Chapter 7
Summary and future perspectives
Summary

Type 2 diabetes remains the leading cause of ESRD, despite novel treatment strategies to reduce renal risk (1). ESRD is a worldwide problem due to its large disease burden, shortened survival and high cost of treatment. Due to the projected increase in the number of patients requiring renal replacement therapies there is a high unmet clinical need for novel interventions for this population (2,3). Drug development in type 2 diabetes and CKD is currently hampered by two problems: 1) a clinical endpoint that requires large and complex studies, 2) unexpected adverse outcomes in late-stage clinical trials that led to early termination of the drug development programs, despite that these drugs had promising effects effect on renal and cardiovascular risk factors such as blood pressure, cholesterol, albuminuria or HbA1c.

The first problem of large and complex studies is a consequence of the slow nature of progression of diabetic kidney disease. From early stages of the disease it may take decades before ESRD manifests (4). Clinical trials in late stage of drug development typically have a follow-up of 3-4 years, and therefore prospective randomized clinical trial data spanning the entire disease progression of diabetic kidney disease is not available (5,6). For example, drugs that intervene in the renin-angiotensin-system (RAS) are known to reduce albuminuria and slow renal function decline in patients with type 2 diabetes and CKD, but whether intervention in early stages of diabetic kidney disease is more beneficial than intervention in later stages is not known. In Chapter 2 we developed and validated a model that simulated the entire course of diabetic kidney disease by connecting data from completed clinical trials spanning different disease stages. From the combined dataset we built models that simulated the progression in diabetic kidney disease from early onset until ESRD, with disease stages that were characterized by combinations of the risk markers albuminuria and eGFR. The estimates from our model indicate that treating early in the course of diabetic kidney disease with drugs that intervene in the RAS (i.e. angiotensin receptor blockers or angiotensin-converting enzyme inhibitors) further delays ESRD compared to treating in later stages of the disease and that this effect is more pronounced in younger patients. We also showed that the beneficial effect of early intervention depends to a large extent on the initial albuminuria-lowering effect.
of the RAS inhibitors, highlighting that monitoring short-term treatment effects is relevant to predict clinical outcome.

The second problem of unexpected adverse clinical outcomes is illustrated by several recent drug failures. For example, the direct renin inhibitor aliskiren lowered blood pressure and albuminuria but did not improve cardiovascular outcomes in patients with type 2 diabetes and CKD (7). The PPAR-γ agonist rosiglitazone reduced HbA1c levels but increased the risk for cardiovascular outcomes in patients with type 2 diabetes (8). The CETP inhibitor torcetrapib decreased LDL cholesterol and increased HDL cholesterol, but ultimately increased the risk for cardiovascular outcomes and death (9). It is therefore clear that currently used risk markers provide insufficient information about a drug’s benefit-risk profile on renal and cardiovascular outcomes. This suggests that in order to prevent early termination of late-stage clinical trials, there is a need to improve early safety and efficacy signals. This may not only speed up the conduct of clinical trials, it may also prevent patients from receiving treatment that is not efficacious, or even harmful. Traditionally, drug efficacy on renal and cardiovascular outcomes is estimated from responses in individual risk markers. However, it is known that multiple risk markers are involved in renal disease progression and drugs have effects on multiple risk markers. For example, RAS inhibitors decrease blood pressure and albuminuria, and thereby slow long-term renal function decline (eGFR), but may also increase serum potassium and serum uric acid. These off-target effects may enhance or offset the benefits of treatment (10,11). Developing a drug response score based on integrating changes in multiple parameters is therefore potentially more accurate than using single risk markers alone.

Such an approach, with the so-called PRE score, has been previously validated for predicting the effect of drugs intervening in the RAS (12), but whether it can be applied to other drugs is unknown. In Chapter 3 we performed a post-hoc analysis of the AleCardio study in which patients with type 2 diabetes were treated with the dual PPAR agonist aleglitazar. The AleCardio trial was terminated early due to futility and an increase in hospitalization for heart failure in the active treatment arm. We showed that more stringent baseline inclusion criteria on the basis of individual or combinations of risk markers would not have prevented the observed risk increase for heart failure. This indicates that patients should be exposed to the drug before an assessment of its long-term effects can be made. Indeed, the PRE
score, that integrated changes in all measured cardiovascular risk markers in AleCardio, predicted the observed risk increase for heart failure and was more accurate than any single risk marker alone. We therefore showed that the approach of using the response in multiple risk markers to predict drug effects can also be used beyond RAS intervention. However, prospective validation is required to confirm the predictive capabilities of the PRE score.

In Chapter 4 we applied the PRE score to a phase II clinical trial (RADAR) in which patients with type 2 diabetes and nephropathy were treated with the endothelin antagonist atrasentan, for which data on long-term clinical outcomes is not yet known. First we established the relationship between multiple risk markers and the clinically meaningful outcomes ESRD and hospitalization due to heart failure in a background dataset of completed clinical trials that matched the RADAR population. We then applied the calculated risk marker-outcome relationship to the risk marker responses seen after 12 weeks of atrasentan treatment to calculate a relative risk change of ESRD and heart failure outcomes, compared to placebo. We showed that, based on the integrated response to all measured cardiovascular and renal risk markers, atrasentan treatment is likely to confer renoprotection without an unacceptable increase in the risk for hospitalization due to heart failure. We also showed that treatment responders (defined as an initial reduction in albuminuria of >30%) are expected to have a more pronounced beneficial effect on renal outcomes than non-responders. The results of this study will be validated by the currently ongoing phase III trial SONAR, that will subject patients with type 2 diabetes and CKD to a 6-week enrichment period in which their response to atrasentan is determined. Only patients that respond with a >30% reduction in albuminuria will proceed to randomization for long-term treatment with either placebo or atrasentan.

In the aforementioned studies, the PRE score was used to test the efficacy of drugs at a group level. For individual patient-level outcomes it was not yet clear whether the score can be used to accurately discriminate between patients that are predicted to benefit from treatment versus those that are not predicted to benefit. In Chapter 5 we determined the response in multiple risk markers to RAS intervention between and within individual patients. Consequently, we determined whether integrating the drug effect on multiple risk markers with the score would provide a better prediction of who would benefit from treatment compared to using single markers alone. Our results showed that there is considerable variation in treatment
response to multiple risk markers, not only *between* individual patients, but also *within* individual patients. Integrating the response in multiple risk markers provides a better prediction of whether an individual patient will reach ESRD compared to any single risk marker alone, such as systolic blood pressure or albuminuria. This indicates that all relevant risk markers affected by treatment should be measured within individual patients, in order to predict who will benefit from treatment and who will not. The results were validated in two independent datasets, strengthening our conclusions.

These findings indicate that there is merit in replacing the current long and complex clinical trials that use clinically meaningful endpoints by a surrogate endpoint based on a model that integrates the response in multiple risk markers. However, providing scientific evidence for using a multiple parameter drug response score, such as the PRE score, as surrogate endpoint is not enough to implement it in practice. There is a long-standing discussion whether surrogate endpoints provide sufficient evidence to support drug marketing authorization, or whether clinically meaningful outcomes should be used before novel therapies can be marketed (13-16). Further complicating the matter is that drug development consists of several stakeholders, including academics, pharmaceutical companies and regulatory authorities such as the FDA and EMA, and there may be differences between how each of these stakeholder groups perceives the use of surrogate endpoints. Therefore in Chapter 6 we conducted a questionnaire targeting these stakeholders to enquire for conditions under which surrogate endpoints can be used and whether novel surrogate endpoints should be developed. Specifically, we asked whether using multiple risk markers would be a better concept than using single markers alone. From the results we conclude that all stakeholders are willing to accept surrogate endpoints in clinical drug development, but that many of the currently used surrogates (e.g. systolic blood pressure, HbA1c) are not valid substitutes for clinically meaningful cardiovascular and renal outcomes. However, the stakeholders agreed that a score that integrates the response to multiple risk markers would provide a more accurate assessment of a drug’s effect than using single risk markers alone, but that it would not be able to replace trials in which clinically meaningful outcomes are used as endpoints.

In this thesis we have provided additional data that support a strategy of integrating short-term changes in multiple risk markers in order to more accurately
Summary and future perspectives

assess drug efficacy on clinical outcomes. These findings have been ascertained for different drug classes, but prospective validation is needed and currently ongoing. Furthermore, our data suggests that approaches that integrate short-term changes in multiple risk markers could also provide a more accurate assessment of the treatment effect within individual patients. Stakeholders in drug development may be willing to accept a multiple parameter drug response score as surrogate endpoints, but several hurdles have to be overcome before it can be implemented.

Future perspectives

What is needed to solve the current problems of long clinical trial duration and unexpected adverse outcomes in the future? The findings presented in this thesis indicate that using short-term changes in multiple risk markers provide a more accurate assessment of the ultimate clinical outcome during long-term follow-up than using single risk markers alone. Using a tool such as the PRE score as a surrogate endpoint in future clinical trials may therefore reduce the chance of unexpected adverse outcomes and at the same time, by using the initial drug-induced changes in multiple risk markers and modeling its long-term effects, reduce the need for conducting trials with clinical endpoints that require an unfeasibly long follow-up.

Regulatory authorities could promote the development and validation of novel surrogate endpoints, such as the PRE score that we proposed, through their existing biomarker qualification programs. However, regulators are also mindful that there is no surrogate for safety outcomes measured in long-term clinical trials. Furthermore, even when a tool such as the PRE score may be well established as a surrogate endpoint, postmarketing studies may still be required to confirm the predicted clinical benefits. Ultimately, the implementation of the PRE score will depend on whether future prospective studies unequivocally show that it provides reliable estimates of the ultimate drug effect on clinically meaningful outcomes.

In case the PRE score is prospectively validated and consensus is reached regarding its usability, several possibilities for implementation in practice exist. For example, the PRE score could be used to predict the effect of novel drugs that target an unmet clinical need. There is consensus among stakeholders in drug development that patients should get timely access to novel drugs that target
diseases for which no effective therapy yet exists, or diseases that impose a large disease burden on patients. Regulators are confronted with increased calls for faster market access for such drugs and have been criticized for being overly sensitive to the risks of novel drugs (17,18). At the same time, there is pressure on regulators to only grant marketing authorization to drugs that are safe. In order to speed up marketing authorization for drugs that target an unmet clinical need, regulatory authorities such as the FDA and EMA have set up special drug development pathways to expedite their marketing authorization (19). The PRE score could aid in the development programs of such drugs by providing more accurate predictions than using single risk markers alone, and therefore reducing uncertainty regarding the benefit-risk profile. For example, the predicted drug-induced risk change for clinically meaningful outcomes as calculated by the PRE score could be used as an endpoint in clinical trials. A threshold could be used to ensure that a predicted treatment benefit is also clinically meaningful. Such an endpoint would be reached earlier than conventional clinical endpoints such as ESRD, but would provide a more complete benefit-risk profile than using single risk markers alone.

Another way to implement the PRE score in regulatory decision making is to use it for adaptive licensing purposes, such as recently proposed by the EMA (20,21). For example, drugs could be granted a limited marketing approval based on their predicted effect as calculated by the PRE score. With its limited licensing, the drug could be given to a select number of patients, for example those at highest risk of developing the clinical outcome. The marketing license may then be extended or broadened by additional evidence derived from long-term clinical trials with endpoints that consist of clinically meaningful outcomes. Such post-marketing studies could be agreed upon between the regulator and drug developer at time of market access.

Tools such as the PRE score could be applied for several other purposes, beyond use as a surrogate endpoint in confirmatory clinical trials and regulatory decision making. For instance, the PRE score could be used by drug developers to guide early decision making regarding novel drugs. By providing an early assessment of the drug's likely effects on clinically meaningful outcomes, drug developers can make a more accurate, evidence-based decision with respect to whether continuation of the drug development process is viable. Another option is to use the score in clinical practice. By measuring and monitoring all relevant risk
markers and integrating them into the PRE score, doctors would get a more reliable estimate of whether their patient is likely to benefit from treatment. Therapy may be titrated to optimize a patient's response as calculated by the PRE score, instead of using single risk markers to determine the optimal dose. For each of these applications the score would require validation in a clinical practice setting.

For the future, the multiple parameter drug response score has potential to be used as an early predictor of drug effects on meaningful clinical outcomes and appears to be more accurate than using single markers alone, either on population level or on individual patient level. The prospective validation of the PRE score, that is currently underway, will ascertain the validity of this approach. The findings presented in this thesis therefore support continuation of the development of multiple parameter drug response scores and to seek ways to implement them in practice.
References


Summary and future perspectives


