10 Summary
This thesis focuses on the early detection by Dutch Preventive Child Healthcare (PCH) of children with psychosocial problems and especially on the question whether short questionnaires can improve this early detection.

The following questions were answered:

1. Can differences in the number of children identified as having psychosocial problems by individual health professionals be explained by differences in the prevalence of problems or background characteristics in the groups of children they examine?

2. What are the psychometric qualities of a number of short questionnaires PCH could use to identify children with psychosocial problems and could they improve the identification by PCH?

3. Is it possible to develop a Computerized Adaptive Test, using items from questionnaires on psychosocial problems, in order to achieve a short, yet accurate assessment of the likelihood of psychosocial problems being present?

In accordance with the conclusion of an expert meeting in 1999 on the prevention of psychosocial problems among children and adolescents by PCH, the term psychosocial problems will be used in this thesis to indicate emotional and behavior problems and problems at school.

Introduction

Chapter 1 describes how PCH started to focus on psychosocial problems more explicitly in the seventies of the last century. In the beginning the extent to which and the way how this was done was largely dependent on local interests, leading to large differences in identification and management. This changed under the influence of the National Working Group on the Early Detection of Psychosocial Problems (Landelijke Werkgroep Signalering van Psychosociale Problematiek bij Jongeren, LSPPJ) and the introduction of the PCH Basic Working Package. The early detection of psychosocial problems belongs to the uniform part of that Working Package and should thus be offered to all children in a standardized way – as much as possible.

Differences between PCH professionals

In Chapter 2 we studied differences between individual PCH professionals in the number of children they identify as having problems. In order to do so data from three studies, using similar data collections methods and measurements were combined. We excluded children currently under treatment because of psychosocial problems, children from ethnic minorities and children with incomplete data. The remaining study sample consisted of over 3,000 children. These children had been examined by 117 PCH professionals in a routine health examination. Their parents had answered the Child Behavior Checklist (CBCL), a widely used and validated questionnaire on emotional and behavioral problems. On average PCH professionals identified 21% of all children as having problems. However, this percentage was shown to differ significantly between individual PCH professionals. Using multilevel analyses we determined whether these differences could be explained by differences in the prevalence of emotional and behavioral problems or in the presence of risk factors between the groups of children examined by each professional. We found that the differences between professionals could not be explained by these factors. In other words, whether a child with problems
was identified as such, did not only depend on the problems present but also on the individual PCH professionals by whom it was examined. This effect was greatest for children with a CBCL Total Problems Score (TPS) slightly above the clinical Cut-off point.

**KIVPA**

Chapter 3 assessed the psychometric qualities and added value of a short questionnaire on psychosocial problems of adolescents, the Korte Indicatieve Vragenlijst voor Psychosociale problematiek bij Adolescenten (KIVPA).

Based on the answers on the KIVPA a simple total sum score is calculated and three sub-scale scores. The sample consisted of over 1,200 pupils in second grade of secondary education, invited for a routine health examination by PCH. They answered the KIVPA and the Youth Self Report (YSR). Their parents answered the CBCL. Factor analyses were done to evaluate the scale structure. The results showed that the KIVPA, essentially, is a one-dimensional instrument. The hypothetical sub-scales were not replicated in the factor analyses. With the cut-off point recommended by the original authors, sensitivity for a clinical CBCL TPS was 0.57. That means that 57% of children with a clinical CBCL TPS, a strong indication of the presence of problems, was detected. Using the YSR and referral because of psychosocial problems as criteria, sensitivity was 0.82, respectively 0.55. The specificity for these criteria were 0.84, 0.85 and 0.83. The KIVPA was more sensitive for Internalizing than for Externalizing problems. The KIVPA was also found to offer added value: it allowed for a more accurate distinction between children with and without problems than was possible using only risk factors known to PCH.

We concluded that the KIVPA can be suitable for PCH, but that adaptations are desirable, in order to make the instrument more sensitive, especially for Externalizing problems\(^a\).

**LSPPK**

In Chapter 4 we evaluated the psychometric properties of a short questionnaire on psychosocial problems for children aged 4 or 5, the Landelijk Signaleringsinstrument Psychosociale Problematiek Kleuters (LSPPK). The LSPPK is answered by parents. Their answers are summarized in a simple sum score, the Parent Index (PI). Following the routine health examination the PCH Professional indicates which problems are present in his or her opinion. Their answers are summarized in the Child Health Professional Index (CHPI). The sample consisted of over 1,200 parents of children invited for a routine health examination by PCH. Parents had answered the CBCL and the LSPPK. Factor analyses resulted in factors that closely corresponded to the theoretical scale structure. Using the cut-off point recommended by the developers of the questionnaires, the sensitivity for the criterion ‘being treated because of psychosocial problems’ was 0.50. In other words, the LSPPK detected 50% of all children being treated for such problems. The sensitivity for a clinical score on the CBCL TPS, a strong indication of the presence of problems, was 0.69. Using a lower

---

\(^a\) In order to prevent misunderstanding: following this study the KIVPA was adapted. This new version has not yet been validated.
cut-off point would improve the sensitivity, but would also result in a steep increase of the percentage of children being ‘detected’ without reason (specificity). The CHPI did not result in a better detection. We also determined whether the questionnaire could offer added value to PCH: we assessed whether using the LSPPK allowed for a better distinction between children with and without problems, than a distinction based only on risk factors, known to PCH. This proved to be the case.

We concluded that the LSPPK can improve the identification of children with psychosocial problems, but that the instrument needs to be improved.

**PSC**

In Chapter 5 the psychometric qualities and added value of the Pediatric Symptom Checklist (PSC) were evaluated. The PSC is a questionnaire for parents and was developed in the USA. Studies in the USA showed it to be a valid and suitable instrument for the detection of psychosocial problems. The concept of added value in this study was extended: we not only determined whether it could improve the identification using known risk factors, but also whether it improved the distinction based on the clinical judgment of PCH professionals after the routine health examination, not supported by validated questionnaires or other methods.

The sample consisted of 674 parents of children aged seven to twelve, invited for a routine health examination by PCH. They answered the PSC and the CBCL. The PSC allows for the calculation of a single sum score. PSC scores for Dutch children were lower than those for children in the USA. Therefore we used not only the cut-off point recommended by the authors, but also an adapted cut-off point. The internal consistency of the PSC was high (Cronbach’s $\alpha = 0.89$). However, a confirmatory factor analysis showed that the single sum score was an inadequate description of the items and their interrelationships. The sensitivity for a clinical CBCL TPS was 0.72. Specificity was 0.93. Sensitivity for CBCL Internalizing and Externalizing problems were quite comparable and somewhat lower than for the CBCL TPS. Using ‘being treated because of psychosocial problems’ as criterion the sensitivity was 0.42, with a specificity of 0.90. The PSC was shown to offer a significant added value for PCH.

Nearly all parents were able to answer the PSC, but 20% made critical remarks on the questionnaire, especially concerning the discrepancies between questions and the answering categories offered.

It was concluded that the PSC is suitable for the identification of children with problems. To prevent missing too many children with problems, a lower cut-off point than the one recommended in the USA should be used.

**Randomized comparison of PSC, SDQ and PSYBOBA**

In Chapter 6 we compared the psychometric qualities and added value of three questionnaires: the PSC, the Strengths and Difficulties Questionnaire (SDQ) and the PSYBOBA. The PSYBOBA is a questionnaire for parents of children aged 7 to 12 and was developed by the LSPPJ, specifically for Dutch PCH. The SDQ was developed in the United Kingdom and was proven to be suitable for the detection of children with

---

b Following this study, the LSPPK was also adapted. The adapted version has not yet been validated.
problems there. We used the SDQ Parent Form. The PSC is a questionnaire for parents and was developed in the USA. Studies in the USA showed it to be a valid and suitable instrument for the detection of psychosocial problems. The sample consisted of more than 2,000 parents of children invited for a routine health examination by PCH. They answered the CBCL and either the PSC or the PSYBOBA or the SDQ. Which parents got which of these three questionnaires was randomized, in order to enable a randomized comparison.

The scale structure of the questionnaires was evaluated with Cronbach’s $\alpha$ and confirmatory factor analyses. The internal consistency of each of the questionnaires was high and – corrected for the number of items – identical. The confirmatory factor analyses showed that for each of the three questionnaires the scales were an inadequate description of the items and their interrelationships. The cut-off point recommended by the original authors resulted in large differences in sensitivity, while the Areas under Curve were quite similar. We therefore defined adapted cut-off points, those scores that resulted in a specificity of at least 0.90. Using these cut-offs the sensitivity for a clinical CBCL TPS, a strong indication of the presence of problems, varied between 0.78 (PSC) and 0.86 (PSYBOBA and SDQ). Differences were not statistically significant. Sensitivity for a borderline CBCL TPS was somewhat lower and, again, we found no significant differences between the questionnaires. Each of the three questionnaires was found to offer significant and substantial added value and thus allowed for a better identification by PCH of children with problems. The PSC performed somewhat less here, but differences were not statistically significant.

Most parents were able to answer the questionnaires. Parents rated the PSC somewhat more often as difficult. Twenty percent of the parents made critical remarks on the PSC, compared to ten percent for SDQ and the PSYBOBA. PCH professionals found using the SDQ more difficult, mainly because of the complexity of the calculation of sub-scales.

It was concluded that each of the three questionnaires could improve the identification by PCH of children with problems. Psychometric qualities do not offer definite arguments to prefer one of the three above the other ones.

**Four SDQ based classification methods**

In Chapter 7 a comparison is made between four different SDQ based methods to distinguish between children with and without problems. These classification methods are:

1. a normal vs. elevated SDQ TPS
2. parents reporting serious problems, yes or no
3. a normal vs. elevated score in one of the SDQ’s problem sub-scales, in combination with an elevated score on the impairment scale
4. a combination of the methods 1, 2 and 3

The sample consisted of more than 700 parents of children aged seven to twelve, invited for a routine health examination by PCH. They answered the SDQ and the CBCL before the routine health examination. We determined the sensitivity and specificity of these four classification methods using 4 criteria: a clinical score on the CBCL TPS, on
CBCL Internalizing Problems, CBCL Internalizing problems and currently being treated because of psychosocial problems.

We found some differences between the four methods, but most of them were not statistically significant. The second method (parent-reported serious problems) was less sensitive for a clinical CBCL TPS, a strong indication of the presence of problems, than the combination method and this method is somewhat more sensitive for a clinical score on Internalizing Problems than method 3 (subscales). The first method did not perform significantly better or worse than the other methods, on each of the criteria used. Each method was shown to allow for a better discrimination between children with and without problems than identification based on clinical judgment and we found no significant differences in this respect.

We concluded that for a first identification of children with problems PCH could use the relatively simple first method: namely determine whether the child has an elevated score on the SDQ TPS.

Computer Adaptive Test

In Chapter 8 we investigated whether the items of the CBCL, SDQ, PSC and PSYBOBA could be used to develop an IRT-based CAT, allowing for an efficient yet accurate identification of children with problems. Using IRT it is possible to determine on which point of a supposed scale items are informative. Suppose, on a test for arithmetic ability, a child has answered the question “713 : 23 =??” correctly. In that case it will not be informative to ask it the question “6 : 3=??” too. Using this information one can select those items that are useful in an individual test administration: children with no or only less serious problems do not have to answer questions that are informative – i.e. distinctive – for children with serious problems and vice versa.

The sample consisted of the same respondents as those in the study reported in Chapter 6. They were randomly divided into two subsamples. The analyses done in the first subsample proved that almost all items could be used in an IRT-based CAT. We developed such a CAT and used that for a simulation study in the second subsample: answers given by parents on the paper and pencil questionnaires were used as if they were given in response to CAT questions. De results showed that discrimination based on the CAT scores between children with and without problems was very accurate. Sensitivity for a clinical CBCL TPS, a strong indication of the presence of problems, was 0.89, with a specificity of 0.91. The CAT procedure was also shown to be effective. On average less than 12 items were needed to determine with 95% accuracy whether a child scored above or below the cut-off point. We concluded that an IRT-based CAT is a promising option for the identification of children with problems. The results of the simulation study, though, have to be replicated in real life, before this method can be implemented in PCH practice.

Discussion and implications

In Chapter 9 the three main research questions of this thesis were answered. The results were discussed and recommendations were made, both for PCH practice and for research.
The answer on the first research question, concerning differences between individual PCH professionals, is that there are indeed significant differences between these professionals and that these differences cannot be explained by differences in the prevalence of problems or risk factors between the groups examined by individual PCH professionals.

Concerning the second research question, we found that the short questionnaires we evaluated, can indeed improve the identification by PCH. The KIPVA and the LSPPK do so less well than the other questionnaires that we evaluated, the PSC, SDQ and PSYBOBA.

The third question concerned the possibility to develop an IRT-based CAT with which an efficient and accurate distinction between children with and without problems could be made. We showed that most items of the CBCL, SDQ, PSC and PSYBOBA were suitable for such a CAT. The simulation study offered strong evidence that such a CAT allows for an efficient and accurate identification of children with problems.

In the discussion we argued that the systematic differences between individual PCH professionals that we found indicate strongly that identification by PCH of children with problems based only on a clinical judgment does not meet the requirements of standardization that must be met for services belonging to the uniform part of PCH’s Basic Working Package. Methods to improve this identification are therefore needed.

The methodological approach of the studies in this thesis allow for a valid generalization of our findings towards the large majority of children under PCH care. However, further research is needed among children from ethnic minorities.

We discussed the value of the criterion measures which were used in this thesis, the CBCL, YSR and being under treatment because of psychosocial problems. We argue that one cannot expect high associations between short questionnaires like the ones we studied and treatment status. The main argument for this is that many children with problems are not being treated. Based on the recommendations of an expert meeting in 1999, we argue that the CBCL and the YSR are relevant and well validated criterion measures for short questionnaires on psychosocial problems. However, they may not be regarded as real golden standards. This is one of the reasons that results of these short questionnaires must be interpreted carefully. Our findings, therefore, do not justify the use of the questionnaires as stand-alone screeners.

The CAT we developed appears a very promising method for the identification of children with problems. However, thus far the evidence for its validity and efficiency is based on a simulation study only. Before this CAT can be implemented, a study is needed assessing its validity and efficiency in real life, as well as its feasibility for parents and in the context of PCH practice.

Sometimes it is argued that the detection of psychosocial problems is only justified when there are sufficient and accessible services available for all problems identified. This proposition must be labeled as shortsighted, as well as in defiance with PCH’s Basic Working Package. It is short sighted, because a high-quality identification will also result in the referral of less children without serious problems. That way, more resources will be available for those who need it most. It is in defiance with the Basic Working Package, because one of the PCH’s tasks is, also, to collect information on the...
population’s health and to provide local and higher authorities with reliable policy information.

Younger children, up to 4 years of age, are seen regularly by PCH. Older children are seen less often. That means that problems developing between the consecutive routine health examinations will not be detected in an early stage. In the context of the so-called school-based ZATs (Zorg en Advies Teams, Care and Advice Teams) methods for continuous monitoring and detection can be developed and applied. PCH is expected to participate in those teams but these activities are not part of the uniform part of PCH’s Basic Working Package. This anomaly should be redressed.

Finally we discuss the way short questionnaires are actually being used in PCH practice. The studies presented in this thesis showed that these questionnaires can improve the identification by PCH. However, the extent to which they will actually do so, not only depends on the quality of the questionnaires, but also on the way they are used in practice. Conversations with PCH professionals and impressions from PCH practice lead us to believe that there are large differences between individual professionals and between individual PCH services. Therefore a careful implementation and continuous quality management is needed.

Implications
Based on the results and the discussion the following recommendations concerning PCH practice are made:

1. For the identification of children with problems PCH should not rely on the clinical judgment of individual professionals alone, but should use methods and instruments which have been shown to be valid.
2. It should use these instruments as they were validated and for the purpose for which they were evaluated; changes should only be allowed after new validation studies.
3. The introduction of instruments in PCH practice should be carefully planned and should be embedded in a continuous implementation and quality management program, ensuring that all staff members know how to use them and continue to use them as intended.
4. A lack of adequate youth healthcare services should be considered as a major argument in favor of high-quality methods to identify children with problems, not used as an argument against it.
5. The identification of children with problems in the ZAT-context must be seen as an essential addition to their identification in the routine health examination. The participation of PCH in these ZATs must be redefined and included in the Uniform Part of the Basic Working Package.

For future research the following recommendations are made:

1. Methods and instruments need to be developed for all groups for which such instruments are not yet available and/or validated, especially children under 5 and adolescents over 14.
2. The quality of the identification of children with problems among ethnic minorities needs to be assessed.
3. Instruments and methods need to be developed and validated which allow for a systematic assessment by PCH of the nature and seriousness of problems and the
need for professional help, when short questionnaires indicate the likelihood of problems.

4. Methods need to be developed to collect data from more than one informant; especially since the Internet and computer adaptive testing may be promising methods in this respect.

5. The question to what extent and under which conditions short questionnaires do actually improve the identification of and care for children with problems in daily practice needs to be addressed.

6. The validity and feasibility of a Computer Adaptive Test in day-to-day PCH practice needs to be investigated.

7. Evaluating the relative merits of a number of comparable questionnaires is to be preferred over the evaluation of a single questionnaire.