Title: Genome-wide association study of stent thrombosis after coronary stenting - bioinformatic analysis and replication.

BACKGROUND: We witnessed a novel treatment of stenosis in coronary arteries in 1986, with the world’s first metallic stent implanted in a coronary artery of a human subject. As the advantages of stent implantation saw its evolution from a ‘bail out’ after complicated balloon angioplasty to a standard treatment strategy, a major limitation with stent therapy that was immediately recognized was an early acute vessel closure due to stent thrombosis (ST). Incidences of late ST (31st day onwards) still remain as a major concern for patients implanted even with the new generation drug-eluting stents (DES).

a) Severity of the problem: A recent retrospective registry by Tada et al., in 2013 including 18,334 subjects undergoing stent implantation has been the largest of its kind and compare the risk of stent thrombosis among patients treated with bare-metal stents (BMS), first-generation drug-eluting stents (G1-DES), and second-generation drug-eluting stents (G2-DES) for a period of 3 years. (1) The cumulative incidence of definite stent thrombosis at 3 years was 1.5% with BMS, 2.2% with G1-DES, and 1.0% with G2-DES. There was a significant excess risk of stent thrombosis at 3 years with G1-DES, driven by an increased risk of stent thrombosis events beyond 1 year. However, G2-DES were associated with a similar risk of stent thrombosis compared with BMS beyond 1 year. (1)

b) Socio-economic impact of ST: Stent thrombosis is a very serious clinical event typically resulting in ST-elevation myocardial infarction in the majority of cases and mortality rates that may be as high as 20–40%. In case of survivors of a ST, most of who result in suffering an MI, most registries report that thrombus aspiration and balloon angioplasty are frequently used with repeat stenting in around 30–50% of cases. In addition to the lives lost by this ST, an additional financial burden incurs on the survivors with a significant amount of lost disability-adjusted life years (DALYs).

c) Mechanisms leading to ST: Delayed and somewhat unpredictable healing of the DES-stented coronary artery segment associated with persistent late platelet accumulation and activation of coagulation appear to play an important role in the process of ST.

d) Traditional Risk factors: Although significant technical shortcomings in the index procedure [eg. stent undersizing or underexpansion, Malapposition (or incomplete stent apposition)] will more likely manifest as early stent failure, such factors can also play an important role in late ST especially after the time point of DAPT discontinuation. Patient-specific risk factors also remain important for late ST, in particular, reduced left ventricular function, diabetes mellitus and impaired response to ADP-antagonist therapy are associated with increased risk.

e) Genetic factors: The CARDIoGRAMplusC4D Consortium in 2013 published the largest meta-analysis of genetic studies to date assessing the impact of common variation on CAD risk which included a discovery and replication cohort. The analyses involving 63,746 CAD cases and 130,681 controls, identified 15 new risk alleles at genome-wide significance, bringing the total number of confirmed CAD susceptibility loci in individuals of European and south Asian ancestry to 45. The only large study on early ST by Cayla et al., 2011 identified 3 genes related to clopidogrel metabolism and platelet receptor function (CYP2C19, ABCB1, and ITGB3) found to be independently associated with early ST. While the problem of ST can very well be multifactorial and could be related to genetic markers of inflammation and immune response
processes in the endothelium, the aforementioned study only focussed on gene variations proven to be associated with thrombosis and platelet inhibition. Whereas several genomic loci have been associated with coronary artery disease and/or myocardial infarction, and to some extent for early ST, gene variations influencing the risk of late ST thus far remain elusive.

**HYPOTHESIS:** Genome-wide case-control association studies (GWAS) have proven very successful in the unbiased discovery of novel genes involved in cardiovascular diseases. There are indications suggest that stent thrombosis to is partially an inherited trait. The fact remains that no GWAS till date has been carried to discover these possible genetic determinants of ST. We thus hypothesized that late ST might have a complex genetic component that could be detected by GWAS.

**AIMS:** We thus aim to discover:

(i) Novel genes leading to late ST.
(ii) Genes involved in endophenotypes of late ST: i.e. endothelialization, peri-stent inflammation, fibrin deposition and platelet inhibition.

**METHODS:**

**Prior experience of the host:**
The host’s group (Deutsches Herzzentrum München - DHM) has extensive experience in morphological and functional research directed at stent thrombosis after coronary stenting. e.g. They coordinated a EU-sponsored (FP7) project called "PRESTIGE" (www.prestige-fp7.eu).

**Preliminary work:**
The DHM hosts the German MI Family Studies (PI: Prof. Dr. Schunkert) who was involved in the identification of most of the to date known genomic loci associated with CAD (CARDIoGRAMplusC4D Consortium). Currently, the DHM oversees a collection of more than 20,000 DNA/blood samples. 320 individuals with late ST have been identified. In these subjects, DNA has been isolated from stored whole blood. The samples underwent genotyping with the Infinium PsychArray BeadChip (Illumina), which incorporates more than 250,000 SNPs from the HumanCore BeadChip. Additionally, we have access to genotyping data from two large clinical trials (Acronyms - PLATO, SOLID) in the scope of a cooperation investigating genetic risk factors for stent thrombosis. Taken together, we have access to more than 800 cases of late ST after coronary stenting which will serve as a discovery cohort for the identification of genetic variants associated with the disease.

**Working program:**

1. **Analysis of the discovery cohort:**
   a) **Data extraction:** The discovery cohort data originates from the PLATO and SOLID trials as well as array genotyping of 320 individuals with late ST, which have been treated in the DHM. In summary, the discovery cohort consists of more than 800 patients with late ST secondary to coronary stenting. Matched controls will be identified in available GWAS data from the DHM, i.e. the German MI Family Studies 5 and 6 (n= 4,000). These studies include individuals with history of coronary stenting without stent thrombosis. Given the relatively low incidence of ST, we will match each case with four controls, i.e. matched patients receiving a stent but stayed free of ST. Data will be subjected to appropriate statistical analysis in order to unravel significant associations.
b) **Pre-imputation quality control (QC):** Pre-imputation QC will be conducted on SNP-level (MAF > 0.01; call rate in cases > 98%; call rate in controls > 98%; p-value for deviation from Hardy-Weinberg Equilibrium > $1 \times 10^{-6}$) as well as individual-level (call rate > 95%; sex check (based on X chromosome); population outlier removal (principal component (PC) or multi-dimensional scaling (MDS) analysis based on autosomal chromosomes); identity by descent (IBD) check (PI-HAT < 0.125); poor sample quality (heterozygosity rate out of range mean ± 5SD).

c) **Imputation:** Pre-phasing will be done using SHAPEIT2. IMPUTE2 will be used for imputation on the 1000 Genomes Phase I integrated haplotypes (NCBI build - b37, release date June 2014).

d) **Association analysis:** For the primary GWAS analysis, we intend to examine additive, dominant and recessive models by fitting logistic regression models that account for genotype imputation uncertainty. For all analyses, ancestry-informative PCs or other study-specific covariates (e.g. study centre) are included. Logistic regression model: ST (all ST cases vs. all controls) = SNP + age + gender + type 2 diabetes + prior myocardial infarction (MI) + ST-elevation MI leading to ST + stent type (bare metal stent; 1st, 2nd generation DES) + PCs.

2. **Replication:**
a) **Identification of replication cohort:** DHM already has blood samples stored for more than 20,000 individuals who underwent coronary stenting. An initial estimation reveals 320 individuals who suffered late ST eligible for analysis.

b) **Isolation of DNA:** Isolation of DNA from blood samples will be conducted. DNA samples will then be subjected to further analysis.

c) **Genotyping:** Genotyping of the discovery findings will be done using Taqman-genotyping on the ViiA7 platform (Life Technologies) at the DHM.

3. **Analysis, manuscript preparation and submission for publication:**
Data gained from the project will be submitted to high-ranking journals, e.g. Lancet, Circulation.

**EXPECTED RESULTS:** We expect that we will be able to identify several genes associated with late ST which will be able to enunciate the multifactorial nature of this problem. We expect that the responsible genes could be much more varied that those related to clopidogrel metabolism and platelet receptor function.

**DISSEMINATION AND IMPACT:** The project will allow us to identify genetic variants associated with late ST. This will be the first and the largest study focusing on the genetic determinants of late ST. Results of the project will be published in high-ranking peer-reviewed journals, e.g. Nature Genetics, New England Journal of Medicine, Lancet. The knowledge gained in the project has the potential to greatly impact the present preventive and therapeutic strategies for late ST.

**REFERENCES:**
TIMELINES AND OBJECTIVES:
This project is to be completed in a stipulated 1 year time. Our two phased study will follow the predefined timelines in order to complete the objectives of our study.

a) Analysis of “Discovery Cohort”:
The first 6 months will be used primarily in the analysis of discovery cohort which includes above explained phases such as “Data extraction”, Pre-imputation quality control, Imputation, Association analysis etc. Identification of replication cohort and Isolation of DNA from stored blood samples will also be done within the first 3 months.

b) Replication cohort:
The later 6 months of the project will focus on genotyping of the replication sample as directed the findings of the discovery cohort. Analysis will be performed and manuscripts ill be prepared for submission in favor of publication in high impact peer reviewed journals.

The comprehensive work program is duly depicted in the table below:

<table>
<thead>
<tr>
<th>Work program</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1a-Data extraction</td>
<td>2</td>
</tr>
<tr>
<td>1b-Pre-imputation quality control</td>
<td>3</td>
</tr>
<tr>
<td>1c-Imputation</td>
<td>4</td>
</tr>
<tr>
<td>1d-Association analysis</td>
<td>5</td>
</tr>
<tr>
<td>2a-Identification of replication cohort</td>
<td>6</td>
</tr>
<tr>
<td>2b-Isolation of DNA</td>
<td>7</td>
</tr>
<tr>
<td>2c-Genotyping</td>
<td>8</td>
</tr>
<tr>
<td>3-Analysis, Preparation of manuscript</td>
<td>9</td>
</tr>
</tbody>
</table>

KEY MILESTONES:
The first key milestone will occur after the first 6 months of the start of the project. We then may have the clarity in terms of the genes we need to focus upon as a result of the analysis of the discovery cohort.

The second key milestone will occur after the end of the project i.e. at 12 months when our analysis will be able to answer the key questions we aimed to pursue from the outset.