Identification of children with psychosocial problems differed between preventive child health care professionals

Antonius G.C. Vogelsa,*, Gert W. Jacobussea,1, Femke Hoekstrab,2, Emily Brugmanc,3, Matty Cronen4, Sijmen A. Reijneveld5

a TNO Quality of Life, Prevention and Health Care, PO Box 2215, 2301 CE Leiden, Zuid Holland, The Netherlands
b Municipal Health Service Amsterdam, Postbus 2200, 1000 CE Amsterdam, The Netherlands
c University of Amsterdam, Clinical Psychology, Roeterstraat 15, 1018 WB Amsterdam, The Netherlands
d University Medical Center Groningen, Department of Health Sciences, PO Box 196, University of Groningen, 9700 AD Groningen, The Netherlands

Accepted 12 December 2007

Abstract

Objective: To assess whether differences between individual Preventive Child Health Care (PCH) professionals in the percentage of children they identify as having psychosocial problems are larger than expected based on chance and whether such differences can be explained by differences in parent-reported problems or risk indicators.

Study Design: We used data from three community-based studies among 3,170 Dutch children aged 8 till 12 years, invited for a routine health examination. Parents filled in the Child Behavior Checklist (CBCL) and questions on demographic characteristics. After the examination, PCH professionals registered whether they had identified any psychosocial problem in the child. We examined differences between professionals in the rate of identified children. We used multilevel logistic regression analysis to assess whether such differences could be explained by parent-reported problems or risk indicators.

Results: Significant differences between PCH professionals were found (P < 0.001). These differences could not be explained by parent-reported problems or risk indicators. The differences were largest for children with a score above the CBCL clinical cut-off point.

Conclusion: Some PCH professionals are more likely to identify psychosocial problems than others, independently from parent-reported problems or other risk indicators.

Keywords: Assessment; Child mental health; Clinical preventive services; Physician decision making; Psychosocial aspects; Public health

1. Introduction

Emotional and behavioral problems are quite common among children, in all age groups. Such problems can lead to serious limitations in daily functioning and are likely to persist in later life [1,2]. However, only a minority of these children are treated for these problems. Verhulst and Ende van der found that only 13% of children with behavioral or emotional problems were referred to mental health care [3].

Community pediatric services, like those in the United States and the Netherlands, offering routine health care services to the population as a whole, are in a unique position to identify children with psychosocial problems. In the Netherlands, this early identification is an explicit task of the existing Preventive Child Health care (PCH). Several studies have shown, however, that identification of children with psychosocial problems in community and primary pediatric care is less than satisfactory [4—7]. Clearly, identification of such problems is not always easy, but when it is done carefully, it can help to improve the prognosis of the children involved [8]. Impresions from clinical practice lead us to believe that individual PCH professionals do not always share the same criteria when rating children’s health. Inspection of data collected in other studies (e.g., [9]) had revealed large differences between individual PCH professionals with respect to the proportion of children they identified as being (rather) unhealthy. In 2005, Crone et al. found significant differences between some PCH centers in the percentage...
of children identified as having problems [10]. They suggest that differences in protocols may be relevant in this respect, but also indicate that this can only partially explain the differences they found. In this study, we want to assess (1) whether there are systematic differences between individual PCH professionals in the proportion of children they identify as having psychosocial problems, (2) whether such differences may be explained by differences in the prevalence of problems or differences in possible risk indicators among the groups of children seen by individual PCH professionals, and (3) to assess how large such differences are in relation to children’s Child Behavior Checklist (CBCL) Total Problem Score (TPS).

### 2. Methods

#### 2.1. Sample

We analyzed data from three previous studies that used similar methodologies with regard to sampling and data collection and the same or very similar measures and included children in the age range between 8 and 12 years. The three studies used a similar two-step data collection procedure. First, we randomly sampled a number of regional PCH centers, after stratification by region and degree of urbanization of their district country. Second, each PCH center provided a sample of children from those who were to be invited for a routine preventive health assessment. Study 1 assessed the prevalence of psychosocial problems to be invited for a routine preventive health assessment. After completion by the parents, the questionnaires were returned to PCH, which sent them, unopened, to the researchers. During the preventive health assessment, the PCH professional obtained additional information from the parents on sociodemographic background, family characteristics, and mental health history. Finally, after the health assessment, the PCH professional answered additional questions regarding his or her assessment of psychosocial problems of the child.

The dependent variables in this study were whether the PCH professional identified the child as having any psychosocial problem and whether he or she identified any problem for which the child was referred to mental health care or youth care. This information was obtained using the PCH questionnaire. The PCH professional was asked: “Does the child have a psychosocial problem at this moment?” (Yes or no). Psychosocial problems were defined as emotional and behavioral problems of the child itself. Children with only risk indicators for the development of psychosocial problems had to be coded as having no problems. PCH professionals were explicitly asked to indicate all problems, whether light, moderate, or severe. They also reported whether the problems identified lead them to refer the child to mental health care.

PCH also asked parents whether the child had ever been treated because of psychosocial problems and if the child was currently under treatment. If so, he asked the institution by which the child had been treated.

In each of the three studies, the parents or PCH provided information on the following possible risk indicators: number of (biological or nonbiological) parents present in the family (one or two), parental employment status (1 or 2 full-time or two part-time jobs, one part-time job, or jobless), parental educational level (father’s, if available, otherwise mother’s: low [up to 8 years in education], medium, [9–12 years in education], or high [more years in education]), age, and gender.

The data sets of each study contained an identification code for the PCH professional who examined the child. One of the PCH professionals involved had participated in two studies. We ensured that this professional’s identification code was identical across the two studies. Finally, we knew the discipline of the PCH profession, either a nurse or a physician.

Each study used the CBCL as a measure for the presence of emotional and behavioral problems. The CBCL was part of the questionnaire answered by parents. The 120-item CBCL was developed originally by Achenbach et al. [14] and has been shown to be reliable and to have substantive concurrent and predictive validity [1,15,16]. In studies 1 and 2, the Dutch 1991 version of the CBCL [17] was used. In study 3, the Dutch version of the 2001 American version of the CBCL [18] was used, which has been revised.
regarding the wording of some items and the construction of some of the subscales. Both versions allow for the calculation of a score on the CBCL Total Problem Scale, which includes information from all items and uses all items. We used the standard clinical cut-off points, as defined in the manuals [17,18] to distinguish between children with and without problems. For the purpose of this study, the TPSs of both versions are assumed to measure a very similar concept and therefore, essentially, to be identical. This assumption was checked in the analyses.

2.3. Analyses

We limited the analyses to children in the age range from 8 till 12 years, of Dutch origin and currently not under treatment for psychosocial problems. The age range of 8 till 12 years was chosen because all three studies covered this age group: study 1 contained 1,231 children in this age group, study 2 contained 1,825, and study 3, 1,351. Children from non-Dutch origin—that is, at least one parent born outside the Netherlands—were excluded because of the following reasons. As people from non-Dutch immigrants tend to live in specific neighborhoods and tend to attend specific schools, some PCH professionals see far more children from non-Dutch origin than others. In our sample, 67% of these children were examined by only 22% of the PCH professionals involved. Reijneveld et al. showed that problem identification by PCH among Dutch children was clearly associated with parent-reported problems; for children from non-Dutch origin such an association was not found [19]. This suggests strongly that problem identification among these groups is a more complicated process and should be studied separately from that among problem identification in general.

We also excluded children currently under treatment because this is an almost certain indication of the presence of psychosocial problems and treatment status was known to the PCH professionals. Finally, children with any missing data on the variables used in this study were excluded. The resulting sample, available for analysis, consisted of 3,140 cases, 949 from study 1, 1,444 from study 2, and 747 from study 3.

First, we assessed whether the differences between individual PCH professionals in the percentages of children identified as having problems were larger than might be expected on the basis of a simple binomial distribution. A simple chi-squared test could not be used because some physicians did only see a very limited number of children (in seven cases, only one child). Therefore, we used a Monte Carlo procedure to attain an unbiased estimate of the exact significance level.

Secondly, we assessed to what extent these differences could be explained by differences in TPSs or the potential risk indicators mentioned before. We used bivariate and multivariate logistic regression analyses to assess the statistical significance of the relation of these characteristics and the discipline of the PCH professional—by themselves and in combination—with the identification of problems by PCH. The variables assessed this way are presented in Table 1. Those factors that showed a significant relationship were included in the multilevel analyses that followed as risk indicators. Then, we performed multilevel logistic regression analyses using three models. The first model only calculated the effect of the identity of the PCH professional as a level 2 factor on the likelihood of children being identified as having problems. In model 2, the CBCL Total Problem Scale was added to the analyses. We plotted the logistic predictions based on the CBCL together with the empirical percentage of children identified by PCH. As the logistic prediction fitted the data very well, we decided to use the continuous CBCL score as a predictor in the model. In model 3, risk indicators were added to the analysis as level 1 factors. This way we assessed whether PCH differences remained significant when TPS and background variables assessed this way are presented in Table 1. Those factors that showed a significant relationship were included in the multilevel analyses that followed as risk indicators. Then, we performed multilevel logistic regression analyses using three models. The first model only calculated the effect of the identity of the PCH professional as a level 2 factor on the likelihood of children being identified as having problems. In model 2, the CBCL Total Problem Scale was added to the analyses. We plotted the logistic predictions based on the CBCL together with the empirical percentage of children identified by PCH. As the logistic prediction fitted the data very well, we decided to use the continuous CBCL score as a predictor in the model. In model 3, risk indicators were added to the analysis as level 1 factors. This way we assessed whether PCH differences remained significant when TPS and background variables assessed this way are presented in Table 1. Those factors that showed a significant relationship were included in the multilevel analyses that followed as risk indicators. Then, we performed multilevel logistic regression analyses using three models. The first model only calculated the effect of the identity of the PCH professional as a level 2 factor on the likelihood of children being identified as having problems. In model 2, the CBCL Total Problem Scale was added to the analyses. We plotted the logistic predictions based on the CBCL together with the empirical percentage of children identified by PCH. As the logistic prediction fitted the data very well, we decided to use the continuous CBCL score as a predictor in the model. In model 3, risk indicators were added to the analysis as level 1 factors. This way we assessed whether PCH differences remained significant when TPS and background
characteristics were added to the equation. We used the random intercept variance (RIV) coefficient, a coefficient calculated in multilevel logistic regression analysis, to assess whether there was an effect of the level 2 professionals’ identity. When the 95% confidence interval around this RIV lies above 0, the second-level effect may be considered significant. We estimated only random intercepts and no random slopes for predictors. The multilevel analyses were done in S-Plus [20]. For parameter estimation, we used Penalized Quasi Likelihood with first order Taylor linearization. These analyses were done using both the identification of any problem and the identification of problems for which children were referred as dependent variable.

Some PCH professionals examined very few children. To determine whether the number of children examined had an effect on the results, we repeated all analyses including only professionals who had examined at least 20 children. To check for the assumption that the two different versions of the CBCL used in this study could be considered as identical, we also repeated the analyses with CBCL version number added as a predictor and checked whether version number showed a significant effect.

As mentioned before, we removed children from non-Dutch origin from the analyses. It may be, however, that those professionals who see many of those children, tend to develop a different frame of reference and thereby contribute disproportionately to interindividual variance. We checked for this by an additional multilevel analysis adding the percentage of non-Dutch children examined by individual PCH professionals (ranging from 0% till 78%) as a professional’s characteristic to the model 3.

Finally, to gain some insight in the magnitude of the differences between PCH professionals we found, we first estimated the relative size of the child-adjusted differences in identification between professionals using the variance partition component (VPC) [21]. The VPC indicates which part of the differences can be explained by the level 2 clustering, that is, interprofessional variability, for each value of the other variables included in the model. For these analyses, we used two simplified models: one with the TPS and gender as factor in the model, the other one with the TPS and past treatment for psychosocial problems as predictors. Second, we calculated the model predicted probabilities of identification, in relation to TPS. We did this for professionals on the 5th, 25th, 50th, 75th, and 95th percentile. The probabilities were calculated using a simplified model with TPS as the sole predictor.

3. Results

Table 1 presents information about the sample used in this study. The 3,140 children involved had been examined by 117 different PCH professionals; 61 of the PCH professionals had examined 20 children or more. The number of children examined by each individual PCH professional varied between 1 and 163, with a mean of 26.8.

About one fifth of the children were identified by the PCH professionals as having some psychosocial problem. Figure 1 shows that this proportion varied widely between individual professionals. The black bars present the observed data. The gray bars present the expected numbers, based on the binomial distribution. Variation in the proportion of identified children between PCH professionals was larger than could be expected based on random fluctuation (P-value based on 10,000 Monte Carlo samples <0.001).

The percentage of children identified as having a psychosocial problem for which they were referred to mental health care or to youth care was 6%. The proportions of PCH professionals that referred none, some but less than 6%, 6–10%, and 10% or more were 40%, 23%, 15%, and 22%, respectively. Again, the variation between PCH professionals was larger than could be expected based on random fluctuation (P-value based on 10,000 Monte Carlo samples <0.001).

In bivariate and multivariate logistic regression analyses, most variables presented in Table 1 showed significant relationships with problem identification by PCH. However, we found no such relationship for job status and being examined by a physician or nurse. So, age, gender, number of parents, educational level, and treatment status were included in the multilevel analyses.

Table 2 shows the results of the multilevel analyses with any problem identified as the dependent variable. Model 1 (first column) is the model with only variation between children and between professionals. The associated RIV was 0.34 with a confidence interval well above zero, indicating a statistically significant variation between professionals. The results of the model 2 (second column) showed that the TPS was indeed related to identification.
by PCH professionals with statistical significance, but the second-level RIV remained statistically significantly higher than zero. The results of model 3 (third column) showed statistically significant effects of the risk indicators too, but again the variation between professionals remained statistically significant.

We repeated these analyses with the identification of problems for which children were referred as the dependent variable. In each of the models, the RIV we found was significant: in the first model it was 0.75 (0.55 to 1.00); in the second model it was 0.83 (0.62 to 1.11); and in the final model it was 0.83 (0.63 to 1.12).

The analyses, which added CBCL version number as a predictor in the model, did not show a statistically significant effect for version number, nor did the second-level RIV associated with the professionals’ identity change substantially. The results of the analyses in which the percentage of non-Dutch children examined by individual PCH professionals was added to model 3, did not lead to other conclusions, either. The same applies for the analyses including only physicians who saw 20 children or more.

Figure 2 presents information about the size of the differences between professionals, in relation to the TPS and the child’s gender. The size of the differences is clearly related to the TPS. For boys, the differences between individual PCH professionals are most marked when they had a TPS of about 40. This is near the CBCL clinical cut-off point (36–40, depending on gender and age). For boys with a TPS around 40, about 9% of the variance in problem identification must be attributed to the identity of the PCH profession by whom they are examined. For girls, the differences are most marked in relation to a TPS of about 50, well above the standard TPS cut-off point.

Figure 3 presents the effect sizes in relation, again, to the TPS and stratified to whether children underwent any treatment for psychosocial problems in the past. The relation to the TPS was similar to that presented in Fig. 2. For children who had been treated, individual differences between professionals were more marked at somewhat higher TPSs than for those who had never been treated.

Figure 4 finally presents the model predicted probabilities of problem identification in relation to TPS. The solid line in the center indicates the mean probability, indicating that, on average, children with a TPS of 40 have a probability of about 40% of being identified as having problems. For 50% of all professionals, those between the two lines indicating the 25th and 75th percentiles, this probability lies between about 30% and 50%. For 5% of the professionals, this chance is less than 20% and also for 5% this probability is more than 65%. Again, we see that the differences at very low or very high TPSs are much smaller. Yet, on a large part of the scale, the differences remain substantial.

![Figure 2](image.png)

**Figure 2.** Degree of interprofessional variability, as measured by the variance partition components, in relation to TPS and gender.

---

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td><strong>95% CI</strong></td>
<td><strong>95% CI</strong></td>
</tr>
<tr>
<td>Intercept</td>
<td>–1.36</td>
<td>–1.51 to 1.21</td>
</tr>
<tr>
<td>RIV level 2</td>
<td>0.34</td>
<td>0.20 to 0.57</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>0.05</td>
<td>0.05 to 0.06</td>
</tr>
<tr>
<td><strong>TPS</strong></td>
<td><strong>95% CI</strong></td>
<td><strong>95% CI</strong></td>
</tr>
<tr>
<td>Age</td>
<td>0.05</td>
<td>0.05 to 0.06</td>
</tr>
<tr>
<td>Gender</td>
<td>0.87</td>
<td>0.56 to 1.18</td>
</tr>
<tr>
<td>No of parents</td>
<td>0.87</td>
<td>0.56 to 1.18</td>
</tr>
<tr>
<td>Past treatment</td>
<td>0.87</td>
<td>0.56 to 1.18</td>
</tr>
<tr>
<td>Educational level 1</td>
<td>–0.63</td>
<td>–1.13 to –0.13</td>
</tr>
<tr>
<td>Educational level 2</td>
<td>–0.85</td>
<td>–1.38 to –0.33</td>
</tr>
<tr>
<td>No of PCH professionals</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>No of children</td>
<td>3,140</td>
<td></td>
</tr>
<tr>
<td>Mean (range) no of children for each PCH profession:</td>
<td>26.8 (1–163)</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** RIV = random intercept variance; CI = confidence interval.
4. Discussion

4.1. Main results

This study showed that there are important differences between individual PCH professionals in the identification of children with psychosocial problems and that these differences cannot be explained by parent-reported problems on the CBCL or by differences in risk indicators in terms of background characteristics of the child. In other words, some PCH professionals tend to perceive more children as having problems than other professionals. This applies both to the identification of any problem and to the identification of problems for which children were referred to mental health care or youth care. The differences in problem identification between professionals were most marked for children with a TPS above the clinical cut-off point, especially for girls—as compared to boys—and for children who had been treated for psychosocial problems in the past, compared to children who had never been treated.

4.2. Limitations

The three studies used data that are representative for all children of Dutch origin in this age group in the Netherlands under care by PCH. All three studies were done in the context of regular care in Dutch PCH. Yet, there are reasons to suppose that the differences we found are an underestimation of the effects in standard practice. First, we excluded children from non-western origin. We previously showed that identifying problems among children from non-Dutch origin probably is more difficult than among children from Dutch origin, due, for example, to language problems, cultural barriers, and so on [19]. We expect therefore that differences between individual PCH professionals have an even greater effect when children from non-Dutch origin are involved. Furthermore, although data collection was done in the context of regular care, PCH professionals who participated in the studies, may be—on average—more interested in the identification of children with psychosocial problems. If this is true, the variation between PCH professionals in the field will probably be larger than the variation in our sample.

The available data sets posed some limitations to the variables we could use in the analyses. For example, some PCH professionals work mainly in socially deprived neighborhoods. Such work setting characteristics may evidently be related to differences in the proportion of children identified as having problems [22,23]. We could not include this factor in the analysis. However, at least part of the effect of this factor will be related to a higher problem level in problem areas and this will be reflected in a higher TPS of the children involved. As far as this is the case, in our analyses the effects of this factor will be attributed to TPS. Therefore, in our view, it seems unlikely that this factor could...
explain all of the systematic interindividual variance we found and we believe, despite the limitations of this study, that our results indicate actual differences between PCH professionals in the Netherlands. Whether our conclusions may be generalized to other settings is difficult to answer, as we found no other studies assessing this problem as we did. However, it would be interesting to investigate this issue in settings, like community-based pediatric centers and primary health care, in which staff with limited psychological and psychiatric training has to identify children or adults with problems.

It would have been very interesting to determine which characteristics of PCH professionals (e.g., experience and education) could explain the differences between them. Unfortunately, we did not have any data on PCH professional characteristics.

4.3. Fit with the literature

To our knowledge, this is the first study that tried to assess differences between individual health care providers in the identification of children with psychosocial problems, as it occurs in large-scale standard practice. Many studies have reported on inter-rater reliability in the field of mental health assessment, but those studies used small samples and mostly concerned specific instruments (e.g., [24–27]). Such studies are less suited for assessing the differences between individual health care providers in real life, as the study design itself introduces deviations from standard care: two or more persons assessing the same child, using video taping, observations through one-way screens, case descriptions, or vignettes etc.

4.4. Implications

Are the differences we found serious? For children with a very low TPS, the differences we found were relatively modest—and most children score in the lower TPS range. When a child has a TPS of about 40, the probability that it will be identified as having problems lies between 30% and 50% when seen by half of the PCH professionals. This is already a remarkable range. Yet, for 25% of all PCH professionals, the probability is still higher and also for 25% this probability is lower. In other words, the variation is largest at relatively high scores, just where the identification by PCH matters most. In our view, therefore, the differences we found are serious, indeed. We must however take into account that most children have relatively low scores and that our model is therefore based mainly on those children. Our study should therefore be replicated in a sample with more higher scoring children.

Other studies reported that PCH did not identify any problem in about half of all children with a TPS above the clinical cut-off point [7,10]. This finding in itself does not necessarily point to underidentification of problems, as a clinical TPS is only a valid and reliable indication for problems, not absolute and certain evidence [17]. The clinical assessment by the PCH professional might be more correct, for example, because (s)he identifies problems not or insufficiently indicated by the CBCL. However, if this assumption would be the main explanation for the lack of agreement between PCH identification and TPS, we would expect it to be randomly distributed over individual professionals.

This study showed that this is not the case. Identification of emotional and behavioral problems by PCH should thus be improved, thereby diminishing interprofessional variation. The question then remains how to reduce the interprofessional variability. The literature offers several suggestions, which can help to do so. Several studies showed that good-quality questionnaires can improve problem identification by PCH [8,12,28,29]. Wiefferink et al. showed that training PCH professionals and the use of detailed protocols led to an increase in the sensitivity and specificity of problem identification by PCH and thus to a reduction in interprofessional variability. These effects diminished in time, though [30]. This shows that attention should be paid to continuation of the method and protocols used. Crone et al. studied the sustainability of a health education program in Well-Baby clinics [31]. They found that factors such as perceived self-efficacy, responsibility, training attendance, participation in the adoption decision, and level of institutionalization were related to the degree to which physicians and nurses used the program. Such findings may be very relevant for protocols for the identification of psychosocial problems, too. We are not aware of such implementation studies in this field.

The results of this and other studies [7,32], clearly indicate that the development of high-quality protocols and tools and an effective implementation program with continuous attention to adherence to the protocol are highly desirable. Moreover, the available time for applying tools available during routine examinations may have to be increased, too. Further research has to show how this could be achieved most effectively.

4.5. Conclusion

There are systematic differences between individual PCH professionals in the number of children they identify as having problems and these differences cannot be explained by problems—as indicated by a clinical TPS—among these children or other risk indicators.

Clearly, PCH needs to improve the identification of problems among children they examine, especially as other studies have shown that such an improvement is possible [8,28–30].

The interindividual variability that we found shows that there are large opportunities to improve the identification of psychosocial problems among children. To do so, high-quality protocols, including the necessary tools, an effective and structural implementation program and sustained attention to adherence to protocols have to be developed.
Acknowledgment

This study was made possible by grants from ZonMw.

References


