Research Paper

Identification of children at risk for mental health problems in primary care—Development of a prediction model with routine health care data

Nynke R. Koning[a,*], Frederike L. Büchner[a], Robert R.J.M. Vermeiren[b,c], Mathilde R. Crone[a], Mattijs E. Numans[a]

[a] Department of Public Health and Primary Care, Leiden University Medical Centre, PO Box 9600 Postzone V0-P/V6-68, 2300 RC Leiden, The Netherlands
[b] Department of Child and Adolescent Psychiatry, Leiden University Medical Centre, Curium-LUMC, The Netherlands
[c] VU University Medical Center, Amsterdam, The Netherlands

ARTICLE INFO

ABSTRACT

Background: Despite being common and having long lasting effects, mental health problems in children are often under-recognised and under-treated. Improving early identification is important in order to provide adequate, timely treatment. We aimed to develop prediction models for the one-year risk of a first recorded mental health problem in children attending primary care.

Methods: We carried out a population-based cohort study based on readily available routine healthcare data anonymously extracted from electronic medical records of 76 general practice centers in the Leiden area, the Netherlands. We included all patients aged 1–19 years on 31 December 2016 without prior mental health problems. Multilevel logistic regression analyses were used to predict the one-year risk of a first recorded mental health problem. Potential predictors were characteristics related to the child, family and healthcare use. Model performance was assessed by examining measures of discrimination and calibration.

Findings: Data from 70,000 children were available. A mental health problem was recorded in 27•7% of patients during the period 2007–2017. Age independent predictors were somatic complaints, more than two GP visits in the previous year, one or more laboratory test and one or more referral/contact with other healthcare professional in the previous year. Other predictors and their effects differed between age groups. Model performance was moderate (*c*-statistic 0.62–0.63), while model calibration was good.

Interpretation: This study is a first promising step towards developing prediction models for identifying children at risk of a first mental health problem to support primary care practice by using routine healthcare data. Data enrichment from other available sources regarding e.g. school performance and family history could improve model performance. Further research is needed to externally validate our models and to establish whether we are able to improve under-recognition of mental health problems.

© 2019 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Research in context

Evidence before this study

Child mental health problems are relatively common and although children are regularly seen in primary care, their mental health problems often remain under-recognised. Currently no structured risk prediction models for overall child mental health problem identification in primary care are available.

Added value of this study

Predictor variables and their effects differed partially between pre-school, primary school, and secondary school aged children. Model performance was moderate, identifying the presence of a first (internalizing) mental health problem correctly in approximately two third of the children. The identified predictors can help mental problem recognition in primary care and mental

health risk assessment, especially when multiple predictors are present.

Implications of all the available evidence

This research is a first promising step in developing a prediction model for the one-year risk of a first child mental health problem that can be used in primary care clinical practice as it is based on readily available routine healthcare data. Further research is needed to improve current identification strategies in order to provide children adequate, timely prevention and treatment if needed and prevent adverse outcomes.

1. Introduction

Mental health problems in children are relatively common, with estimated prevalences ranging from 10 to 20% worldwide [1]. Mental health problems are generally characterised by some combination of abnormal thoughts, emotions, behaviour and relationships with others, and they can range from problems with mild to severe impairment. Half of all lifetime mental health problems occur by the age of 14 years and 75% by the age of 24 years [2]. Most children visit primary care professionals, usually general practitioners (GPs) or paediatricians, at least once a year despite different healthcare systems across the world [3–5]. Although children are regularly seen in primary care, mental health problems often remain under-recognised, with nearly half of the children with mental health problems not being recognised as such [1,6–9]. Early identification of mental health problems in children is important as they often have a negative effect on children's everyday functioning and wellbeing. It is also known that they have long lasting effects, resulting in for instance a higher risk of impairment due to a DSM-diagnosis later in life and poorer performances at school and/or on the job market [10,11]. Adequate treatment has fortunately proven to be effective and alleviate these long lasting effects [3].

In order to provide adequate and timely treatment for children, identification of mental health problems has to be improved. Risk prediction models based on a number of patient and disease characteristics available in medical registrations are an integral part of current clinical practice in primary care [12,13] and might provide an efficient solution to improve early mental health problem recognition. Prediction models for anxiety and depression in adolescents [14] and adults [15,16] in primary care have been developed and have shown good discriminative properties, with only the study on depression in adolescents solely based on readily available routine healthcare data. To our knowledge, models based on readily available routine healthcare data that help identifying mental health problems in children and adolescents in primary care are not available yet. Such a model estimating the probability of a child having a mental health problem in the next year might help professionals to better recognise problems in daily practice, thereby improving timely recognition. Specific mental health problems have a higher incidence at different ages, which means that risk factors for mental health problems may vary across childhood and adolescence [17,18]. During childhood and adolescence, children might also experience events that alter their prognosis for a first mental health problem from that time onwards. The aim of our current study is therefore to develop a prediction model for the one-year risk of a first recorded mental health problem in general and internalizing problems (i.e. depression, anxiety or somatization) in particular in children and adolescents presenting in primary care; taking into account age and time-varying factors. We developed different risk prediction models for children in different age groups.

2. Methods

2.1. Study design and setting

We performed a population-based cohort study among primary care patients aged 1–19 years who were registered with 76 practice centers (107 GPs) that were affiliated with the ELAN primary care network (Extramural Leiden Academic Network) of the Leiden University Medical Centre (LUMC), the Netherlands. The participating practices are located in the greater Leiden area and are representative for Dutch primary care.

In general, all residents of the Netherlands, including children, are registered with a GP in his/her neighborhood. Primary care is free of charge for children and no private primary healthcare system exists in the Netherlands. Dutch children visit their GP on average once a year. All children registered with participating GP practice centers were included in our study regardless of whether they have visited the GP during our study. The GP is the gatekeeper of the Dutch healthcare system and to enter secondary care, a referral from the GP is needed.

Our data consisted of the routine healthcare data anonymously extracted from the electronic medical records (EMRs) from the participating practices [19]. Available patient data included demographics, consultation dates, symptoms and diagnoses coded according to the WHO International Classification of Primary Care (ICPC), prescribed medication coded according to the Anatomical Therapeutic Chemical (ATC) classification, laboratory test results, and descriptive or coded information of referrals and correspondence with other healthcare professionals e.g. profession/specialty of the other professional and date of referral and correspondence [20,21].

2.2. Study population

All patients aged 1–19 years on 31 December 2016 and registered with participating practices between 1 January 2007 and 1 January 2017 were part of our cohort. We excluded patients who had missing data on gender ($n = 11$), registration date with practice ($n = 961$), patients with a negative follow up ($n = 852$), or a missing postal code ($n = 1274$). Patients who had a recorded mental health problem before 1 January 2007 ($n = 3415$), or with an undated mental health problem diagnosis ($n = 7$) were also excluded.

For each patient we determined an entry date to the cohort, which was the earliest of either date of registration with the practice plus one year or the beginning of the study period (1 January 2007). Patients were censored at the date of their first mental health problem, death, deregistration with a practice in the cohort, last upload of EMR data, or the study end date (31 December 2016).

2.3. Outcomes

Our main outcome was a first recorded child mental health problem based on the presence of at least one of the following: a recorded mental health problem, a referral to child mental healthcare and/or a mental health medication prescription between 1 January 2007 and 1 January 2017 (Supplement Table 1). We defined a recorded mental health problem when ICPC codes from the P (psychological) chapter or ICPC code T06 ('anorexia nervosa/bulimia') were present, including both mental health symptoms as well as hypothesised and confirmed disorders. Related mental health medication prescriptions were defined as prescriptions coded with ATC codes N05A, N05B, N05C, N06A, N06BA02, N06BA04, N06BA09, N07BA, or N07BB. Referrals to child mental healthcare were defined as referrals to a psychologist, psychiatry, or psychotherapy. We also investigated model development for first

recorded internalizing mental health problems specifically, consisting of the most common registered internalizing mental health problems depression, anxiety, and somatisation symptom and disorder ICPC codes an medication ATC codes (Supplement Table 1).

2.4. Predictor variables

As predictor variables we included characteristics related to the child (e.g. gender, age, somatic complaints and co-morbidities), social context (e.g. family history of mental health problems and parental divorce recorded in the child's EMR) and healthcare use (e.g. number of visits, referrals, and laboratory tests). As it is likely that interactions exists between the variables somatic complaints and chronic disease, we investigated this in all models. The predictor variables were identified based on a systematic review we conducted regarding predictors for identified mental health problems in primary care [22], current guidelines including risk factors for mental health problems and an expert panel consisting of authors NK and MC, two GPs, a preventive youth health physician, a paediatrician, a pharmacist, and two researchers from the Netherlands Centre for Youth health [18]. We operationalised the predictor variables according to the available data from the EMRs based on ICPC coded diagnoses, ATC coded prescriptions, and count variables (Supplement Table 2). Prior to the data analysis, the count variables were dichotomised according to expert opinion into more than two visits, one or more prescription(s), one or more laboratory test(s), and one or more referral(s)/correspondence with other healthcare professionals.

Every first occurrence of a predictor was taken into account. As predictor variables for mental health problems may vary across childhood and adolescence, we investigated models for the following age groups separately: pre-school aged children (aged 1–3 years), primary school aged children (aged 4–11 years), and secondary school aged children (aged 12–19 years) [9,23].

The same set of predictor variables was examined in the different age groups, however we required the prevalence of a predictor to be >1% per age group with regard to the clinical usefulness of the predictor. Continuous variables are presented as mean (SD) or as median (IQR) when appropriate. Categorical variables are presented as counts (%).

2.5. Model development

Statistical analyses were carried out with the programs SPSS (version 23) and R (version 3.5.1). To obtain the one-year risk of a first recorded mental health problem, we developed a multilevel logistic regression model per age group; pre-school aged children, primary school aged children, and secondary school aged children. Firstly, the data were split according to the children's age; age 4 years, age 5 years and so on. For every age the status of all predictor variables was updated at the same time at that specific age. We obtained a prediction model per age group by combining the data from those years (e.g. age 4–11 years) and fitting a logistic regression model including a cluster effect on the patient level. This to adjust for using different age years of one patient, for instance at age 4 years and age 5 years [24].

2.6. Model performance and internal validation

Model performance was evaluated by determining measures of discrimination and calibration. Discrimination, the ability of the model to distinguish between children who are diagnosed with a first mental health problem and those who are not, was assessed using the *c*-statistic.

The in-sample calibration of the model was assessed by the calibration plot of actual probabilities versus predicted probabilities. The models were internally validated using bootstrap resampling (500 bootstrap samples) and estimating a shrinkage factor. Brier scores were calculated to assess the average prediction error [25].

The Ethics Committee of the Leiden University Medical Centre issued a waiver of consent (G16.018).

2.7. Role of the funding source

This study was supported by ZonMW, the Netherlands, Organization for Health Research and Development (grant 839110012). ZonMw did not have any role in study design, the collection, analysis, and interpretation of data, the writing of the report and the decision to submit the paper for publication.

3. Results

3.1. Baseline characteristics and prevalence of mental health problems

Our cohort consisted of 70,000 children with a median age of 10.0 years (IQR 10 years) and 35,595 (50.9%) were male (Table 1). The median follow up was 6.4 years. In 19,420/70,000 (27.7%) patients a mental health diagnosis was recorded in the electronic medical record (Table 2). An internalizing problem was recorded in 3501 (5.0%) patients. A first mental health problem was recorded in 3.2–4.4% of children aged 1–3 years, in 4.7–6.7% of children aged 4–11 years, and in 3.8–6.4% of children aged 12–19 years. Most recorded MHPs (55%) were based on the presence of 1 of the 3 following criteria: a recorded ICPC code (81%), ATC code (9%) or a referral for a MHP (10%; Table 3). A first recorded internalizing mental health problem was recorded in 0.5–0.7% of children aged 1–11 years and in 1.0–3.7% of children aged 12–19 years. In adolescents aged 17 years and older, first internalizing mental health problems counted for over half of the first general mental health problems.

3.2. Prediction of a first mental health problem diagnosis

Predictors for a first recorded mental health problem one year later in all age groups were somatic complaints, and the healthcare use related variables more than two GP visits in the previous year, one or more laboratory test and one or more referral/contact with other healthcare professional in the previous year (Table 4). Boys aged 1–3 years (OR 1.60, 95%CI 1.43–1.77) and boys aged 4–11 years (OR 1.65, 95%CI 1.61–1.70) were more likely to have a first recorded mental health problem than girls, while boys aged 12–19 years were less likely to have a first recorded mental health problem than girls (OR 0.82, 95%CI 0.75–0.89). Chronic disease was only positively associated with a first recorded mental health problem in children aged 4–11 years. The co-occurrence of somatic complaints and chronic disease was not associated with a first record mental health problem one year later.

Lower neighborhood socioeconomic status was positively associated with a first recorded mental health problem one year later in children age 1–3 and 12–19 years. A difficult temperament, such as excessive crying or feeding problems (OR 1.27, 95%CI 1.07–1.48) was associated with mental health problems in pre-school aged children but not in school-aged children. Prior developmental problems such as growth delay and speech disorders were related to a first recorded mental health problem in children aged 1–11 years, but not in the eldest age group. Life events were only associated to a first recorded mental health problem in children aged 12–19 years (OR 1.79, 95%CI 1.58–1.99) as they were not reported frequently enough to be included in our analyses for the younger age groups.

Table 1
Baseline characteristics study population.

Characteristics	Children age 1–3 years $N = 27,831\% (n)$	Children age 4–11 years $N = 44,622\% (n)$	Children age 12–19 years $N = 22,629\% (n)$
Male gender	51.3 (14,276)	49.7 (22,178)	47.0 (10,628)
Low socioeconomic status	5.0 (1396)	4.4 (1975)	4.2 (957)
Perinatal morbidity	5.6 (1550)	2.2 (995)	0.3 (69)
Congenital anomaly	10.5 (2928)	12.4 (5519)	14.6 (3307)
Disabilities	0.9 (240)	0.9 (430)	0.9 (195)
Neoplasms	2.0 (558)	4.7 (2086)	6.8 (1534)
Chronic disease*	39.9 (11,098)	38.8 (17,302	38.8 (8770)
Somatic complaints**	20.3 (5650)	33.5 (14,953)	49.8 (11,280)
Tension headache***	0.2 (59)	3.7 (1632)	9.3 (2096)
Migraine***	0.0 (3)	0.4 (164)	2.4 (536)
Abdominal pain***	3.3 (917)	12.9 (5759)	17.5 (3953)
Constipation***	12.0 (3335)	13.9 (6186)	11. 2 (2532)
Tiredness***	1.3 (353)	4.9 (2193)	13.7 (3096)
Other somatic complaints***	6.5 (1804)	11.2 (4991)	27.5 (6216)
Life event	0.4 (109)	0.8 (376)	1.9 (421)
Academic problem	0.0 (1)	0.1 (62)	0.4 (82)
Developmental problem	3.5 (964)	7.1 (3161)	3.7 (839)
Difficult temperament	9.7 (2711)	3.6 (1600)	0.1 (21)
>2 Visits	85.5 (23,789)	82.9 (37,002)	84.4 (19,101)
≥1 Medication prescript	72.4 (20,144)	68.5 (30,562)	69.6 (15,784)
≥1 Laboratory test	12.4 (3457)	23.2 (10,362)	35.5 (8044)
≥1 Referral/correspondence other healthcare prof.	64.8 (18,036)	64.9 (28,942)	66.4 (15,017)

* Chronic disease when present one or more of the following: asthma, eczema, psoriasis, inflammatory bowel disease, epilepsy, diabetes mellitus, cystic fibrosis, rheumatoid arthritis.

** Somatic complaint when present one or more of the following: tension headache, migraine, abdominal pain, constipation, tiredness, irritable bowel syndrome IBS, musculoskeletal symptoms, dizziness, nausea, hyperventilation syndrome, palpitations, fainting.

*** Separate somatic complaints do not add up to the total amount of somatic complaints as a child can have multiple somatic complaints.

Table 2
First recorded (Internalizing) MHPs per age.

Child age (years)	Nr of children without previous MHP	Children with first recorded MHP % (n)	Children with first recorded Internalizing MHP % (n)
1	6193	3.1 (191)	0.7 (41)
2	22,935	3.9 (903)	0.6 (129)
3	23,065	4.4 (1020)	0.5 (114)
4	23,006	4.7 (1070)	0.5 (122)
5	22,878	5.9 (1348)	0.5 (125)
6	22,209	6.0 (1322)	0.5 (122)
7	21,700	6.7 (1464)	0.8 (183)
8	21,054	6.1 (1278)	0.9 (189)
9	20,530	5.8 (1190)	1.0 (203)
10	20,180	4.9 (995)	1.1 (213)
11	20,020	4.6 (912)	1.0 (197)
12	19,861	3.8 (757)	1.0 (206)
13	17,770	4.1 (720)	1.1 (190)
14	15,611	4.8 (750)	1.6 (242)
15	13,425	4.8 (647)	1.7 (229)
16	11,200	5.3 (591)	2.3 (254)
17	9033	6.4 (575)	3.6 (322)
18	6898	6.1 (421)	3.7 (252)
19	4956	5.4 (266)	3.4 (168)

MHP = mental health problem.

Table 3
Characteristics of first recorded MHP.

MHP based on the presence of	Percentage of children with first recorded MHP ($n = 19,420$)
1 criteria: either ICPC code or ATC code or referral	55
2 of the following 3 criteria: ICPC code or ATC code or referral	30
All 3 criteria: ICPC code, ATC code and referral	15
*MHP based on 1 of 3 criteria present:	Percentage of children
Only ICPC code present	81
Only ATC code present	9
Only referral to psychologist, psychiatry or psychotherapy present	10

MHP = mental health problem.

Table 4
Results of adjusted logistic regression analysis for the one-year risk of MHPs.

Covariate	Age 1–3 years 52,193 person years, nr of events 2114				Age 4–11 years 171,577 person years, nr of events 8204				Age 12–19 years 98,754 person years, nr of events 5947			
	Coefficient	Odds ratio	Robust SE	95% CI	Coefficient	Odds ratio	Robust SE	95% CI	Coefficient	Odds ratio	Robust SE	95% CI
Intercept	-4.23				3.60				-3.51			
Male gender	0.48	1.60	0.09	1.43–1.77	0.50	1.65	0.02	1.61–1.70	-0.19	0.82	0.03	0.75–0.89
Low SES	0.47	1.57	0.05	1.49–1.66	0.21	1.23	0.51	0.23–2.23	0.19	1.20	0.07	1.06–1.34
Congenital anomaly	0.06	1.04	0.09	0.86–1.21	0.15	1.15	0.03	1.09–1.21	0.11	1.11	0.04	1.03–1.18
Perinatal morbidity	0.31	1.34	0.07	1.21–1.47	0.12	1.12	0.07	0.99–1.26	NA	NA	NA	NA
Developmental problem	0.38	1.44	0.08	1.29–1.59	0.20	1.21	0.04	1.15–1.29	0.02	1.01	0.09	0.83–1.19
Difficult temperament	0.26	1.27	0.10	1.07–1.48	0.02	1.01	0.06	0.89–1.12	NA	NA	NA	NA
Life events	NA	NA	NA	NA	NA	NA	NA	NA	0.58	1.79	0.10	1.58–1.99
Chronic disease*	0.11	1.11	0.05	0.99–1.19	0.10	1.10	0.03	1.05–1.15	0.03	1.02	0.04	0.94–1.10
Neoplasms	-0.01	0.96	0.15	0.68–1.25	0.06	1.06	0.05	0.97–1.15	-0.02	0.97	0.07	0.84–1.10
Somatic complaints**	0.20	1.19	0.06	1.06–1.32	0.19	1.20	0.02	1.16–1.24	0.20	1.21	0.02	1.17–1.25
>2 Visits	0.31	1.34	0.08	1.19–1.49	0.23	1.26	0.03	1.20–1.31	0.21	1.22	0.04	1.14–1.30
≥1 Medication prescript	0.05	1.02	0.05	0.92–1.13	0.10	1.10	0.02	1.05–1.15	0.27	1.30	0.04	1.23–1.37
≥1 Laboratory test	0.17	1.16	0.07	1.23–1.31	0.09	1.09	0.04	1.02–1.16	0.17	1.17	0.04	1.09–1.25
≥1 Referral/correspondence other healthcare prof.	0.21	1.21	0.05	1.11–1.30	0.29	1.33	0.02	1.28–1.38	0.26	1.28	0.03	1.22–1.35
Somatic complaints* Chronic disease	0.10	1.08	0.08	0.91–1.24	-0.04	0.95	0.03	0.90–1.01	0.01	1.00	0.03	0.94–1.06
Shrinkage factor, B = 500	0.97				0.99				0.99			
C-statistic corrected	0.63				0.62				0.63			
Brier	0.04				0.05				0.05			

MHPs = mental health problems.
NA = not applicable, when predictor was present in <1% of the children a particular age group.
* One or more of the following: Tension headache, Migraine, Abdominal pain, Constipation, Tiredness, Irritable bowel syndrome IBS, Musculoskeletal symptoms, Dizziness, Nausea, Hyperventilation syndrome, Palpitations, Fainting.
** One or more of the following: Asthma, Eczema, Psoriasis, Crohn, Inflammatory bowel disease IBD, Epilepsy, Diabetes Mellitus DM, Cystic Fibrosis CF, Rheumatoid Arthritis RA.

One or more medication prescript was only associated with a first recorded mental health problem in the school aged children. Academic problems and disabilities were not recorded often enough to be included in the analyses for all age groups. In addition, family (mental) health problems were not registered with a specific ICPC code and could therefore not be included in our analyses.

3.3. Prediction of a first internalizing mental health problem diagnosis

Among boys aged 12–19 years, internalizing mental health problems were relatively less often found (OR 0.59, 95%CI 0.48–0.69) compared to girls aged 12–19 years (Table 5), whereas boys aged 4–11 years had an increased risk of a first recorded internalizing mental health problem one year later (OR 1.60 95% CI 1.56–1.64). The healthcare use related variables showed various associations with a first recorded mental health problems. The variables more than two visits in the previous year and one or more referral/contact with other healthcare professional in the previous year were only associated with a first internalizing mental health problem one year later in the school-aged children. One or more medication prescript in the previous year increased the risk of having a first recorded internalizing mental health problem in all age groups. One or more laboratory test in the previous year only resulted in more first recorded internalizing mental health problems in children aged 12–19 years old.

Somatic complaints, chronic disease and congenital anomaly were related to a recorded internalizing mental health problem among the school-aged children. A lower socio-economic status and the co-occurrence of somatic complaints and chronic disease were negatively associated with a first recorded internalizing mental health problem in children aged 4–11 years. A difficult temperament or perinatal morbidity were not associated with internalizing mental health problems in all age groups. Life events were associated with a first recorded mental health problem in children aged 12–19 years (OR 1.59, 95%CI 1.27–1.91) and were not included in the analyses in the younger age groups due to a low prevalence in our data. Again, academic problems and disabilities were not recorded often enough to be included in the analyses for all age groups.

3.4. Model performance

Internal validation for the models for a first recorded mental health problem showed shrinkage factors of 0.97–0.99. The model's discriminatory accuracy for the general mental health problem models was moderate with corrected c-statistics of 0.62–0.63 (Table 4). The Brier scores were 0.04–0.05, indicating good accuracy of probabilistic predictions. Most children had predicted probabilities of a first recorded mental health problem ≤8% with a good calibration (Fig. 1AC). A minority of the children had higher predicted probabilities, which were overestimated.

The shrinkage factor for the model of a first recorded internalizing mental health problem in age group 1–3 years was with 0.81 lower than in the two older age groups 0.96 (age 4–11 years), and 0.98 (age 12–19 years). The corrected c-statistics of the models for a first recorded internalizing mental health problem were 0.64 (age 1–3 years and age 4–11 years), and 0.68 (age 12–19 years), (see Table 4). The Brier scores were low. Most children aged 1–11 years had a predicted probability ≤1% with good calibration (Fig. 1D–F). Children age 12–19 years mostly had a predicted probability ≤4% with good calibration.

Table 5
Results of adjusted logistic regression analysis for the one-year risk of internalizing MHPs.

Covariate	Age 1–3 years 52,193 person years, nr of events 284				Age 4–11 years 171,577 person years, nr of events 552				Age 12–19 years 98,754 person years, nr of events 1853			
	Coefficient	Odds ratio	Robust SE	95% CI	Coefficient	Odds ratio	Robust SE	95% CI	Coefficient	Odds ratio	Robust SE	95% CI
Intercept	−5.98				−3.60				−4.58			
Male gender	0.11	0.94	0.10	0.74–1.13	0.49	1.60	0.02	1.56–1.64	−0.51	0.59	0.05	0.48–0.69
Low SES	−0.05	0.77	0.24	0.29–1.24	0.20	1.19	0.05	1.09–1.28	0.19	1.19	0.12	0.96–1.42
Congenital anomaly	−0.08	0.74	0.97	0.42–2.65	0.14	1.11	0.03	1.06–1.17	0.22	1.23	0.0	1.11–1.34
Perinatal morbidity	0.23	1.08	0.18	0.72–1.44	0.12	1.09	0.08	0.96–1.21	NA	NA	NA	NA
Developmental problem	0.17	1.00	0.24	0.53–1.47	0.19	1.17	0.04	1.10–1.25	−0.06	0.93	0.14	0.64–1.21
Difficult temperament	0.07	0.88	0.15	0.59–1.18	0.01	0.97	0.06	0.86–1.09	NA	NA	NA	NA
Life events	NA	NA	NA	NA	NA	NA	NA	NA	0.047	1.59	0.16	1.27–1.91
Chronic disease*	0.17	1.00	0.12	0.77–1.23	0.10	1.06	0.03	1.01–1.11	0.17	1.18	0.07	1.05–1.30
Neoplasms	−0.05	0.76	0.44	0.09–1.61	0.06	1.02	0.05	0.93–1.11	0.01	0.99	0.11	0.78–1.20
Somatic complaints**	0.17	1.01	0.15	0.71–1.30	0.18	1.16	0.02	1.12–1.20	0.28	1.31	0.03	1.25–1.37
>2 Visits	−0.08	0.74	0.18	0.38–1.09	0.22	1.21	0.03	1.16–1.27	0.30	1.34	0.07	1.20–1.47
≥1 Medication prescript	0.44	1.41	0.14	1.13–1.68	0.10	1.06	0.02	1.01–1.11	0.35	1.40	0.06	1.29–1.51
≥1 Laboratory test	0.06	0.87	0.16	0.56–1.19	0.09	1.05	0.03	0.99–1.12	0.30	1.33	0.06	1.21–1.45
≥1 Referral/correspondence other healthcare prof.	0.27	1.14	0.13	0.92–3.36	0.28	1.28	0.02	1.24–1.33	0.14	1.13	0.05	1.03–1.23
Somatic complaints** Chronic disease*	0.14	0.97	0.18	0.61–1.32	−0.41	0.63	0.03	0.58–0.68	−0.04	0.93	0.04	0.85–1.01
Shrinkage factor, B = 500	0.81				0.96				0.98			
C-statistic corrected	0.64				0.63				0.68			
Brier	0.005				0.008				0.02			

MHPs = mental health problems.

NA = not applicable, when predictor was present in <1% of the children a particular age group.

* One or more of the following: Tension headache, Migraine, Abdominal pain, Constipation, Tiredness, Irritable bowel syndrome IBS, Musculoskeletal symptoms, Dizziness, Nausea, Hyperventilation syndrome, Palpitations, Fainting.

** One or more of the following: Asthma, Eczema, Psoriasis, Crohn, Inflammatory bowel disease IBD, Epilepsy, Diabetes Mellitus DM, Cystic Fibrosis CF, Rheumatoid Arthritis RA.

4. Discussion

In this population-based cohort study among primary care patients we investigated the possibilities to predict the one year risk of a first recorded general mental health problem and a first recorded internalizing mental health problem in children aged 1–3 years, 4–11 years, and 12–19 years based on readily available routine healthcare data. Predictors in all ages were the presence of somatic complaints, more than two GP visits in the previous year, one or more laboratory test and one or more referral/contact with other healthcare professional in the previous year. The occurrence of other potential predictors differed between age groups, advocating for the development of partially different models for different age groups. The models' discriminatory accuracy was moderate.

A recent case-control study with UK routine healthcare data investigating a prediction model for depression in males and females aged 15–19 years found a similar performance compared to our model [14]. Similar patient characteristics like somatic complaints appeared to be predictive in that study [14]. The models in the UK study also contained a more extensive set of predictors including mental health problem symptoms and family-related and social predictors. Healthcare use related variables were not investigated, which were important predictors in our study. Information on academic problems and family mental health problems were not well reported in our study and could unfortunately not be included in our analyses. Investigating the value of additional information on for instance school performance and family history might improve our models [14].

Age-dependent predictors we found were in line with the literature. Boys had a higher risk of a first mental health diagnosis in pre-school and primary school-aged children than girls, probably due to the higher prevalences of externalizing mental health problems (e.g. Attention Deficit and Hyperactivity Disorder) in boys. In adolescence, girls had a higher risk than boys due to a higher occurrence of internalizing mental health problems in girls as is shown in other studies [17]. A history of developmental and temperament problems added to the prediction of a first recorded mental health problem, but only in younger children. At a younger age, developmental problems, such as growth delay and speech disorders, and difficulties in temperament, such as excessive crying or feeding problems are most prevalent and have been found to be related to mental health disorders at a later age, e.g. attention deficit [26,27]. In primary school-aged children a difficult temperament was not associated with internalizing problems, confirming the association between difficult temperament and externalizing mental health problems [26,27]. In adolescence, the registered life-events seemed to play a more prominent role in the identification of a first (internalizing)mental health problem, but they were not often enough recorded in our data to be included in the analyses in the younger age groups.

The combination of somatic complaints and chronic disease diagnoses decreased the likelihood of a recorded mental health problem in high school-aged children and was not significantly associated in the younger age groups. A possible explanation for this might be that physicians relate occurring problems to physical and not mental health issues.

The healthcare use related variables more than two GP visits in the previous year, one or more laboratory test and one or more referral/contact with other healthcare professional in the previous year were all associated with a first recorded mental health problem one year later. One or more medication prescriptions was only associated with a first record mental health problem one year later in the school-aged children. GPs might want to exclude a somatic cause for instance by consulting another healthcare professional or performing laboratory tests before relating problems to a mental health issue. An example for instance is tiredness, which can be
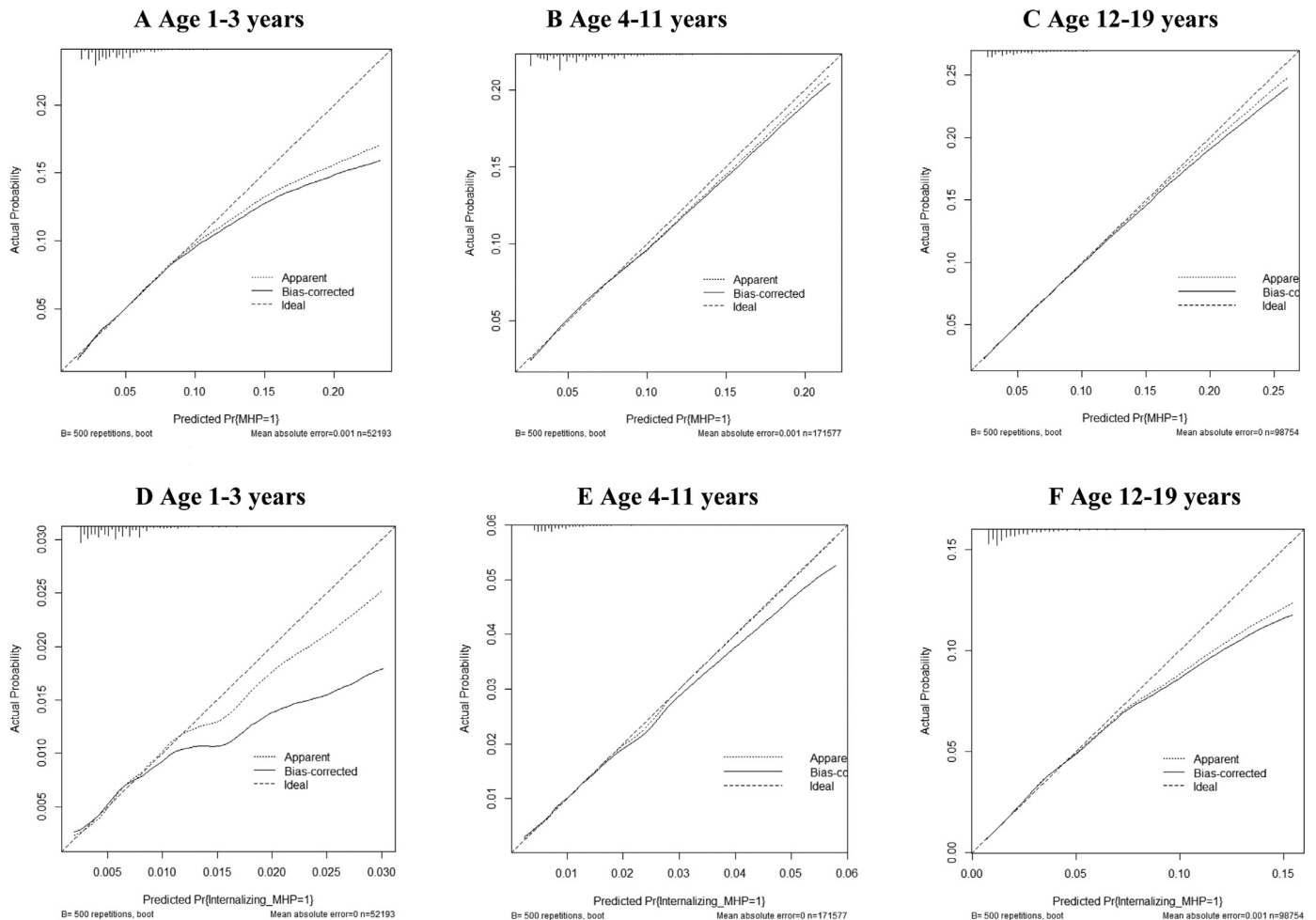
Fig. 1. Calibration plots for predicting the 1-year risk of a first recorded general mental health problem (A, B, C) and internalizing mental health problem (D, E, F). In each plot, the actual observation and predicted probabilities were drawn on the y- and x-axes respectively. The 45-degree dotted line depicts complete agreement between the actual and predicted probabilities.

caused by a somatic problem, but can also be a symptom of a mental health issue. It is common practice to perform laboratory tests to rule out a somatic cause before considering other possible causes. In addition, it might be that the visits, laboratory tests, contact/referral with other healthcare professionals and medication prescriptions are explained in the context of a co-occurring chronic disease or other somatic complaints. It would be interesting to assess the electronic medical records of children who are diagnosed with mental health problems in detail, including the complete free text, to see the course of symptoms, visits, medication prescriptions, referrals and performed laboratory tests to gain more insight in the actual process of diagnosing mental health problems in primary care.

Our study included over 70,000 children in primary care, allowing us to investigate a substantial number of potential predictors. The data consisted of readily available routine healthcare data reflecting daily practice in 'average' primary care. This makes the results potentially more suitable for implementation in practice compared to models requiring (additional) questionnaire information [15,16]. The key advantage of our approach is that it takes into account the time-varying effects of predictor variables, which to our knowledge has not been done in previous research.

A limitation of using routine healthcare data is that possibly useful information might be missing. When the patient consults his GP, the patient presents his symptoms in a specific manner to the GP. The GP then records the information in the medical

record and codes this information. The information is not consistently recorded by GPs. A possible effect of this information bias might be an underestimation of the association between the outcome and for the patient less troublesome or less notable symptoms. This information recording process might also be an explanation for the low presence of school problems, life events, and family mental health problems in our data, variables that have shown to be important risk factors for child mental health problems [14,18], but that will not always be recorded in the EMR of the children. On the other hand, overestimation of the association between outcome and predictors might occur when GPs already suspect mental health problems. For this study, we only had coded information available, we did not have full access to free text notes of the history of a patient for privacy reasons. However, we did have information about the presence of some often used words in the free text of the patient's history, such as 'divorce' or 'school problem'. It is likely that information regarding school problems or life events such as a divorce, if they are registered, are recorded in the free text of the patients record.

It turned out that these words were not often recorded in the free text of the child's medical record and were not of influence on our predictions.

The extent to which the definitions used for our outcomes corresponded to an officially classified mental health disorder needs to be further investigated. For the definition of (internalizing) mental health problems, we included both mental health prob-

lem symptoms and recorded disorders, as according to our expert panel, GPs are cautious of labeling a child with an actual mental health disorder ICPC code. Our models intended to support the early identification of children at risk for mental health problems. It is known that almost half of the children with a mental problem are not being recognised as such in primary care [1,6–9]. Early identification and if needed treatment has shown to improve long-term prognosis. The inclusion of symptoms of mental health problems as outcome in our prediction model might therefore enable the early prevention of adverse outcomes. Research comparing our model estimations with screening tools for child mental health problems or official diagnosis from secondary mental healthcare is needed to investigate whether our models improve primary care identification rates. In addition, the used definition for internalizing mental health problems does not include all children with mental health problems according to the DSM 5 classification and referrals to psychology/psychiatry could not be included in this outcome definition. Our aim was to explore the usefulness of the data in the development of a prediction model for the most commonly registered internalizing mental health problems.

Our data give a fair representation of Dutch primary healthcare. As this research is performed in Dutch primary care, external validation is needed to investigate model performance in other populations with possible other healthcare systems.

Our developed predictions models estimating the one-year risk of a first recorded (internalizing) mental health problems in primary care showed a moderate performance. At this moment we are of the opinion that the models are not good enough yet to be applied in daily clinical practice. The next steps would involve investigating model performance when additional information is included about predictors which from literature are known to be important predictors for child mental health problems such as school performance, life events and family mental health problems. These predictors were not well recorded in the EMR data of the GPs. This information could be added by linking registry data from other sources, for instance from preventive youth healthcare. In addition, more research is needed to investigate whether our models improve primary care identification rates and whether our models are identifying the right children, i.e. children who have an actual mental health problem. This can be done for instance by comparing our model estimations with screening tools for child mental health problems or official diagnosis from secondary mental healthcare. The healthcare use related variables in general were important predictors for a first recorded child mental health problem one year later. Research about the actual diagnostic process of mental health problems could give more insight in the course of symptoms, referrals, laboratory tests and prescriptions. Furthermore external validation, a key element in the development of a prediction model for use in daily clinical practice, is needed to validate the prediction model with external data [28].

In conclusion, our models estimating the one-year risk of a first recorded (internalizing) mental health problem identified in around two thirds of the children correctly whether a first mental health problem was present or not. Especially when multiple predictors are present, the identified predictors can aid mental health problem recognition in primary care. Further research is needed to investigate whether additional information e.g. regarding school performance and family history can improve the performance of the developed models and whether the models also aid mental health problem recognition in the children that are currently not being recognised with a mental health problem by their GP. Also, external validation is needed to investigate the generalizability of our findings.

Declaration of Competing Interest

The authors declared to have no declaration of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eclinm.2019.09.007.

References

[1] Kieling C, Baker-Henningham H, Belfer M, et al. Child and adolescent mental health worldwide: evidence for action. Lancet 2011;378(9801):1515–25.

[2] Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 2005;62(6):593–602.

[3] Vallance AK, Kramer T, Churchill D, Garralda ME. 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):301–9.

[4] Wieske RC, Nijnuis MG, Carmiggelt BC, Wagenaar-Fischer MM, Boere–Boonekamp MM. Preventive youth health care in 11 European countries: an exploratory analysis. Int J Public Health 2012;57(3):637–41.

[5] Bezem J, Theunissen M, Buitendijk SE, Kocken PL. A novel triage approach of child preventive health assessment: an observational study of routine registry–data. BMC Health Serv Res 2014;14(1):498.

[6] Wildman BG, Kizilbash AH, Smucker WD. Physicians' attention to parents' concerns about the psychosocial functioning of their children. Arch Fam Med 1999;8(5):440.

[7] Brugman E, Reijneveld SA, Verhulst FC, Verloove-Vanhorick SP. Identification and management of psychosocial problems by preventive child health care. Arch Pediatr Adolesc Med 2001;155(4):462–9.

[8] Sheldrick RC, Merchant S, Perrin EC. Identification of developmental-behavioral problems in primary care: a systematic review. Pediatrics 2011:2010–3261 peds.

[9] Ormel J, Raven D, van Oort F, 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):345–60.

[10] Hofstra MB, Van Der Ende J, Verhulst FC. Child and adolescent problems predict DSM-IV disorders in adulthood: a 14-year follow-up of a Dutch epidemiological sample. J Am Acad Child Adolesc Psychiatry 2002;41(2):182–9.

[11] Veldman K, Reijneveld SA, Ortiz JA, Verhulst FC, Bültmann U. Mental health trajectories from childhood to young adulthood affect the educational and employment status of young adults: results from the TRAILS study. J Epidemiol Community Health 2015;69(6):588–93.

[12] Fontein D, Klinten Grand M, Nortier JW, et al. Dynamic prediction in breast cancer: proving feasibility in clinical practice using the TEAM trial. Ann Oncol 2015;26(6):1254–62.

[13] Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? Bmj 2009;338:b375.

[14] Nichols L, Ryan R, Connor C, Birchwood M, Marshall T. Derivation of a prediction model for a diagnosis of depression in young adults: a matched case—control study using electronic primary care records. Early Interv Psychiatry 2018;12(3):444–55.

[15] King M, Bottomley C, Bellón-Saameño J, et al. An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: predictA. Psychol Med 2011;41(8):1625–39.

[16] King M, Walker C, Levy G, et al. Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study. Arch. Gen. Psychiatry 2008;65(12):1368–76.

[17] Zwaanswijk M, van Dijk CE, Verheij RA. Child and adolescent mental health care in Dutch general practice: time trend analyses. BMC Fam Pract 2011;12(1):133.

[18] de Wolff M, Theunissen M, van Rooijen K. JGZ-richtlijn psychosociale problemen. JGZ Tijdschrift voor jeugdgezondheidszorg 2017;49(4):90–2.

[19] STIZON - Stichting Informatievoorziening voor Zorg en Onderzoek. https://www.stizon.nl/.

[20] Lamberts H, Wood M. ICPC, international classification of primary care. USA: Oxford University Press; 1987.

[21] Methodology WCCfDS. ATC index with DDDs. Norway: WHO Collaborating Centre for Drug Statistics Methodology Oslo; 2002.

[22] Koning NR, Buchner FL, Verbiest MEA, Vermeiren R, Numans ME, Crone MR. Factors associated with the identification of child mental health problems in primary care-a systematic review. Eur J Gen Pract 2019:1–12.

[23] Verhulst FC. De ontwikkeling van het kind. Uitgeverij Van Gorcum; 2005.

[24] Bouwmeester W, Twisk JW, Kappen TH, van Klei WA, Moons KG, Vergouwe Y. Prediction models for clustered data: comparison of a random intercept and standard regression model. BMC Med Res Methodol 2013;13(1):19.

[25] Royston P. Explained variation for survival models. Stata J 2006;6(1):83–96.

[26] Winsper C, Wolke D. Infant and toddler crying, sleeping and feeding problems and trajectories of dysregulated behavior across childhood. J Abnorm Child Psychol 2014;42(5):831–43.

[27] Hemmi MH, Wolke D, Schneider S. Associations between problems with crying, sleeping and/or feeding in infancy and long-term behavioural outcomes in childhood: a meta-analysis. Arch. Dis. Child. 2011;96:622–9.

[28] Bleeker S, Moll H, Steyerberg E, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003;56(9):826–32.