

Design of Class Structures and Inter Class Communication Interface

Computer Methods used in Non-linear Mixed
Effects Modeling

Pontus Pihlgren



UPPSALA
UNIVERSITET

**Teknisk- naturvetenskaplig fakultet
UTH-enheten**

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
<http://www.teknat.uu.se/student>

Abstract

Design of Class Structures and Inter Class Communication Interface

Pontus Pihlgren

The field of pharmacokinetics and pharmacodynamics describes the interaction between drugs and the human body from the time the drug is absorbed in the blood stream, distributed in the body, metabolized and finally excreted. The behavior can be described with nonlinear mixed effects models. NONMEM is a software for estimating parameters of such models and simulation of data. This thesis describes the design of Perl-speaks-NONMEM (PsN) which is a software framework for control of different aspects of NONMEM such as creating input, performing execution and output interpretation and presentation. The framework aims to simplify implementation of a range of statistical tools not featured in NONMEM. PsN includes the Bootstrap for measuring statistical accuracy and precision, the Case Deletion Diagnostics (CDD) for identifying influential individuals in medical studies and the Log Likelihood Profiling (LLP) for estimating asymmetrical confidence intervals for parameter estimates. PsN interfaces with various GRID systems: Sun Grid Engine, Plattform LSF and openMosix to enable parallel execution of NONMEM instances. The implementation of PsN was aided with custom built tools for generating class definitions, constructors and from UML diagrams. PsN also provides a set of command line interface tools that simplify usage for researchers without programming experience.

Handledare: Lars Lindbom
Ämnesgranskare: Olle Eriksson
Examinator: Anders Jansson
IT 08 004
Tryckt av: ITC

Populärvetenskaplig sammanfattning

Inom farmakokinetisk forskning utforskas vad som sker i kroppen efter ett läkemedel har administrerats. Ett vanligt läkemedel kan vara livshotande i stora doser och verkningslöst i små doser. Det är därför intressant att förstå de mekanismer i kroppen som avgör hur stor koncentration av ett administrerat läkemedel som kommer finnas i kroppen efter en viss tid.

För att göra det samlar man in data från studier med både människor och djur. Den datan behandlas för att förstås bättre. En hjälpmedel är att skapa en modell för hur en typisk människokropp påverkas av ett läkemedel. Det är också intressant är att förstå hur människokroppen påverkar läkemedlet genom olika typer av nedbrytning.

Den farmakokinetiska modellen är en matematisk beskrivning för hur läkemedlets koncentration ändras över tid. För att förbättra modellen utnyttjas olika automatiska sökningsalgoritmer som anpassar modellens parametrar för att passa den data som samlats in.

För att veta att modellen beskriver verkligheten bra nog så kan man utföra olika tester. Ett test går ut på att systematiskt utesluta individer ur datan och utföra den automatiska sökningen för varje ny uppsättning data. På så sätt kan man avgöra om en individ påverkar modellen för mycket eller kanske inte alls.

Det finns många parameterer som påverkar kroppens reaktion på ett läkemedel och kan inkluderas i en modell. Exempel är individens allmänna tillstånd som vikt, ålder och kön. Mer specifika variabler såsom om individen är sjuk, är rökare eller tar andra läkemedel samtidigt kan vara inflytelserika. Man kan då tänka sig automatiska metoder som lägger till en eller fler parametrar till modellen och utvärderar skillnaden mellan grundmodellen och den nya större modellen. Är den nya modellen bättre, enligt något mått, så kan metoden fortsätta med att lägga till parametrar, tills modellen inte kan förbättras.

Ett tredje moment som är intressant i modellbyggandet är att utvärdera modellens eller datan rent statistiskt. Att modellen visar att läkemedlets koncentration är i det intervall där det gör nytta kan vara meningslöst om modellen är osäker. Modellen är osäker om det också finns en sannolikhet att koncentrationens maximala eller minimala värde ligger utan för det nyttiga intervallet. För att hitta ett mått på modellens osäkerhet kan man utnyttja en metod som slumpmässigt väljer individer i den ursprungliga datan och skapar många nya varianter av datan. För varje variant utvärderar man sedan modellen och får en mängd olika resultat, med hjälp av dessa kan man beräkna ett mått av modellens osäkerhet.

Det är förhållandevis välkända metoder som används i flera olika forskningsområden, men dessvärre är implementationen, det faktiska utförandet, ganska specifikt till dessa områden. Inom farmakokinetisk forskning har man bra verktyg för de sökningsalgoritmer som anpassar modellen till insamlade data, men verktyg för statistisk utvärdering, automatisk modellbyggnad och automatisk manipulation av data har inte funnits tillgängliga.

Syftet med detta examensarbete är att skapa en grund för implementation av dessa olika verktyg. Genom att samla funktioner för hantering av de filformat som används inom området, funktioner för resultatpresentation och kombinerad av olika metoder i ett programbibliotek siktar arbetet på att förenkla för forskare inom farmakokinetik att implementera traditionella och nya metoder.

För att ytterligare underlätta implementation av olika metoder så har man valt att skapa programmet med hjälp av ett grafiskt verktyg, Dia, utifrån vilket datorkod kan generas. Dock saknades den översättare från Dias filformat till datorkod varvid ett sådant också utvecklats inom ramen för detta arbete.

De metoder som beskrivits ovan kräver alla en mängd beräkningsarbete vilket kan ta mycket lång tid för en enskild dator att utföra. Därför implementerades en funktion för att utföra dessa beräkningar på olika typer av så kallade datorkluster eller GRID-miljöer som är en samling av

datorer tillgängliga för den här typen av beräkningar.

Slutligen har ett par av metoderna ovan implementerats i programbiblioteket och för de mest efterfrågade metoderna har ett användarvänligt kommandoradsgränssnitt utvecklats för att öka tillgängligheten bland de forskare som inte besitter programmeringskunskaper.

Resultatet av arbetet har varit lyckat, programvaran har använts för att implementera ett handfull olika metoder som används av flera akademiska institutioner och företag i läkemedelsindustrin. Resultatet har lett till publikation av en artikel om programbiblioteket i en vetenskaplig tidskrift. Det har även fått ett gott motagande på kurser för avancerad modellering där det använts som hjälpmedel i undervisningen.

Table of Contents

| | | |
|-------|--|----|
| 1 | Introduction | 6 |
| 1.1 | NONMEM | 6 |
| 1.2 | Perl-speaks-NONMEM (PsN) | 7 |
| 2 | Purpose | 8 |
| 3 | Method..... | 8 |
| 4 | PsN – In detail | 8 |
| 4.1.1 | Model..... | 9 |
| 4.1.2 | Data | 10 |
| 4.1.3 | Output | 10 |
| 4.1.4 | Multiple Run Script | 10 |
| 4.2 | Existing tools | 10 |
| 4.2.1 | Bootstrap | 10 |
| 4.2.2 | Case Deletion Diagnostics..... | 11 |
| 4.2.3 | Gam42 | 11 |
| 4.2.4 | LLP..... | 11 |
| 5 | Results | 12 |
| 5.1 | Language choice | 12 |
| 5.2 | UML diagrams and code generation..... | 12 |
| 5.2.1 | Dia | 12 |
| 5.2.2 | dia2code..... | 13 |
| 5.2.3 | Additional development software | 14 |
| 5.3 | Class hierarchy..... | 15 |
| 5.4 | Tool interconnection | 16 |
| 5.5 | Grid execution..... | 17 |
| 5.5.1 | NorduGrid | 17 |
| 5.5.2 | Platform Load Sharing Facility, and general grid communication | 18 |
| 5.5.3 | Sun Grid Engine | 18 |
| 5.6 | Resuming broken runs | 18 |
| 6 | Discussion..... | 19 |
| 6.1 | Memory management | 19 |
| 6.2 | UML Diagrams | 19 |
| 7 | Future plans | 20 |
| 8 | Conclusion..... | 20 |

