**	-2.6	-4.1	-0.9	1.8	-1.0	-0.9
	TN (Sp)	589 (95.3)	601 (95.2)	667 (95.0)	**684 (98.3)**	476 (93.5)	444 (92.5)	644 (94.7)	656 (95.2)	669 (94.8)	670 (94.8)
Autism ICPC n=215	TP (Se, PPV)	18 (17.1, 8.4)	14 (15.2, 6.5)	1 (4.8, 0.5)	1 (3.7, 0.5)	**117 (54.7, 54.4)**	68 (28.0, 31.6)	9 (20.9, 4.2)	4 (11.8, 1.9)	1 (5.9, 0.5)	4 (25.0, 1.9)
	st.adj.res.	-3.1	-3.3	-2.5	-3.0	**9.5**	-0.7	-1.3	-2.3	-2.2	-0.4
	TN (Sp)	421 (68.1)	430 (68.1)	488 (69.5)	482 (69.3)	**411 (80.7)**	333 (69.4)	474 (69.7)	478 (69.4)	492 (69.7)	496 (70.2)
ADHD ICPC n=270	TP (Se, PPV)	20 (19.0, 7.4)	16 (17.4, 5.9)	6 (28.6, 2.2)	6 (22.2, 2.2)	70 (32.7, 25.9)	**158 (65.0, 58.5)**	11 (25.6, 4.1)	4 (11.8, 1.5)	1 (5.9, 0.4)	5 (31.3, 1.9)
	st.adj.res.	-4.2	-4.2	-0.8	-1.7	-1.7	**10.9**	-1.6	-3.2	-2.7	-0.5
	TN (Sp)	368 (59.5)	377 (59.7)	438 (62.4)	432 (62.1)	309 (60.7)	**368 (76.7)**	421 (61.9)	423 (61.4)	437 (61.9)	442 (62.5)
Behavi-oural ICPC n=203	TP (Se, PPV)	20 (19.0, 9.9)	11 (12.0, 5.4)	5 (23.8, 2.5)	1 (3.7, 0.5)	75 (35.0, 36.9)	64 (26.3, 31.5)	**29 (67.4, 14.3)**	3 (8.8, 1.5)	1 (5.9, 0.5)	**10 (62.5, 4.9)**
	st.adj.res.	-2.2	-3.7	-0.4	-2.9	2.7	-0.7	**5.9**	-2.6	-2.1	**3.1**
	TN (Sp)	435 (70.4)	439 (69.6)	504 (71.8)	494 (71.0)	381 (74.9)	341 (71.0)	**506 (74.4)**	489 (71.0)	504 (71.4)	**514 (72.7)**
Persona-lity ICPC n=52	TP (Se, PPV)	9 (8.6, 17.3)	**15 (16.3, 28.8)**	**5 (23.8, 9.6)**	1 (3.7, 1.9)	12 (5.6, 23.1)	16 (6.6, 30.8)	6 (14.0, 11.5)	**8 (23.5, 15.4)**	1 (5.9, 1.9)	**9 (56.3, 17.3)**
	st.adj.res.	0.6	**3.6**	**3.0**	-0.7	-1.1	-0.5	1.8	**3.8**	-0.2	**7.7**
	TN (Sp)	575 (93.0)	**594 (94.1)**	**655 (93.3)**	645 (92.7)	469 (92.1)	444 (92.5)	634 (93,2)	**645 (93.6)**	655 (92.8)	**664 (93.9)**
Somati-sation ICPC n=32	TP (Se, PPV)	8 (7.6, 25.0)	**11 (12.0, 34.4)**	0 (-)	0 (-)	8 (3.7, 25.0)	4 (1.6, 12.5)	0 (-)	1 (2.9, 3.1)	**5 (29.4, 15.6)**	0 (-)
	st.adj.res.	1.7	**3.8**	-1.0	-1.1	-0.6	-2.6	-1.5	-0.4	**5.1**	0.9
	TN (Sp)	594 (96.1)	**610 (96.7)**	670 (95.4)	664 (95.4)	485 (95.3)	452 (94.2)	648 (95.3)	658 (95.5)	**679 (96.2)**	675 (95.5)

Columns: frequency of diagnoses as classified in specialized mental healthcare. Rows: frequency of tentative diagnoses in RLs. TP= true positives, number of youth that had an indication of the diagnosis they are classified with in their RLs. Se= sensitivity, ratio TPs to the frequency of the diagnoses, in percentages. PPV= positive predictive value, ratio TPs to the frequency of diagnoses in RLs. St.adj.res.= standardized adjusted residual, computed as a measure of co-occurrence beyond chance. Depicts the discrepancy between observed and expected values and suggest statistical significance at the level of $p<.05$ when it exceeds $|1.96|$. Given the many comparisons made, we set a higher boundary of $|3.0|$.[1] TN= true negatives, number of youth without the classification and no indication in the RL. Sp= specificity, ratio TN to the whole sample without the disorder group, in percentages. ASD=autism spectrum disorders. ADHD= attention-deficit hyperactivity. Re. attachm. & Disin. s. eng. d.= reactive attachment & disinhibited social engagement disorder.

Results suggest that a quarter of children referred for mood problems were later classified with an anxiety disorder (24.3%, 36/105). The reverse association, i.e. referred with anxiety then classified with depression, was not foud. Nor was it accounted for by co-occurrence of depression in those classified with an anxiety disorder: after exclusion of the comorbid cases (n=16) the number of children with anxiety disorders that were referred for depressive disorders remained similar (28.1%, 25/89). A similar pattern was seen for those eventually diagnosed with behavioural disorders, as they were equally likely to be referred for suggested behavioural problems (14.3%), and or trauma (13.9%, 5/43). Again this association remained after exclusion of the cases with comorbid PTSD (16.0%, 4/25). Although high raw values were found for some other disorder groups, the frequencies were no more than what could be expected by chance.

Assuming that the reason of consult for Reactive attachment disorder and Disinhibited social engagement disorder might differ from those with PTSD, we included only PTSD from the Trauma and stressor-related disorders chapter in the tabulation with indications of trauma made in RLs. Adjustment disorders were not cross-tabulated as a result of their small sample size and since they can not be the only diagnosis in, or reason of referral to, specialized mental healthcare according to the Dutch legislation. For their relative severe nature, personality disorders were mentioned relative frequently in the RLs, but had a low sensitivity (23.5%). A suggestion of personality or attachment problems in RLs was significantly related to a classification of depression (28.8%) and PTSD (9.6%). This association decreased with about a third, to respectively 20.5% (9/44, st.adj.res.=2.1.) and 6.8% (3/44, st.adj.res.=2.0), when those with a co-occuring personality disorders were excluded. An indication of somatisation disorder in RLs was to some extent related to a later classification of somatisation disorder (PPV=15.6%). When next to somatisation disorder (which has a specific ICPC-code: P75), other bodily symptoms were also counted as an indication of somatic symptom and related disorders, both the sensitivity and PPV increased, to respectively 58.8% (10/17), and 25% (10/40). This (combining P75-somatisation disorder and indications of other bodily symptoms and physical complaints) did not change associations with other disorder groups.

Supplementary Table 7. Cross tabulation of reasons of referral and clinical status for the low prevalence disorder groups that were not included in the main manuscript

	PPV of the disorder specific labels	Sensitivity when all labels from the chapter are combined
Intellectual disability	1/8 (12.5)	16/21 (76.2)
Communication disorder	1/12 (8.3)	17/18 (94.4)
Motor disorders	5/ 14 (35.7)	11/14 (78.6)
Specific learning disorder	11/30 (36.7)	28/38 (73.7)
+ Other Neurodevelopmental Disorders + ASD + ADHD	322/435 (74.0)	
+ high IQ	323/439 (73.6)	
Separation anxiety disorder	-	4/8 (50.0)
Specific phobia	0/2 (0.0)	2/6 (33.30
Social anxiety disorder	1/7 (14.3)	10/16 (62.5)
Panic disorder	1/5 (20.0)	5/8 (62.5)
Agoraphobia	-	1/1 (100.0)
Generalized anxiety disorder	0/2 (0.0)	24/47 (51.1)
Anxiety disorder not otherwise specified	-	6/28 (21.4)
OCD	5/15 (33.3)	5/8 (62.5)

Separation anxiety disorder, agoraphobia, and anxiety disorder NOS do not have specific ICPC-codes. Hence no disorder specific PPVs' are depicted in these rows. Sensitivity is depicted on the level of the neurodevelopmental disorders chapter, the anxiety disorders chapter and for obsessive compulsive disorders (including trichotillomania n=2). Extant yet infrequent reasons of referral were emotion dysregulation (n=21), self-image (19) and game addiction (4).[2]

Supplementary Table 8. Extended table reasons of referral per disorder groups

	Anxiety disorders n=105	Mood disorders n=92	PTSD n=21	Eating disorders n=27	ASD n=214	ADHD n=243	Behavioral Disorders n=43
Study problems	11 (13.1%)	5 (6.0%)	1 (1.2%)	2 (2.4%)	32 (38.1%)	39 (46.4%)	7 (8.3%)
n=84 St. adj. res.	-0.4	-2.0	-1.0	-0.7	1.8	2.6	1.0
School attendance problems	12 (42.9%)	8 (28.6%)	0	0	7 (25.0%)	4 (14.3%)	0
n=28 St. adj. res.	4.3	2.6	-0.9	-1.1	-0.5	-2.2	-1.4
Perfectionism/ fear of failure	6 (30.0%)	3 (15.0%)	1 (5.0%)	1 (5.0%)	6 (30.0%)	4 (20.0%)	1 (5.0%)
n=20 St. adj. res.	2.0	0.3	0.6	0.3	0.0	-1.3	-0.2
High IQ	4 (25.0%)	2 (12.5%)	0	0	4 (25.0%)	4 (25.0%)	2 (12.5%)
n=16 St. adj. res.	1.2	0.0	0.7	-0.8	-0.4	-0.7	1.1
Intellectual disabilities	0	1 (12.5%)	1(12.5%)	0	3 (37.5%)	2 (25.0%)	0
n=8 St. adj. res.	-1.2	0.0	1.6	-0.6	0.5	-0.5	-0.7
Learning disorders	0	0	0	0	8 (26.7%)	16 (53.3%)	1 (3.3%)
n=30 St. adj. res.	-2.3	-2.1	-1.0	-1.1	-0.4	2.3	-0.6
Communication problems	1 (8.3%)	0 (0.0%)	0	0	5 (41.7%)	5 (41.7%)	0
n=12 St. adj. res.	-6	-1.3	-0.6	-0.7	0.9	0.6	-0.9
Somatic symptoms	8 (25.0%)	11 (34.4%)	0	0	8 (25.0%)	4 (12.5%)	0
n=32 St. adj. res.	1.7	3.8	-1.0	-1.1	-0.6	-2.6	-1.5
Headaches	0	2 (28.6%)	0	0	0	0	0
n=7	-1.1	1.3	-0.5	-0.5	-1.7	-1.9	-0.7
Pain-Fatigue	6 (35.3%)	6 (35.3%)	0	0	6 (35.3%)	3 (17.6%)	0
n=17	2.5	2.8	-0.7	-0.8	0.5	-1.4	-1.0
Stomach/ Bowel	2 (22.2%)	3 (33.3%)	0	0	2 (22.2%)	0	0
n=9	0.7	1.9	-0.5	-0.6	-0.5	-2.1	-0.8
Fainting/ Powerlessness	2 (66.7%)	1 (33.3%)	1 (33.3%)	1 (33.3%)	0	0	0
n=3	2.6	1.1	3.1	2.7	-1.1	-1.2	-0.4
Hyperventilation	1 (50.0%)	0	0	0	0	1 (50.0%)	0
n=2	1.4	-0.5	-0.2	-0.3	-0.9	0.5	-0.4
Problems Sleeping	4 (22.2%)	4 (22.2%)	2 (11.1%)	1 (5.6%)	2 (11.1%)	7 (38.9%)	0
n=18 St. adj. res.	0.9	1.2	2.1	0.4	-1.7	0.5	-1.1
Suicidal ideation	10 (18.9%)	23 (43.4%)	1 (1.9%)	0	14 (26.4%)	8 (15.1%)	2 (3.8%)
n=53 St. adj. res.	0.9	7.0	-0.5	-1.5	-0.5	-3.0	-0.7
Self Harm	7 (25.0%)	12 (42.9%)	3 (10.7%)	3 (10.7%)	8 (28.6%)	6 (21.4%)	1 (3.6%)
n=28 St. adj. res.	1.6	4.9	2.5	2.0	-0.1	-1.4	-0.5
(Sexual) Violence	1 (12.5%)	1 (12.5%)	5 (62.5%)	1 (12.5%)	0	1 (12.5%)	0
n=8 St. adj. res.	-2.0	0	10.1	1.3	-1.8	-1.3	-0.7
Problems with parents	15 (17.2%)	14 (16.1%)	8 (9.2%)	3 (3.4%)	16 (18.4%)	25 (28.7%)	11 (12.6%)
n=87 St. adj. res.	0.8	1.0	3.7	-0.2	-2.4	-1.3	2.8
Bullied-Social relatedness	5 (9.8%)	6 (11.8%)	1 (2.0%)	1 (2.0%)	20 (39.2%)	22 (43.1%)	1 (2.0%)
n=51 St. adj. res.	-1.0	-0.2	-4.0	-0.7	1.6	1.5	-1.2

Frequency (%) of the reasons of referral per disorder group, i.e. referred with the reason of referral in the row and classified with the disorder group in de column. Below each row percentage are standardized adjusted residual values depicted. A case could be referred for multiple reasons as well as be classified with multiple disorders.

References belonging to supplementary material of study 2

[1] Haberman, S. J., *Adjusted st. res. the Analysis of Frequency Data,*. 1974, Chicago: University of Chicago Press.

[2] Scottish Association for Mental Health and Information Services Division Scotland and NIHS, *Rejected Referrals Child and Adolescent Mental Health Services (CAMHS): A qualitative and quantitative audit*, Group Scotland, T. S., Edinburgh EH6 5NA PPDAS433246 (06/18), Editor. 2018, The Scottish Government.

Supplementary material belonging to Study 3

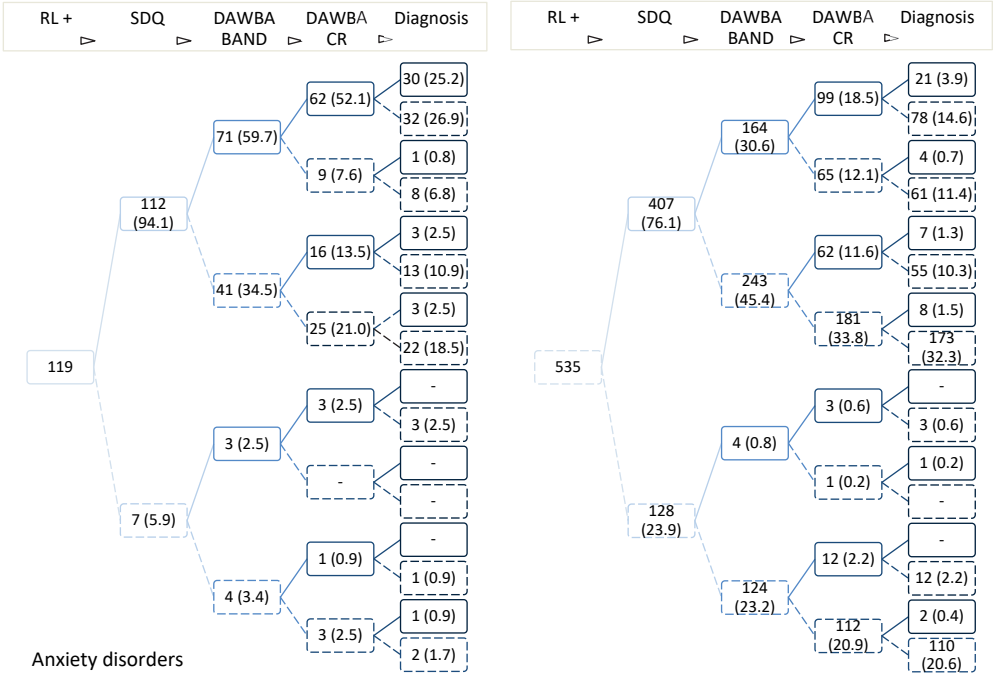
Supplementary table 1. Two by two cross-tabulation of the instruments per disorder group presenting positive predictive values

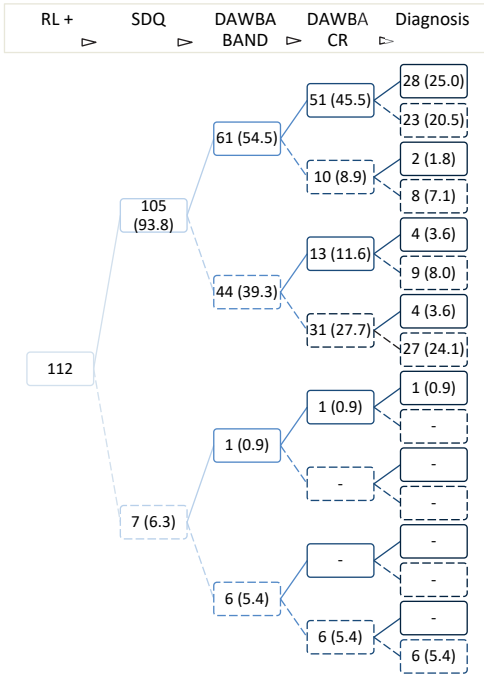
		Anxiety disorders		Depressive disorders		ASD		ADHD		Behaviour disorders	
		+	-	+	-	+	-	+	-	+	-
RL	+	38 (31.9)	81 (68.1)	39 (34.8)	73 (65.2)	108 (54.8)	89 (45.2)	114 (53.5)	99 (46.5)	26 (14.3)	156 (85.7)
	-	43 (8.0)	492 (92.0)	26 (4.8)	516 (95.2)	89 (19.8)	361 (80.2)	90 (20.5)	350 (79.5)	18 (3.8)	455 (96.2)
SDQ	+	77 (14.8)	442 (85.2)	62 (11.9)	457 (88.1)	140 (40.7)	204 (59.3)	181 (44.0)	230 (56.0)	38 (10.4)	328 (89.2)
	-	4 (3.0)	131 (97.0)	3 (2.2)	132 (97.8)	57 (18.8)	246 (81.2)	23(9.5)	219 (90.5)	6 (2.1)	283 (97.9)
Band	+	57 (23.6)	185 (76.4)	45 (32.4)	94 (67.6)	18 (78.3)	5 (21.7)	121 (60.8)	78 (39.2)	16 (6.6)	225 (93.4)
	-	24 (5.8)	388 (94.2)	20 (30.8)	495 (96.1)	179 (28.7)	445 (71.3)	83 (18.3)	371 (81.7)	28 (6.8)	384 (93.2)
CR	+	62 (24.2)	194 (75.8)	49 (32.0)	104 (68.0)	151 (49.5)	154 (50.5)	170 (51.8)	158 (48.2)	26 (11.5)	200 (88.5)
	-	19 (4.8)	379 (95.2)	16 (3.2)	485 (96.8)	46 (13.5)	296 (86.5)	34 (10.5)	291 (89.5)	18 (4.2)	411 (95.8)

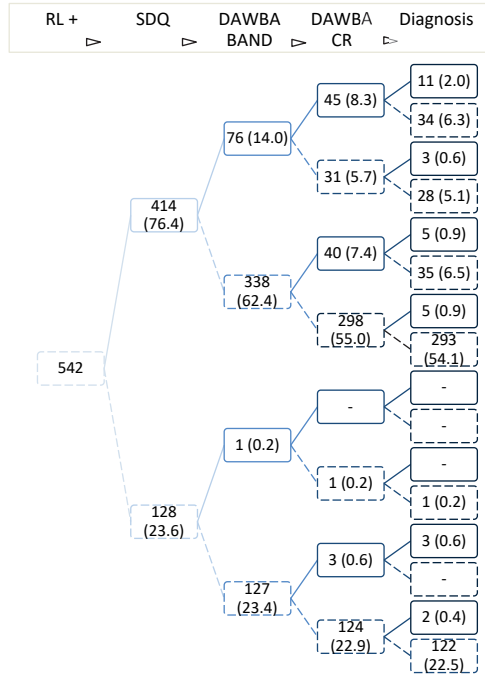
Frequency (%) of the positive and negative indications made per instruments and per disorder group, as a ratio of the total number of positive and negative indications made in the considering instrument. Number of diagnoses and sample size were as follows: anxiety disorders n=81 and N=654; depressive disorder n=65 and N=654, autism spectrum disorders (ASD) n=197 and N=647; attention-deficit hyperactivity disorder (ADHD) n=204 and N=653; behaviour disorders n=44 and N=655.

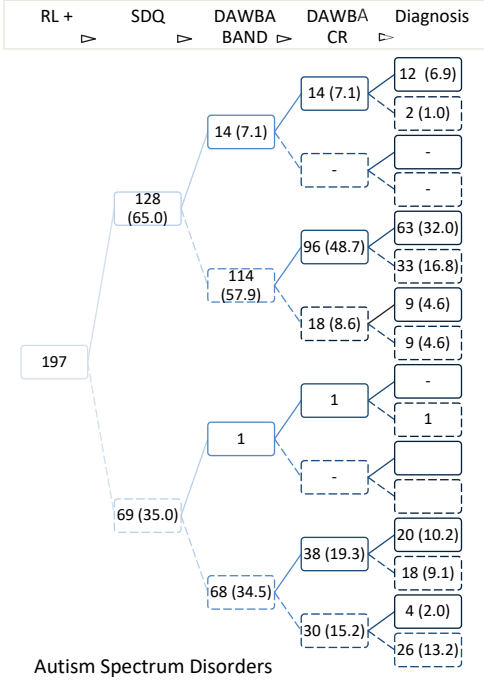
Supplementary Figure 1 Diagnostic trajectory

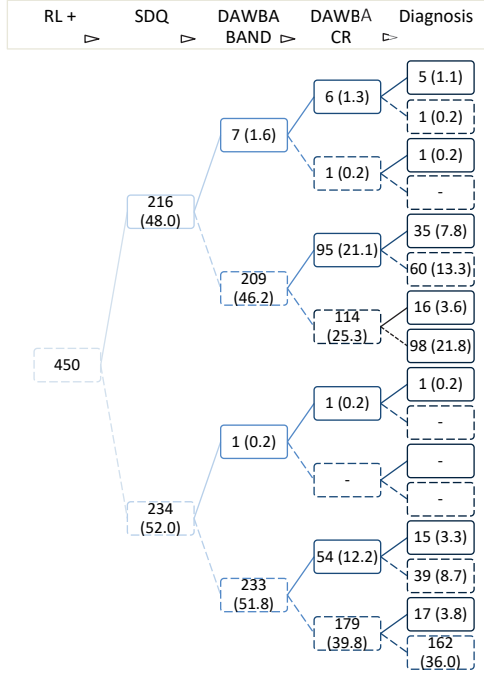
Number (%) of cases detected through RLs, SDQ, DAWBA band, and CR scores, per disorder group, as a ratio of the total frequency of the positive RLs (figure on the left) or negative RLs (right). The continuous lines present inflow, i.e. those that score positive on the concerning instrument. The dotted lines present outflow, i.e. those that score negative on the concerning instrument. Computed in the dataset with complete datapoints. Number of datapoints (N) for Anxiety: 654, Depression: 654, ASD:647, ADHD: 551, Behaviour disorders: 655.



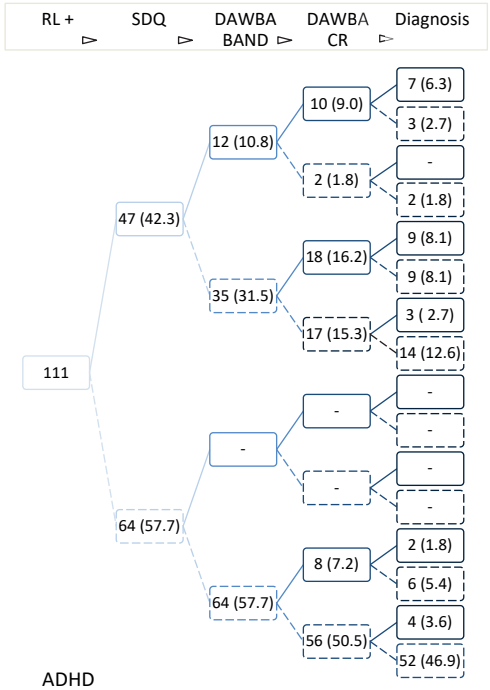




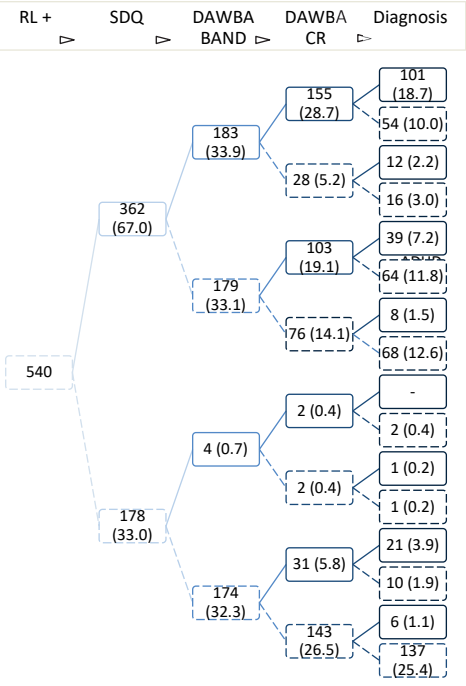


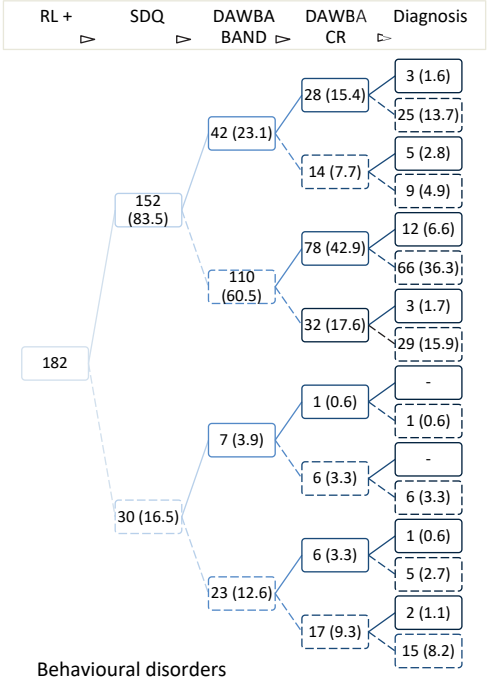


Autism Spectrum Disorders

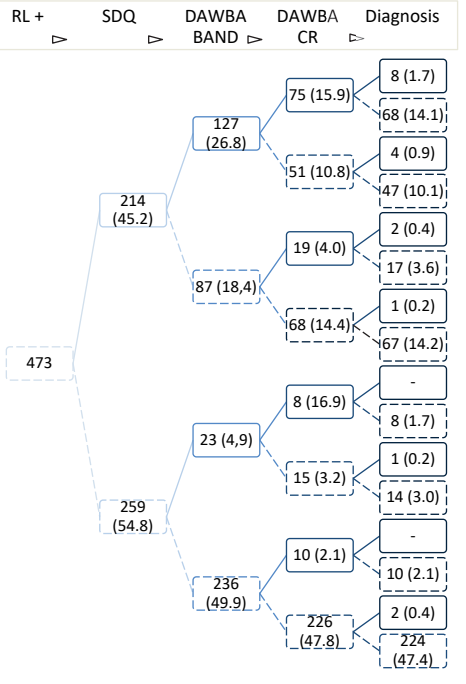


ADHD





Behavioural disorders



Nederlandstalige samenvatting

Tussen kijken en zien
Herkennen, verwijzen en evalueren van psychosociale problematiek bij jeugdigen in de eerstelijnszorg en de geestelijke gezondheidszorg

Als ouder en professional verzamel en evalueer je continu informatie over de jongeren die je ziet. Dit wordt idealiter vertaald in een hulpvraag in geval van aanhoudende opvallendheden rondom het functioneren en welzijn van een jongere. In werkelijkheid zijn er echter vele stappen en uitdagingen tussen het bestaan van een ondersteuningsbehoefte en het daadwerkelijk verkrijgen van adequate ondersteuning en zorg. Een eerste grote uitdaging is het herkennen en beoordelen van de klachten: het herkennen van de signalen en erkennen van de ervaringen alsook het evalueren van de sterke en zwakke kanten van de jeugdige en diens systeem. Uit eerder onderzoek weten we dat zowel ouders als professionals hier moeite mee hebben. In dit proefschrift presenteren we de resultaten van een reeks empirische studies waarin we ons hebben gericht op onbeantwoorde vragen rondom dit thema van herkennen, verwijzen en evalueren van psychosociale problematiek bij kinderen en jongeren. We hebben dit met name gedaan met een focus op angststoornissen, alhoewel we ook waarden presenteren voor andere veel voorkomende classificaties zoals gedragsproblemen of depressieve stoornissen. In de eerste studie hebben wij onderzocht of huisartsen in voldoende mate denken aan de mogelijkheid dat een jongere last heeft van angststoornissen wanneer zij op consult komen. Wij hebben naast vragen naar de aard van de klachten, aan huisartsen ook gevraagd waar de verschillende soorten problematiek het beste behandeld kunnen worden (hoofdstuk 2). In de volgende studie hebben we verwijsbrieven die naar de geestelijke gezondheidszorg (ggz) zijn verstuurd geanalyseerd om de vraag te beantwoorden wat de verwijsredenen van huisartsen zijn als ze eenmaal het besluit hebben genomen een jongere te verwijzen. In die studie hebben we ook gekwantificeerd hoe vaak andere meer contextuele redenen (zoals bijvoorbeeld problemen op school of thuis) werden benoemd in de verwijsbrieven (hoofdstuk 3). In de derde studie hebben we een sequentiële benadering voor diagnostiek tussen de eerste lijn en de ggz onderzocht. Hierbij was de vraag wat de toegevoegde waarde van de opeenvolging van onder andere verwijsredenen, een screeningsvragenlijst en een uitgebreide evaluatie

methode kan zijn in het voorspellen van de uiteindelijke diagnose (hoofdstuk 4). Hieronder volgt een samenvatting van elk hoofdstuk in dit proefschrift.

Hoofdstuk 1: In de algemene introductie omschrijven we de status quo van mentale gezondheidsproblemen onder jongeren, het zorg landschap in Nederland en de rol van professionals en familie in de herkenning en uiteindelijke behandeling van deze jongeren.

Epidemiologisch onderzoek, dus onderzoek naar de getallen van ziekten, suggereert dat ongeveer een op de vier jongeren bij uitvragen zou voldoen aan de diagnostische kenmerken van een psychiatrisch beeld. Studies onder volwassenen onderschrijven dat ruim driekwart van de volwassenen met klachten op dit vlak, deze al vóór de jongvolwassenheid ervoer. Het is dan ook niet vergezocht om te concluderen dat het bevorderen van het welzijn van jongeren meerdere generaties goed zal doen. Gelukkig geldt voor een deel van deze personen dat de klachten tijdelijk zijn en na verloop van tijd verdwijnen. Voor degenen die wel baat zouden kunnen hebben bij begeleiding of zorg, geldt helaas dat ongeveer twee derde geen passende zorg weet te vinden. In hun bekende filter-model omschrijven Goldberg en Huxley (1982) de verschillende niveaus van zorg en hoe deze als filters of knelpunten kunnen werken in de zoektocht naar gepaste zorg. Onderzoekers die dit model verder hebben uitgebreid en getoetst onderschrijven dat het verkrijgen van begeleiding of hulp voor jongeren een meerstappenprocedure is. Dit niet alleen door de hoeveelheid betrokkenen (zowel het jeugdige alsook diens systeem), maar ook door enerzijds de uitdagingen van het onderscheiden van normale variatie en ontwikkelingsuitdagingen van psychopathologie als anderzijds de ingewikkeldheid van het zorgsysteem. Het zorgsysteem bestaat uit een variëteit aan zorgpaden, welk het resultaat is van tientallen jaren aan hervormingen in de zorg om de toenemende kwantitatieve en kwalitatieve eisen bij te benen. Vanaf de jaren ’50 werd een onderscheid gemaakt tussen de zogenoemde eerstelijnszorg en tweedelijnszorg. Doel was om relatief minder complexe zorgvragen laagdrempelig en efficiënt te beantwoorden in de eerstelijnszorg en expertise doelgericht te benutten in de tweedelijnszorg. Meer recent en gericht op de zorg voor jeugdigen werd in Nederland in 2015 de Jeugdwet van kracht. Met de Jeugdwet werd bedoeld om de schotten te slechten tussen de verschillende niveaus van zorg en al het bestuurlijke verantwoordelijkheid te verplaatsen naar daar waar de leefwereld van het gezin zich

afspeelt: het lokale niveau van de gemeente in plaats van voorheen een centrale rol van het rijk of de regio. Hiermee is ook de specialistische jeugdhulp gedecentraliseerd; dat wil zeggen dat het onder de verantwoordelijkheid van de gemeente is komen te staan. De Jeugdwet was dus in leven geroepen om de toegenomen fragmentatie van de zorg te verminderen en om de verschillende ondersteuningsbehoeften van gezinnen op een geïntegreerde wijze te beantwoorden. Evaluaties van de Jeugdwet concluderen dat alhoewel men in sommige gemeenten goed op weg is naar voornoemd doel, er ook veel uitdagingen zijn in veel gemeenten. Kenmerkend aan deze gemeenten is dat er binnen een gemeente veel verschillende niet geïntegreerde werkwijzen bestaan en er sprake is van rolambiguïteit. Men concludeert dat het zaak is te faciliteren wat werkt op het niveau van de individuele zorg, nog los van hoe het zorgsysteem is georganiseerd. Gegeven de variëteit aan ondersteunings- en zorgbehoeften zullen er altijd verschillende specialisaties zijn en per definitie ook een keten van verschillende zorgleveranciers. Inherent hieraan zal communicatie altijd een uitdaging vormen. Om betekenisvol en samenhangend zorg te leveren moeten de sterke en zwakke kanten van jongeren en diens gezin expliciet gemaakt worden. Voor een meer uitgebreide omschrijving van wat bedoeld wordt met evaluatiemethoden (‘assessment’) verwijzen we u graag naar hoofdstuk 1 van dit proefschrift. We concluderen daarin dat zonder inzet van enige vorm van assessment, we in feite niet kunnen spreken van een doelgerichte behandeling. Immers, om gericht te beantwoorden aan de hupvraag is er altijd een bepaalde mate van genereren en doorcommuniceren van hypothesen nodig. In de volgende hoofdstukken gaan we in op de opzet en resultaten van studies waarin we een drietal hoofdvragen bijdragen rondom dit thema van betrouwbaar genereren en doorcommuniceren in en tussen de eerstelijnszorg en de specialistische ggz.

Hoofdstuk 2: Hier worden de resultaten gepresenteerd van een poll gehouden onder 229 huisartsen die hebben deelgenomen aan een landelijk nascholingscongres. Wij hebben aan hen casusomschrijvingen, oftewel gevalsbeschrijvingen, gepresenteerd van jeugdigen met een mogelijke angststoornis welke niet letterlijk als zodanig benoemd werd door de ouders of het kind. Deze casusomschrijvingen waren zodanig ontwikkeld dat er ruimvoldoende angstklachten genoemd werden. De hypothese was hierbij dat angststoornissen al in de eerste interpretatie van de klachten door de huisarts onvoldoende als zodanig herkend en

verwezen zullen worden. Wat nieuw is aan dit design met een mondelinge presentatie is dat het de mogelijkheid bood om zicht te krijgen op het eerste oordeel en interpretatie van de professional, terwijl factoren buiten de huisarts om (denk aan vermijding door het kind) kon worden uitgesloten. Naast de vraag wat de huisarts gepast vindt qua verwijzing van de casusomschrijvingen, hebben wij hen ook gevraagd wat zij als de meest gepaste behandelomgeving ervaren voor de verschillende vaak voorkomende klachtengroepen (zoals depressies, ADHD, angststoornissen e.d.). Doel was om hiermee zicht te krijgen op eventuele discrepanties tussen wat zij doen (bij de casusbeschrijvingen) en wat zij zeggen dat ze zouden doen (bij de vraag wat ze een toepasselijke behandelplek vinden). Dezelfde casusbeschrijvingen en vragen zijn ter vergelijking ook voorgelegd aan een groep professionals uit de ggz. Dit resulteerde in 1128 evaluaties van casusomschrijvingen door huisartsen en 45 door ggz professionals. We vonden dat huisartsen bijna vier keer zoveel kans hebben om angststoornissen niet te herkennen in vergelijking met ggz medewerkers (OR=0.26, 95% BI 0.15 - 0.46). Noemenswaardig hierbij is dat er veel variatie was binnen de groep van huisartsen; alhoewel enkele huisartsen zelfs in alle casusomschrijvingen aan mogelijke angstklachten dachten, herkende 44.1% geen keer angst als mogelijk onderliggende problematiek. Op de vraag waar jeugdigen met angstklachten behandeld zouden moeten worden reageerde 63.9% van de huisartsen met de ggz als een adequate behandelomgeving. In tegenstelling tot hun reactie op deze vraag waarin naar hun bewuste voorkeuren werd gevraagd, koos enkel 12.0% van de huisartsen die angststoornissen als een mogelijke klacht had herkend in een casusomschrijving voor een behandeling in de ggz voor desbetreffende casus. Deze bevindingen impliceren dat tijdens een consultsituatie waarin klachten ongestructureerd en niet voldoende expliciet, doch alom vertegenwoordigd, worden gepresenteerd, huisartsen al bij hun eerste evaluatie onvoldoende rekening houden met angst als onderliggend problematiek. Overigens, ook de ggz professionals herkenden angst onvoldoende, iets wat ook eerder in de literatuur is omschreven. De studieresultaten suggereren dat de bekendheid van professionals met de symptomatiek behorend bij angststoornissen vergroot moet worden, en vooral ook dat ze meer bekend zouden mogen raken met hoe vaak angststoornissen voorkomen (base-rate) in hun dagelijkse praktijk. Immers, huisartsen herkenden angststoornissen minder vaak dan hoe vaak ze voor angststoornissen hadden mogen kiezen als ze enkel af waren gegaan op de statistiek: hoe

vaak jeugdigen met angststoornissen gemiddeld verschijnen bij hen op de praktijk. Een belangrijke tekortkoming van de studie is dat het een momentopname toetst en geen recht doet aan de continuïteit van de zorg en diagnostiek. Tenslotte bediscussiëren we hoe het enerzijds niet gaat om hoe de huisarts verfijnde diagnostiek verricht (en geen onderscheid zou hoeven te maken tussen verschillende stoornisgroepen) en anderzijds dat het wel degelijk van belang is dat juist voor klachtengroepen die gekenmerkt worden door 'stille gedragingen' (anderen hebben relatief weinig last van en zicht op bijvoorbeeld een angstig persoon) deze gerichter gesignaleerd worden.

Hoofdstuk 3: In dit hoofdstuk delen we de resultaten van dossieronderzoek verricht met als doel het omschrijven van wat huisartsen schrijven op de verwijsbrieven van jeugdigen die behandeld worden in de ggz. Alhoewel verwijsbrieven een centrale rol spelen in de doorverwijsstromen in de zorg, was er eerder geen studie gepubliceerd naar wat er zoal staat in de verwijsbrieven naar de jeugd ggz en wat de verwijsredenen van jeugdigen zijn geweest die uiteindelijk zijn gediagnosticeerd met een variëteit aan klachten. Dit terwijl het versturen en ontvangen van verwijsbrieven onderdeel is van de dagelijkse klinische praktijk van de professional. Als een eerste inspectie van dit onderwerp hebben wij dit onderzoek gericht op wat de verwijsredenen zijn zoals genoemd in de verwijsbrieven. Wij hebben een totaal van 723 brieven geïnspecteerd die verstuurd waren vanuit de huisartsenzorg naar de ggz. De omschrijving van de verwijsredenen bestonden gemiddeld uit 59 woorden (SD=41, bereik 2 tot 246). In de meeste verwijsbrieven werden één (25.0%), twee (32.2%) of drie (24.8%) verschillende soorten symptomen of mogelijke diagnosen genoemd. In 5.0% van de verwijsbrieven werd gevraagd om voorrang en in 6.9% werd de ernst van de situatie benadrukt. Een vergelijking van het niveau van ervaren belemmeringen (uitgedrukt in zogenoemde CGAS scoren van 10 tot 100) zoals werd ingeschat door de clinicus die uiteindelijk een diagnose heeft gesteld in de ggz onder de groep jeugdigen met een brief waarin gevraagd werd om voorrang (M_{CGAS}=47.27, SD=8.12) en waarin de ernst van de situatie werd onderstreept (M_{CGAS}=48.83, SD=8.01), met jeugdigen waarin dit niet werd omschreven (M_{CGAS}=51.35, SD=7.12), toonde dat de eerste twee groepen jeugdigen enigszins hoger scoren. De vergelijking van de verwijsredenen met de classificaties die uiteindelijk in de ggz werden gesteld liet zien dat in 50.8% van de brieven op zijn minst 1 verwijsreden genoemd werd (zoals bijvoorbeeld: "mogelijke depressieve stoornis, graag uw diagnostiek

en behandeling") welke ook uiteindelijk werd geclassificeerd (zoals bijvoorbeeld een classificatie van een depressieve stoornis). In 7.3% van de brieven werden twee, en in 0.3% van de brieven werden drie mogelijke diagnosen correct omschreven. De betrouwbaarheid (de 'sensitiviteit' en 'specificiteit') van de verwijsredenen varieerde van voldoende tot goed voor de verschillende stoornisgroepen. Zo varieerde de 95% betrouwbaarheidsinterval van de sensitiviteit tussen 60.9-70.6 voor angststoornissen tot 90.5-100 eetstoornissen). De positief voorspellende waarde van de verwijsredenen was vooral laag voor gedragsproblemen (95% BI: 11.6-17.5) en het hoogst voor eetstoornissen (95% BI: 54.1-78.7). De resultaten van de logistische regressie analysen toonden geen significante associaties met factoren als leeftijd, geslacht, de ernst van de klachten of de geschatte lengte van het behandelvoorgeschiedenis; behalve voor een effect van leeftijd op de betrouwbaarheid van ADHD als verwijsreden: de verwijsbrieven waren enigszins beter in het voorspellen van de diagnose van ADHD naarmate leeftijd toenam (OR=1.14, 95% BI: 1.03-1.27). In veel brieven werden ook contextuele factoren zoals problemen op school of thuis genoemd, maar ook andere waardevolle informatie als schoolweigering of lichamelijke klachten. We concluderen dat verwijsbrieven inhoudelijk waardevolle informatie bevatten, alhoewel er ook ruimte is voor verbetering. Niet alleen inhoudelijk, maar ook procesmatig hebben verwijsbrieven een te verwaarlozen potentieel als communicatiemiddel om de ervaren fragmentatie in de zorg te verminderen en om bij te dragen aan de overbrugging van de ene naar de andere zorginstelling indien ze niet worden gezien als enkel een administratief stuk papier. Een belangrijke kanttekening hierbij is dat wij in samenwerking met een huisarts een codeersysteem hebben ontwikkeld om vervolgens alle verwijsbrieven uitgebreid te kunnen inspecteren en te coderen. Deze resultaten zijn dus behaald na ontmanteling van wat er in de brieven staat. We sluiten af met dat vervolgonderzoek nodig is om de taal van de verwijsbrief en besluitvorming aan beide zijden van de zorg in kaart te brengen en om de gedeelde taal te versterken vanuit vertrouwen.

Hoofdstuk 4: Verschillende informatiebronnen en methoden worden gebruikt om tot een voldoende betrouwbaar beeld van de sterke en zwakke kanten van jeugdigen te komen. Voorbeelden hiervan zijn screeningsvragenlijsten en meer uitgebreide gestructureerde of ongestructureerde methoden. We weten uit eerder onderzoek dat elk instrument en informant andere voor- en nadelen met zich meebrengt. Een korte screeningsvragenlijst is

bijvoorbeeld toepasselijk in de eerste lijn waar de focus ligt op herkenning met beperkte middelen en tijd, maar het heeft als nadeel dat het relatief vaak resulteert in vals-positieven. Een zogenoemde multi-informant assessment heeft juist als voordeel dat het zowel de ouders, het kind zelf, als de meesters en juffen van het kind betrekt bij het proces van vergaren van informatie. Echter, het heeft als nadeel dat het meer tijd kost en dat de interpretatie minder rechtlijnig is. Over de verschillen tussen informanten weten we dat daar waar de ouders en docenten de zogenoemde internaliserende problematiek minder goed herkennen dan jeugdigen zelf, zij juist beter zijn in het herkennen van bijvoorbeeld gedragsproblemen. In klinische richtlijnen wordt geadviseerd om de keuze voor de verschillende methoden en informatiebronnen te laten leiden door wat er nodig is daar waar de jeugdige zich bevindt. Dit resulteert erin dat verschillende methoden worden ingezet op verschillende momenten in de zorg. Vanuit het perspectief van efficiëntie rijst hiermee de vraag wat de onafhankelijke waarde van elk van deze methoden is. Om deze vraag te beantwoorden hebben wij de toegevoegde en unieke voorspellende waarde getoetst van verwijsbrieven, resultaten van een screening, een multi-informant assessment methode en de online beoordeling van voorgaande informatie door een clinicus, in het voorspellen van de classificaties zoals gesteld in een gespecialiseerde kind-en jeugd ggz instelling. Daartoe hebben we dossiers onderzocht van 1259 jeugdigen die tussen januari 2015 en december 2017 zijn verwezen. Hun verwijsbrieven, responsen op de 'Strenghts and Difficulties Questionnaire' (de SDQ), resultaten volgend op de gesloten vragen de Development and Well-being Assessment (DAWBA), en het door de clinicus beoordeelde totaal plaatje inclusief de responsen op de open vragen in de DAWBA omgeving, werden gekoppeld and de classificaties gesteld na een intake afspraak en psychiatrisch onderzoek. Na multipele imputatie van de ontbrekende datapunten werden logistische regressie analysen gedraaid met voornoemde vier bronnen als voorspeller en als uikomst de vijf meest voorkomende classificaties in de ggz – angststoornissen, depressie, autisme spectrum stoornissen, aandachtstekort hyperactiviteitsstoornissen, of gedragsstoornissen. Voor elk werd een aannemelijkheidsquotienttoets (likelihood-ratio test) en wedverhouding (diagnostic odds ratio) berekend. De resultaten van alle vier de methoden toonden een significante associatie met de uitkomstmaat, de classificaties. Opeenvolgend toevoegen van elk van de vier methoden bleef resulteren in een significante mate van toegevoegde waarde voor de

voorspelling van de uitkomstmaat, met uitzondering van de waarde van de clinicus voor de voorspelling van de classificatie van een gedragsstoornis. De voorspelling van deze classificatie werd niet beter door toevoeging van het oordeel van de clinicus. Alle vier de methoden toonden unieke associaties met de uitkomst wanneer er gecontroleerd werd voor de effecten van de andere drie methoden, behalve de SDQ in het voorspellen van een depressie of gedragsstoornis. Naast deze toetsende aanpak hebben we ook descriptief bekeken wat de waarde is van vier opeenvolgende instrumenten. In deze exploratieve analyse zagen wij dat van de jongeren bij wie alle vier de methoden een mogelijke angststoornis indiceerde, er 48.8% uiteindelijk als zodanig werd geclassificeerd. Dit percentage was 54.9% voor opeenvolgende indicaties van een depressieve stoornis, 85.7% voor vier opeenvolgende verdenkingen van een autisme spectrum stoornis, 70.0% voor ADHD en 10.7% voor opeenvolgende indicaties van gedragsstoornissen. In tegenstelling tot de algemene literatuur waarin gesuggereerd wordt dat de meer uitgebreide methoden sterkere voorspellende waarden zouden kunnen tonen, zagen wij geen eenduidige significante verschillen. De studieresultaten ondersteunen dus niet het idee dat het loont (in termen van detectie) om het meest uitgebreide instrument direct in te zetten en ondersteunen juist een stapsgewijze aanpak. We concluderen dat het gestructureerd verkrijgen en geïntegreerd gebruik van verschillende methoden van waarde kan zijn voor de diagnostiek in en tussen de eerste- en tweedelijnszorg. Dit zou nu in een prospectieve studie verder onderzocht moeten worden.

Hoofdstuk 5: in deze algemene discussie plaatsen we de bevindingen in een groter geheel met als doel het maken van uitspraken over de klinische implicatie en praktische waarde van de studieresultaten. Doel van deze dissertatie was het beantwoorden van een drietal hoofdvragen over het thema van herkennen, verwijzen en het evalueren van psychische problematiek onder kinderen en jeugdigen. Vele eerdere studies hebben al aangetoond dat de problemen die een jeugdige ervaart met zijn of haar psychisch welbevinden, met name onvoldoende worden herkend wanneer het zogenoemde internaliserende problematiek betreft. Dit type klachten zijn naar binnen gericht en leggen zodoende minder snel een tol op de omgeving van het kind. Deze klachten worden dan ook niet vaak als zodanig herkend of omschreven door ouders of niet expliciet geuit door jeugdigen zelf. Omdat het voor vroege herkenning van belang is ook een beeld te hebben van wat er aan het begin van het

diagnostisch proces gebeurt aan de kant van de huisarts, hebben wij in een tot nu toe unieke opzet onderzocht wat de eerste diagnostische overwegingen van huisartsen zijn wanneer zij jeugdigen met angststoornissen zien. We vonden dat zij angststoornissen niet alleen minder vaak herkennen dan wat wij hadden verwacht op basis van de inhoud van de casusbeschrijvingen, maar ook minder dan wat had gemogen op basis van hoe vaak angststoornissen voorkomen in de huisartsenpraktijk of de populatie. Voor de klinische praktijk betekent dit dat er verbeteringen behaald zouden kunnen worden door kennis over de prevalentie, symptomatiek en impact van psychische problematiek te vergroten. Alhoewel het niet de taak van de huisarts is om precisie-diagnostiek uit te voeren, onderstreept de bevinding dat het minder vaak geselecteerd werd dan wat verwacht kan worden op basis van enkel kennis van hoe vaak het voorkomt het belang van het verbeteren van hun kennis. Juist in praktijken waarin gestructureerde vragenlijsten niet worden ingezet is kennis aan de kant van de professional cruciaal. Zonder bewustwording van dat er mogelijk een probleem speelt, zal herkenning uitgesteld worden en zo ook de erkenning van de ondersteuningsbehoefte. Wij vroegen aan de huisartsen ook waar jeugdigen met een angststoornis het beste behandeld kunnen worden. Hieruit maakten we op dat zij angststoornissen in ieder geval niet minder prioriteren dan andere stoornisgroepen. We concluderen dat de onderdiagnostiek van angststoornissen niet een bewust proces van beperkt problematiseren is, maar dat het vooral te maken heeft met onderherkenning van het probleem. Ook in onze studie van de verwijsbrieven vonden we dat met name jeugdigen die in de ggz behandeld werden voor een angststoornis, niet daarvoor verwezen waren. In ruim de helft van de verwijsbrieven was wel een kernsymptoom of mogelijke diagnose genoemd dat in lijn lag met de uiteindelijke classificatie. Ook de context van het kind werd vaak omschreven in de brieven. Op basis hiervan concluderen wij dat verwijsbrieven wel degelijk relevante informatie bevatten. In tegenstelling tot de huidige houding van professionals en de ondertoon dat verwijsbrieven klinisch niet van nut zouden zijn, betogen wij dat verwijsbrieven de communicatie tussen de eerste- en tweedelijn alsook de beleving van het proces door gezinnen zou kunnen verbeteren. Om dit te bereiken zijn nodig een heldere verdeling van de verantwoordelijkheden van zowel verwijzer als ontvanger en zoals het in sommige andere landen gebeurt de verwijsbrieven op te nemen in de curricula en training van professionals aan beide zijden van de zorg. In de laatste studie deden we een

poging om een belangrijke psychometrische waarde, namelijk de toegevoegde voorspellende waarde van een stapsgewijze evaluatiemethode te onderzoeken. Studies naar geïntegreerde zorg suggereren dat professionals behoefte hebben aan ondersteunende methoden voor de besluitvorming en diagnostiek, maar dat de selectie en interpretatie van deze methoden een barrière vormt. Deze moeite is ook niet verbazingwekkend als we het plaatsen in het licht van dat een recente literatuurstudie 672 publicaties over verschillende vragenlijsten heeft gevonden, maar dat enkel 4 van deze voldeden aan de criteria dat het psychometrisch van voldoende kwaliteit, kort, gratis en makkelijk vindbaar is. Er is dus behoefte aan methoden die wetenschappelijk onderbouwd en ook praktisch implementeerbaar zijn. Dat wil zeggen voldoen aan de psychometrische eisen van betrouwbaarheid, alsook toepasbaar zijn binnen de beperkte tijd en expertise van de praktische (eerstelijns-) zorg. Middels data verkregen van een zogenoemde 'best-practice' (=een praktijk die probeert de procedures vorm te geven op basis van onderzoeksresultaten), hebben wij de toegevoegde waarde van vier instrumenten onderzocht, elk met potentieel voor verschillende doeleinden in en tussen de eerste en tweede lijn. De kracht van deze aanpak is dat het voldoet aan voornoemde vier criteria, dat het informatie van verschillende informanten (jeugdigen zelf, ouders en bijvoorbeeld docenten) combineert, het bestaat uit verschillende elementen die passen bij de behoeften van zowel de eerstelijns- als de tweedelijnszorg, en het een rapport produceert voor de betrokkenen en zodoende een communicatiemiddel kan vormen. De resultaten suggereren dat alle vier de instrumenten unieke informatie toevoegen, maar dat geen van hen voldoende is om de classificatie op te baseren. Prospectief onderzoek naar de toepasbaarheid en effectiviteit van deze stapsgewijze aanpak is nodig. We sluiten af met een nabespreking van de sterke en zwakke kanten van de studies in dit proefschrift en adviezen voor toekomstige studies en de praktijk. We wijzen op de risico's van een dogmatische benadering van evidence-based assessment; de voordelen van een meer bottom-up benadering door belangrijke maten zoals "toegevoegde waarde" te definiëren vanuit het perspectief van alle betrokkenen; de rol van de POH-GGZ (praktijk ondersteuner huisartsen-ggz), en de vraag of een focus op een specifieke stoornisgroep behandelinhoudelijk een wezenlijk verschil maakt.

Acknowledgements

I would like to express my sincere gratitude to my (co-) supervisors: prof.dr. P.M. Westenberg, prof.dr. M.E. Numans, prof.dr. R.R.J.M. Vermeiren, dr. M.R. Crone, drs. B.M. Siebelink, to the members of the promotion committee and the opposition committee: prof.dr. A.W.M. Evers, prof.dr. V. Hendriks, prof.dr. E.M.W.J. Utens, prof.dr. M.H. Nauta, dr. I.A. Arnold, and dr. C.H.A. Verhaar.

Special thanks to the participants and organizers of the 2018 Boerhaave Vorderingen en Praktijk. Also, I am grateful for the fruitful discussions with dr. E. Dusseldorp, and dr. J.R. van Ginkel for the studies described in chapter 3 and 4. Furthermore, many thanks to dr. J. Eekhof for his feedback on our coding scheme, to M. van Velzen, Msc., and B. Peeters, Msc., for extracting the referral letters for the study described in chapter 3.

Curriculum vitae

Semiha Aydın was born in Rotterdam on August 16th, 1992. She started studying Psychology at Leiden University after graduating from her secondary education at Libanon Lyceum in 2010. In 2012 she finished with pleasure her Bachelor thesis on the topic of math strategies in elementary education at the department of Methods and Statistics of Leiden University. When in the same year a work-group facilitator asked Semiha whether she would like to give workgroups, her interest in combining the roles of researcher, practitioner and teacher grew. After finishing her masters in Clinical Psychology, Health Psychology and courses on Clinical Neuropsychology in 2014, she started working in clinical practice and facilitating workgroups. Whilst working in clinical practice, she started also as a research assistant at the Institute of Pedagogical Sciences, Leiden University. In 2017 she started her multi-disciplinary PhD project at the Institute of Psychology (unit Developmental and Educational Psychology) of Leiden University, the departments of Public health and primary care, and Curium Child and adolescent psychiatry, Leiden University Medical Centre. Starting from 2018, Semiha also developed and implemented an ongoing programme for aspiring first generation students, aimed at easing the gap between pre-university education and the university. As of 2021, Semiha started working as a therapist again. After her PhD, she will continue her work as a postdoctoral researcher on her project on schema-therapy for adolescents and young adults with social anxiety disorder, as a therapist in practice, and by guiding students doing for example their clinical internship or writing their thesis.

List of publications

Aydin, S., Siebelink, B.M., Westenberg P.M., Vermeiren R.R.J.M. & Crone M.R. (2023), Optimal scoring methods for autism spectrum disorders in the strengths and difficulties questionnaire (SDQ). Manuscript in preparation.

Kubicki, E., Grootendorst, N.H. & Aydin, S. (2023), Personality Traits and Mental Illness Exposure Among Volunteers in Mental Health and Sport Organisations. Manuscript in preparation.

Aydin S., Siebelink B.M., Crone M.R., Ginkel J.R. van, Numans M.E., Vermeiren R.R.J.M. & Westenberg P.M. (2022), The diagnostic process from primary care to child and adolescent mental healthcare services: the incremental value of information conveyed through referral letters, screening questionnaires and structured multi-informant assessment, British Journal of Psychiatry Open 8(3): e81.

Aydin S., Crone M.R., Siebelink B., Vermeiren R.R.J.M, Numans M.E. & Westenberg P.M. (2021), Angststoornissen bij jongeren herkennen, Huisarts & Wetenschap 64(9): 24–27.

Aydin S., Crone M.R., Siebelink B.M., Numans M.E., Vermeiren R.R.J.M. & Westenberg P.M. (2021), Informative value of referral letters from general practice for child and adolescent mental healthcare, European Child and Adolescent Psychiatry, doi:10.1007/s00787-021-01859-7.

Aydin S., Crone M.R., Siebelink B.M., Vermeiren R.R.J.M., Numans M.E. & Westenberg P.M. (2020), Recognition of anxiety disorders in children: A cross-sectional vignette-based survey among general practitioners, BMJ open 10(4): e035799.

The societal toll and human misery associated with mental health disorders is well established. Nonetheless, only about one in five minors with mental health difficulties access adequate professional support. Care pathways and procedures have to facilitate timely recognition and adequate evaluation of patients' needs to navigate those who can benefit towards 'the right service in the right place, at the right time, and delivered by the right person' - a meaningful Chinese proverb often used by authors from the field of healthcare.

To reach this meaningful goal, professionals should be able to 'look' at a patient and 'see' patients' needs. This by relying on their sufficient knowledge to recognize a probable mental health need, their skills and experience to enquire further, known methods to reliably assess strengths and weaknesses, and readily available resources to translate what they see into an adequate support or management plan.

In this PhD thesis we present a series of studies aimed to fill gaps in empirical knowledge on this topic of recognition, assessment, and referral of youth with mental health problems at the interface of primary care and mental healthcare. The results of the studies could inform clinicians on the status quo considering recognition and referral of minors with mental health problems. Moreover, the findings could serve policy and curriculum makers, thereby improving effective practices in child and adolescent mental healthcare.

In between looking and seeing