Chapter 9

General discussion and conclusion
SUMMARY OF CONCLUSIONS

This thesis addressed several aspects of long-term sequelae of bacterial meningitis (BM) during childhood, with a focus on prediction. Briefly summarized, the following conclusions were drawn:

1. Systematic review of the literature proposed the following plausible and important prognostic factors for the prediction of mortality or sequelae after BM during childhood: complaints >48 hours before admission, coma/impaired consciousness, (prolonged duration of) seizures, (prolonged) fever, shock, peripheral circulatory failure, respiratory distress, absence of petechiae, causative pathogen *Streptococcus pneumoniae*, young age, male gender, several cerebrospinal fluid (CSF) parameters and low white blood cell (WBC) count. Due to limited quality and heterogeneity meta-analysis was not possible, and therefore findings must be interpreted and used cautiously. This emphasizes the need for additional well-conducted prognostic studies on this subject.

2. Children at risk for hearing loss after childhood BM can be identified with the now externally validated clinical prediction model. Independent risk factors included in the model are: duration of symptoms prior to admission, petechiae, CSF glucose level, *S. pneumoniae* as a causative pathogen and ataxia. This model is not a replacement for good audiological follow-up protocols, but is a valuable adjuvant. Further, it can be used in areas with scarce availability of adequate follow-up and support in case of auditory deficits where it can facilitate selection of the group of children that urgently needs hearing evaluation. But, additional local validation and probably updating of the prediction model is essential.

3. Although addition of genetic risk factors did not significantly improve the clinical prediction model for hearing loss after BM, addition of TLR9-1237 SNPs and the combination of TLR2+2477 and TLR4+896 SNPs improved the clinical prediction model. However, this did not reach significance. Therefore, genetic factors contribute to a good clinical prediction model, but to optimize the effect they must be incorporated in the development of the model.

4. External validation of the clinical prediction model for academic or behavioral limitations after childhood BM was not successful, and the model is therefore not ready for use in clinical practice. The combined outcome measure should be disentangled and two new prediction models must be developed.

5. The simulated situation of a vaccinated population with a seven-valent conjugate vaccination against *S. pneumoniae* resulted in good reproducibility of both the prediction
model for hearing loss and for academic or behavioral limitations after BM. Performance of the prediction rules remained good in the otherwise identical population as they were developed in. One can assume that the prediction rules will be applicable in the reality of a vaccinated population. But vaccination does not provide full coverage and serotype replacement might occur. Therefore the prediction models must be updated in a vaccinated cohort.

6. When the Strengths and Difficulties Questionnaire (SDQ) was used as a diagnostic tool, an increased prevalence of emotional and behavior problems in school-age BM survivors was found. Comparing the SDQ with the gold standard, the Child Behavior Checklist (CBCL), the SDQ proved to be an acceptable screening tool for emotional and behavioral problems of BM survivors.

7. Based on experimental animal studies and case series in adults it was hypothesized that children with a history of childhood BM without major sequelae but with proven learning or behavioral problems may have atrophy of the hippocampus. However a study using cerebral MRI-study found no differences in hippocampal volume between (now teenage or adolescent) BM survivors and their controls. There were also no anatomical differences in other regions of the brain. While structural differences were not found, functional imaging (e.g. fMRI or SPECT) should be the next step in research.

The aim of this chapter is to place these results in a broader context and to provide an overview of the aspects that link the individual studies. Using this context as a starting point and with some critical considerations from my side on the strengths and limitations of the work presented in this thesis, I hopefully can make some fair recommendations for future research.

THE IMPORTANCE OF CONTINUING RESEARCH ON BACTERIAL MENINGITIS

There is no reason to believe that currently available research on (the effects of) BM is outdated or irrelevant. As stated repeatedly in this thesis, the incidence of BM has decreased dramatically in the Western world in the last two decades, making it a relatively rare disease. It is often thought that vaccination is predominantly responsible for this trend, but it is only part of the explanation. Worldwide, three bacteria (Haemophilus influenzae type b (Hib), S. pneumoniae, and Neisseria meningitidis) are responsible for the majority of cases of BM. In developed countries almost full elimination of Hib and serogroup C meningococcal meningitis has been achieved by vaccine introduction. But one should not forget that a spontaneous decrease of N. meningitidis serogroup B infections played a key role in the total decrease in incidence of BM in Western countries. Recent epidemics are caused by the introduction of
new virulent clonal groups of serotypes into a population in combination with high transmission rates due to multiple factors like close contact or climatic circumstances. These virulent serotype groups have a cycle of rise, diversification, and decay through transformation and genetic recombination, immune selection, and mutation. Although progress is made with the development of vaccines against local strains of this serogroup, it is essential to realize that epidemics of \textit{N. meningitidis} serogroup B infections are still lying in wait, as the Dutch and the New Zealand epidemics in the last two decades of the 20th century prove.

Next there is the process of serotype replacement in which carriage of non-vaccine serotypes increases when vaccines are introduced. Serotype replacement with increasing incidence of infections has not been described for \textit{N. meningitidis} yet, but new outbreaks with non-vaccine-serotypes have occurred in history and may occur again. Although data varies between different regions in the world, there seems to be a net reduction of invasive \textit{S. pneumoniae} infections due to vaccination, but an increase in infections with non-vaccine-serotypes is observed.

Regardless of decreasing incidence, BM still ranks high in the causes of death in (young) children in high-income countries. An estimated 4100 cases and 500 deaths from BM occurred annually in the United States between 2003-2007, and despite significant decline of childhood BM, the incidence did not decrease in infants less than 2 months of age. To support this, similar observations were made in the Netherlands where the incidence of neonatal and infant BM was stable between 2004 and 2010. Although the situation in the Western world is improving, BM is still a persistent major cause of mortality and morbidity in developing countries. For example, in the 2009 epidemic season, 14 African countries reported 88,199 suspected cases, including 53,52 deaths, showing that meningococcal meningitis is still epidemic in Africa: The World Health Organization calculated that low- and middle-income countries account for 98\% of the estimated 5.6 million disability-adjusted life years (DALYs) attributed to meningitis globally.

Besides incidence and mortality there is the devastating impact of short- and long-term morbidity of BM, which is extensively described in this thesis. Briefly, sensorineural hearing loss, epilepsy, motor handicaps, hydrocephalus and mental retardation are well known major sequelae with a high incidence and a significant impact. But also more subtle outcomes like academic and behavioral problems are frequently observed. This type of sequelae appeared to be present in more than 20\% of childhood BM survivors.

In conclusion research on BM is still relevant, and researchers, governments and society should continue investing money and effort in this disease.
Chapter 9

GENERAL ASPECTS OF THE SEQUELAE STUDIED IN THIS THESIS

Hearing loss
In this thesis the need for good follow-up of hearing abilities after childhood BM was emphasized repeatedly since even mild impairment in hearing abilities may impair auditory, linguistic, communication and learning skills in young children. Therefore, early identification and treatment of hearing loss is indispensable in the acquisition of normal linguistic development. Delay in diagnosis and treatment can be disastrous where cochlear ossification may complicate cochlear implantation 26-29. Studies presented over the last two decades showed that the evaluation of hearing was often not part of standard follow-up, but in the last years awareness of the problem has increased. At present, some good standardized protocols for hearing assessment after BM have been developed and published 22, 27, 30-33. However, to our best knowledge, there are no recent studies that evaluate the compliance to and the impact of these protocols on outcome. Therefore, until it is unequivocally proven that a situation of optimal follow-up and care is reached with the current recommendations, we believe that a good clinical prediction rule has added value in the identification of children that need (prolonged) hearing assessment.

Academic and behavioral limitations
Until recently, academic and behavioral limitations were probably the most underestimated sequelae of childhood BM, both in incidence and impact on quality of life. In the Netherlands good research on these long-term effects was mainly initiated by the research group of Koomen and Van Furth et al. in the first decade of this century. It is notable that parental complains about this subject were the main reason to start studying these problems. Along with comparable work from other countries, these studies were highly successful to create an increasing awareness of the fact that children that seem to survive BM in an apparently good condition have a high chance to develop more subtle but potential disabling problems later in life 17, 21, 24, 34-37. This thesis confirms that academic and behavioral problems after BM are a relevant problem with high incidence.

In addition to the studies on parental perception, incidence and impact on quality of life, Koomen et al. developed a clinical prediction model to identify the children at high risk of academic or behavioral limitations. Unfortunately, we were not able to validate this model externally. As discussed in Chapter 5 there are methodological limitations of the study that can partially explain the validation failure, but in my opinion the main problem lies in the fact that the combined outcome measure “academic or behavioral limitations” is too complex to define and to reproduce. Therefore, two separate clinical prediction models should be developed, one for academic limitations and one for behavioral problems.

It has been recognized that the problem is too important to give up on: in the Netherlands, and probably also in many other Western countries, much was gained in the recognition and
guidance of children with academic or behavioral problems. Unfortunately, in recent years many of the investments made have been reversed due to the economic recession. Besides the impact on quality of life of the individuals affected/involved, it might be a classic case of “penny wise, pound foolish”: poor guidance of children with learning disabilities or behavioral problems in general will probably lead to more school dropout with a kickback on the economy and on society as a result. Good clinical prediction models can play a role in the optimization of efficiency regarding identification of those children who needs more attention and guidance.

Neuroimaging and long-term sequelae after bacterial meningitis

In this thesis the hypothesis that hippocampal atrophy might occur after BM and that it can result in learning disabilities was tested. As a secondary outcome measure other structural differences in the brain were searched for. The hypothesis was based on multiple experimental studies in animals and in scarcely available human (adult) post-mortem pathology and neuroimaging data 38-42. *In vivo* studies investigating the long-term post-BM cerebral defects in humans are rare, and to our best knowledge only two studies by Free *et al.* and by Schmidt *et al.* looked for cerebral lesions and atrophy in adults after BM. In contrast to our results, these studies did find hippocampal and cerebral atrophy respectively in adults some years after BM 41, 43, 44. It is hypothesized that these differences can be explained by the facts that Free *et al.* included subjects with a much higher severity of the disease then our patients, and that both studies included adults instead of children, emphasizing the enhanced ability of the child’s brain to recover from neuronal damage 36, 38, 39.

MRI and other imaging techniques are most often used successfully to study acute BM manifestations and complications in humans and animals 45-49, and gadolinium-enhanced MRI was used successfully to predict hearing loss after BM 50. But the field of research focusing on neuroimaging for long-term sequelae after BM seems to be unexplored. In my opinion it is regrettable that due to financial limitations our group was not able to extend this pilot study. I strongly belief that a large amount of valuable information can be obtained from good imaging studies. The future lies in higher spatial resolution of MRI and in functional imaging, like Single Photon Emission Computed Tomography (SPECT) (which has been used for studying acute BM and tuberculous meningitis 51-56), and functional MRI (fMRI).

THE USE AND MISUSE OF CLINICAL PREDICTION MODELS

The main goal of this thesis was external validation of two previously developed clinical prediction models. Clinical prediction models are regression models that use clinical variables derived from medical history, physical examination, laboratory or other diagnostic tests, etc. to define the individual probability of a specific outcome 57-59. For these models many names...
are introduced in literature (for instance prognostic or predictive model, (clinical) prediction model or rule). This thesis attempts to consistently use the term “(clinical) prediction models” for the regression models, and “(clinical) prediction rules” for the risk scores derived from these models. Risk scores and nomograms are formats used for the clear presentation of the original regression model.

Prediction models are part of the relatively young field of prognosis research, in which the risk of future outcome in individuals with a certain health condition is studied. The recently published “PROGRESS” series, in which international experts in the field describe the complete “why and how” of prognosis research, state: “Prognosis research is thus the investigation of the relations between future outcomes (endpoints) among people with a given baseline health state (start point) in order to improve health” ⁵⁷. Prognostic studies might help to understand pathophysiology and causality, but it is important to understand that this is not the main goal of this field of research. As Moons et al. state: “Every causal factor is a predictor, but not every predictor is a cause” ⁶⁰.

Clinical prediction models are potential powerful instruments that can help to inform individuals about the their possible future health condition and to support clinicians and their patients in decision making ⁵⁷, ⁶⁰. But to achieve this goal a good prediction model must meet certain requirements: first, there is the development phase. In a development study the important predictors are identified and weighed for contribution with regression techniques. Selection of candidate predictors is an important step and should start with expert knowledge about the outcome of interest. All variables possibly associated with the outcome measure are candidate predictors, but as mentioned before, this association does not need to be causal. After the choice of candidate predictors the selection process of these variables towards the final model starts using multivariable regression. There are multiple methodological and statistical ways to deal with this issue, but yet there is no consensus on what the best approach is. However, it is essential that data derived from the predictors is of high quality: definition of the variable must be clear, and measurements must be performed in a standardized and reproducible way. The same of course also applies to the outcome measure. A prospective approach is therefore preferable.

Predictive performance of the resulting model must be defined by the assessment of discrimination and calibration. Then the model must be corrected for overfitting and optimism. Overfitting is the phenomenon that the data used in the development is well described, but that predictions are not generalizable to new but similar individuals. This leads to a too optimistic performance of the model. To correct for this optimism internal validation must be performed with bootstrap techniques to calculate a correction factor (shrinkage factor) to adjust the regression coefficient to more realistic values.

When an adequate model is developed, external validation is essential: prediction models perform less in a different population than the one it is developed in. Therefore its predictive abilities must be tested in a new cohort. External validation can be temporal (in a cohort in
the same center as the development cohort, by the same researchers but at a later time) or geographic (in a cohort in a different center or in a different region or country). A validation study is said to be fully independent when the participants and the investigators of the development cohort are not related to those of the validation cohort. It is obvious that the generalizability of a prediction model is good when the performance remains strong in validation cohorts with relatively big differences in patient characteristics, case mix etc. compared to the development cohort. Then, when a good prediction model is presented and implemented in clinical practice, its value must be evaluated with impact studies. These studies quantify the effect of the model on decision-making by clinicians and on patient outcome.

Next, there is the aspect of model updating. Because of the fact that prediction models do perform less in validation cohorts there is always the possibility that a model cannot be used in this cohort. Therefore, it can be desirable to add new predictive factors, such as (new) biomarkers, genetic factors or results of imaging techniques. Since the development of a new prediction model is time consuming and expensive, updating of the existing model with recalibration of the used predictors or addition of new ones is a more efficient strategy. As stated by Steyerberg et al. in the aforementioned PROGRESS series: “Ideally there should be an ongoing process of model validation and updating”.

At last it should be needless to say that all the steps discussed must have a good methodological quality and must be performed in cohorts with adequate sample size and a sufficient number of events.58-68.

Besides all the methodological and technical aspects of development and validation, a good prediction model must also have good practical usability. Before it will be adapted in daily practice, support from workers in the field is essential. The predictors in the model must be easily available and the model must be as simple as possible and easy to use. For this reason many models are transformed to clinical prediction rules, with a user-friendly risk score or nomogram. Increasingly web based forms are used where only the value of the predictors needs to be filled in and a probability is automatically calculated. It depends on the discipline the model is developed for and the outcome measure it predicts whether and which biomarkers have an additive value. For instance, for decision making in an emergency room only biomarkers that are part of routine care and quickly available are useful. But in oncology, where prognosis and decisions on which therapy to start are often more a matter of days or weeks, biomarkers such as genetic factors can be of great value, even though determination is more time consuming.59.

In recent years there has been an exponential increase in the amount of published clinical prediction models. But after development the majority is never validated and an even smaller number of models has been implemented in clinical practice. Good evaluation of impact is even more rare. Further, the quality of methodology, statistical analysis and reporting is often poor.59,69. For prediction models developed for children the situation is even worse because the quality and the number of validated models is lower than in those developed for adults.69.
Placed in the context discussed above, the research presented in this thesis has some strengths and limitations to address. Two external validation studies, one model updating study and one simulation model updating study were described. After the foregoing discussion, it can be concluded that the fact that both validation and updating were performed is a major strength. In addition, methodology has always been state of the art in accordance with the insights of the time the studies were designed and performed.

The major limitation of the studies presented in this thesis is the fact that sample sizes were rather small. As discussed in the relevant chapters, a rule of thumb for validation studies was that the cohort should be at least half the size of the development cohort. But new insights were gained in recent years indicating this might be too optimistic: to detect minor changes and avoid type II errors (the Null hypothesis of equal model performance is falsely not rejected) in calibration and discrimination, the validation cohort should contain at least 100 events and 100 non-events \(^{58,70}\). With a negative result in the validation of the prediction model for academic or behavioral limitations, type II error is not an issue, but minor differences in performance might have been missed in the validation of the model for hearing loss. But the question is then how clinically relevant these minor differences are?

**CLINICAL PREDICTION MODELS AND GENETIC PREDICTIVE FACTORS**

In Chapter 4 we attempted to improve the clinical prediction model for hearing loss after BM by updating it with host genetic risk factors known for their association with susceptibility to bacterial meningitis and with hearing loss in survivors of bacterial meningitis \(^{71,72}\). But addition of these genes did not result in a significant improvement of the model. It is a known problem in the updating process that addition of new biomarkers may result in no, or only marginal, additional benefit. The originally developed models generally include strong predictors, therefore, the independent effects of new prognostic factors need to be strong before an improvement in performance is achieved \(^{58,59}\). In recent years this phenomenon was found in a number of studies that tried to improve prediction models with genetic predictive factors \(^{73,74}\). With these negative results in mind, the question arises if we should continue to explore the additive value of genetic risk factors in clinical prediction models. In my opinion the answer should be “yes”. Genetic risk factors will play an increasing key role in understanding pathophysiology of disease and outcome, and will become rapidly cheaper and faster available. To find the genetic factors with strongest predictive value new techniques like next-generation sequencing and genome-wide association studies (GWAS) can be used next to the candidate gene approach used to find the SNPs associated with the outcome used in this thesis. Sequentially, the found genetic risk factors must be implemented in the development process of clinical prediction models to avoid the aforementioned problem difficult improvement of already strong models. Large sample sizes and again, international collaboration is hereby essential.
CONCLUSIONS AND FUTURE PERSPECTIVES

First, it can be concluded that research on the long-term effects of childhood BM is and stays essential. Further, prediction models have the ability to be a valuable additive in management and follow-up of many diseases in general and of BM in particular, and they can support decision making. But it is vital to develop strong, usable models for daily practice. An ongoing process of validation, updating and impact evaluation in many different settings and global regions must confirm and maintain the quality.

For the prediction model for hearing loss after BM this means the implementation in Dutch follow-up protocols for hearing evaluation, an update in a recent cohort with children vaccinated according to current guidelines and evaluation of impact after some time.

The prediction model for academic or behavioral limitations needs to be redeveloped. This model must be split in two, creating a model for academic limitations and one for behavioral problems. Genetic factors should be included from the beginning, and the strongest genetic factors may be found with new sequencing techniques and GWAS.

To reach optimal performance and generalizability of the models all these steps should be performed in prospective constructed cohorts with large sample sizes. With the present low incidence of childhood BM in western countries international collaboration is the only way to achieve these goals.

Where it comes to neuroimaging, the future lies in higher spatial resolution of MRI and in functional imaging, like SPECT and fMRI.
REFERENCES


