“I need the numbers now!”
Programming ad hoc requests in a project centric environment

Michaela Kohring

Phuse SDE Frankfurt (Main), 10th of May 2016

Boehringer Ingelheim

Agenda

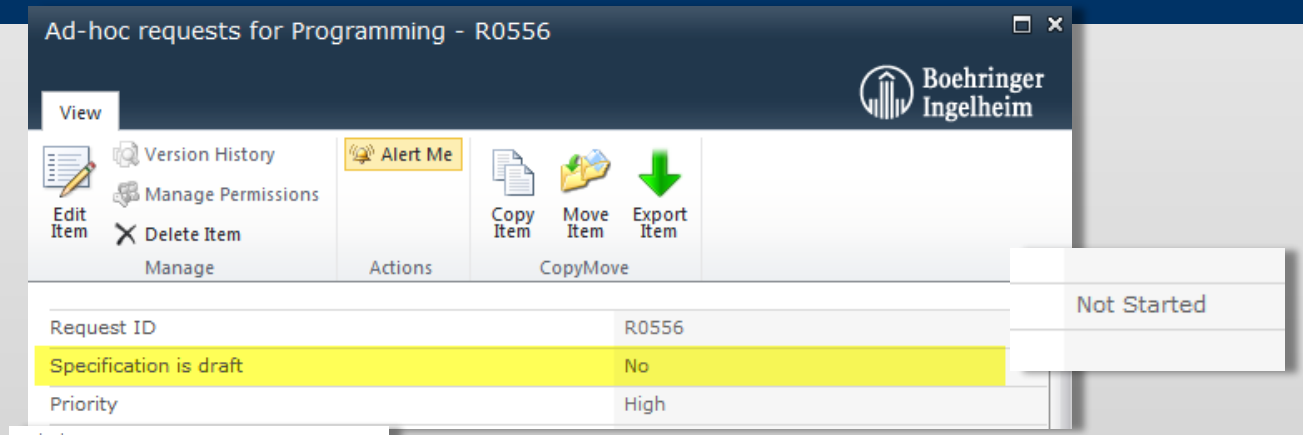
- Challenges of ad hoc requests;
- Specification;
- Centralized data preparation;
- Central TFL library;
- Implementation;
- Validation;
- Oversight;

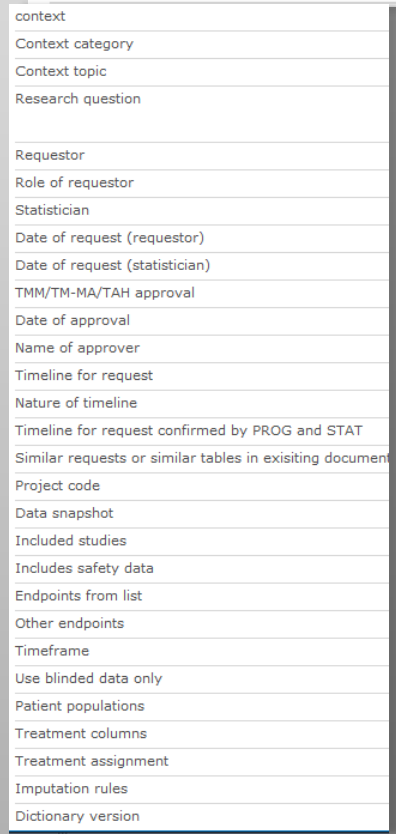
Challenges of an ad hoc request

- Short timelines;
- Explorative analyses;
- Changing priorities;
- Standard packages (CTR etc.) need to be delivered in parallel;
- Topics may vary from request to request depending on the purpose (authority questions, publication, re-imbursement etc.);
- None of your customer has time to wait for the results;

Ad hoc request

- centralized specification

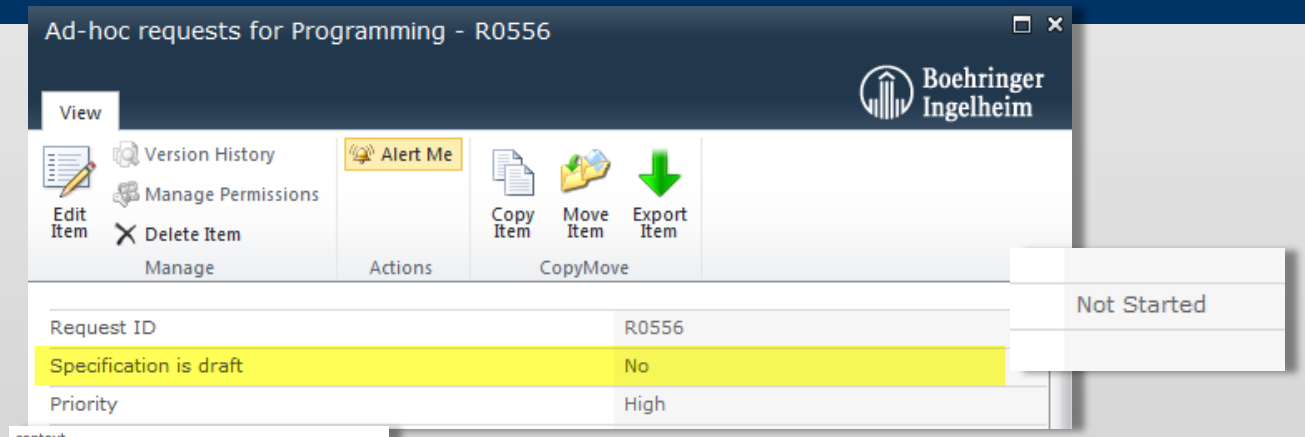


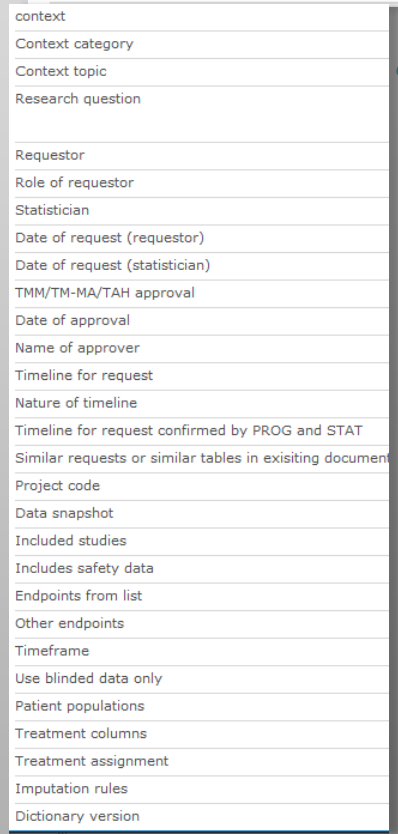


- The specification for ad hoc requests is stored in a central location, shareroom list including mandatory information;
- Dedicated programmers track incoming new requests;
- Each request is assigned to a responsible programmer;
- All information around the request is maintained in the shareroom entry, e.g. first / second programmer, status of ad hoc (in progress, validation outstanding, etc.);

Ad hoc request

- centralized specification





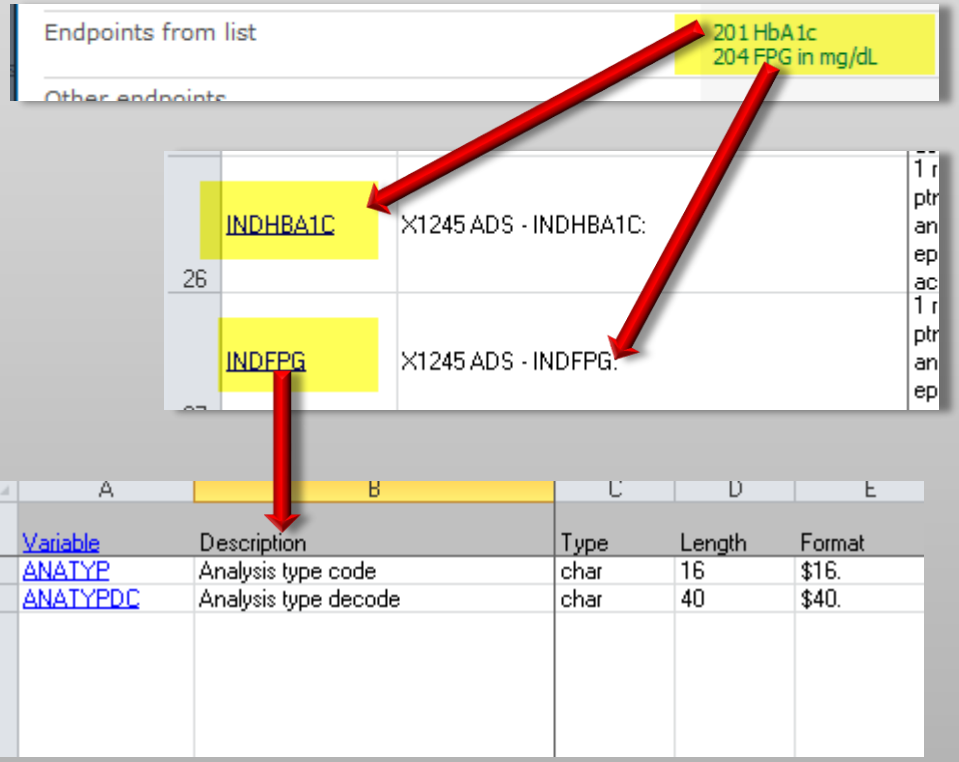
- PROs:
  - Status tracking is centralized, i.e. progress of requests can be followed up with simple tools;
  - Workload of programmers can be figured out;
  - Information is not spread over different trial and projects, documents, folders, etc.;

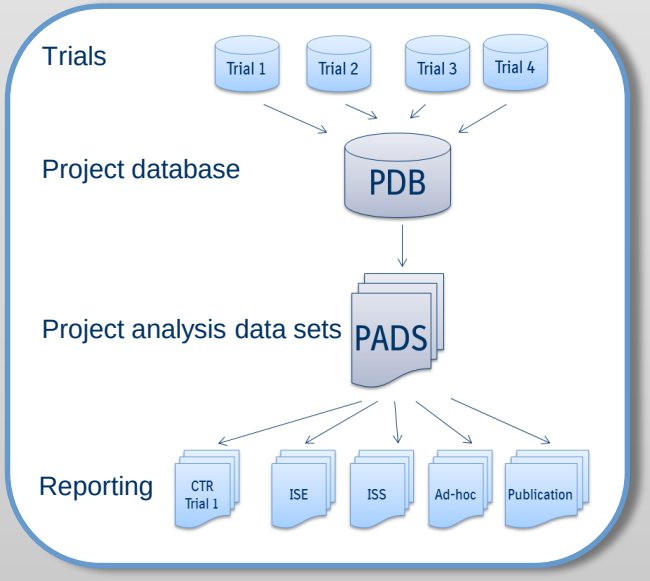
Ad hoc request

- centralized data preparation

In Progress

- PCA (project centric approach) – all data derivation will be provided by the project team for trials and projects in analysis datasets (ADS);
- All information is collected in a central project ADS plan;
- Each derivation done at any time point can be found in one document;



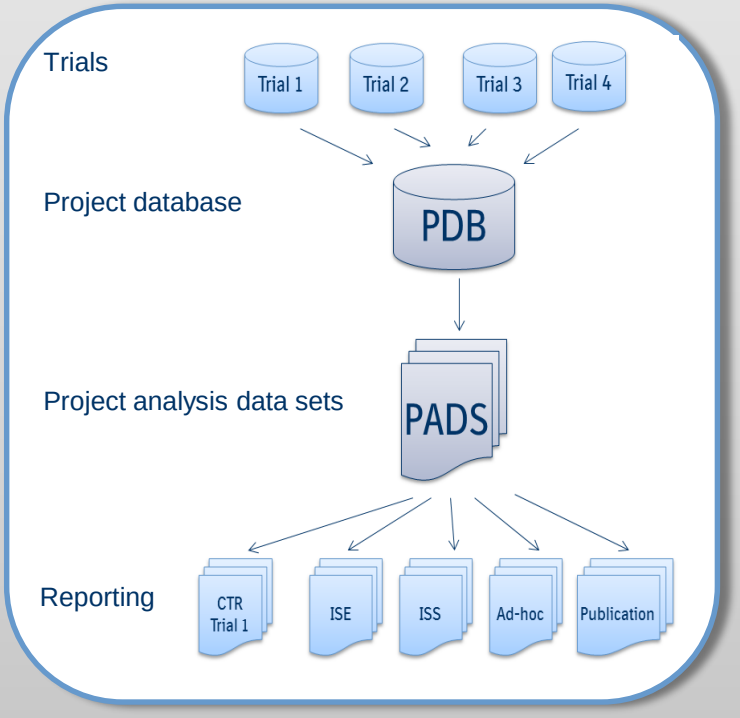


Ad hoc request

In Progress

- centralized data preparation

- PROs:
  - All derivation ever implemented is stored in a central place;
  - In case a derivation, population set etc. is not available for a specific study, the programs can be easily amended for additional studies;
  - Variables are stable throughout the whole project – the derivation of e.g. baseline eGFR is always the same and doesn't differ between the trials;
  - Same holds true for variable naming, a specific variable will always contain the same information throughout the project;

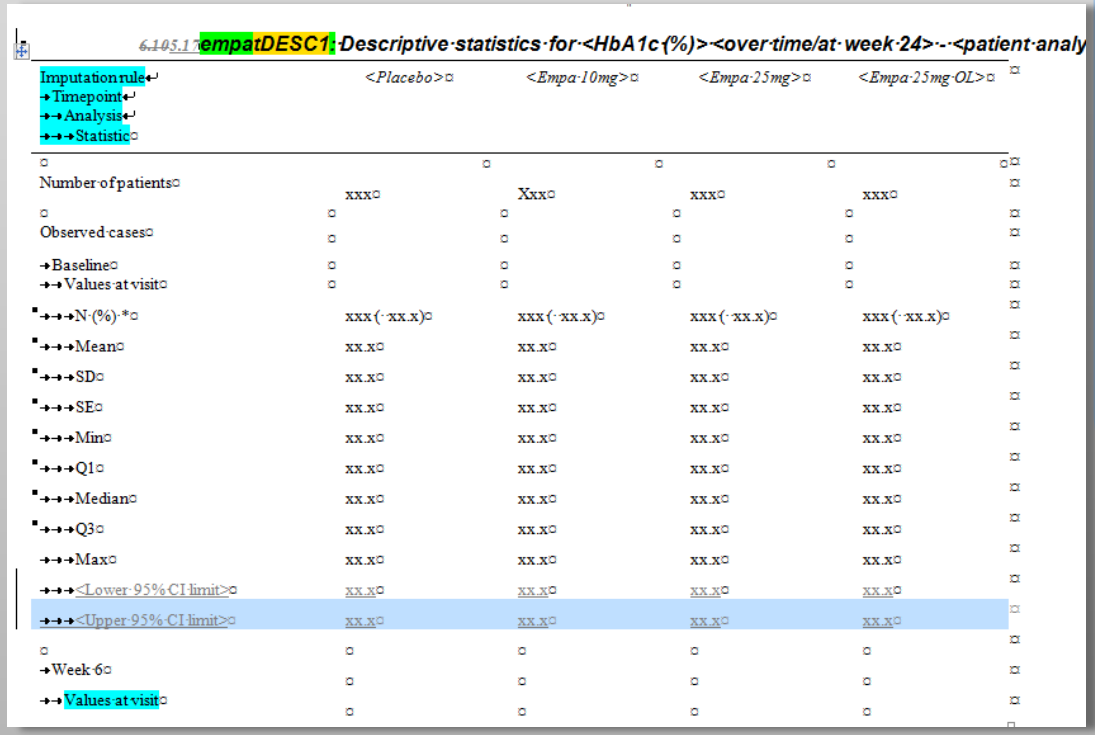


Ad hoc request

- central TFL library

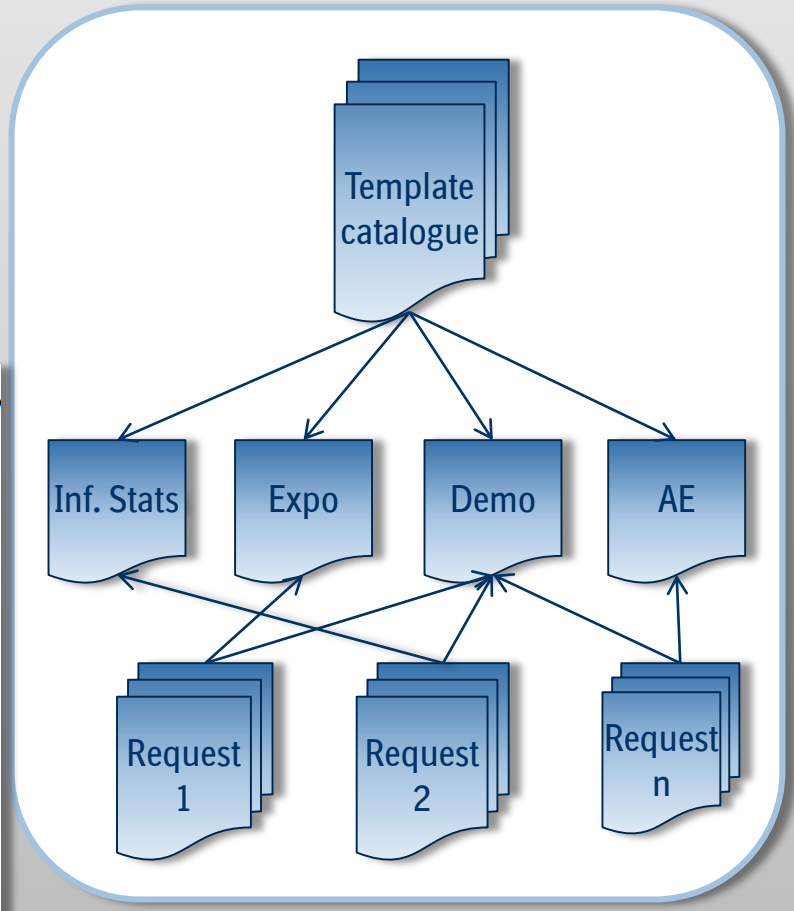
In Progress

- Inside PCA all trials and projects use the same TFL shell catalogues;
- Customers (statisticians) can select, what they need, from the central documents;



6.105.1 empatDESC1: Descriptive statistics for <HbA1c (%)> <over time/at week 24> - <patient analy

Imputation rule → Timepoint →→ Analysis →→→ Statistic	<Placebo>	<Empa 10mg>	<Empa 25mg>	<Empa 25mg OL>
Number of patients	xxx	Xxx	xxx	xxx
Observed cases				
→ Baseline				
→→ Values at visit				
→→→ N (%) *	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
→→→ Mean	xx.x	xx.x	xx.x	xx.x
→→→ SD	xx.x	xx.x	xx.x	xx.x
→→→ SE	xx.x	xx.x	xx.x	xx.x
→→→ Min	xx.x	xx.x	xx.x	xx.x
→→→ Q1	xx.x	xx.x	xx.x	xx.x
→→→ Median	xx.x	xx.x	xx.x	xx.x
→→→ Q3	xx.x	xx.x	xx.x	xx.x
→→→ Max	xx.x	xx.x	xx.x	xx.x
→→→ <Lower 95% CI limit>	xx.x	xx.x	xx.x	xx.x
→→→ <Upper 95% CI limit>	xx.x	xx.x	xx.x	xx.x
→ Week 6				
→→ Values at visit				



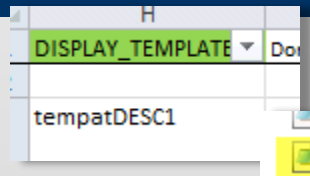
Ad hoc request

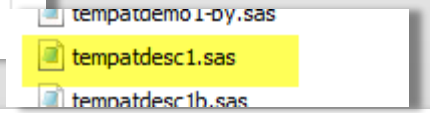
In Progress

- central TFL macro library

Boehringer Ingelheim

- Due to PCA in ADS standardized macros for output creation can be used;
- Programs to create outputs are set up as macros;
- The programs are stored in a central place;
- Standardized setup – data building, parameter names, result datasets for validation;
- In case an adaption is needed, this will be straight forward due to the standardized way of programming in the template macros;
- The use of macros is supported by technical documentation;






```
%macro tempatDESC1( _inpop  = ads.popu,
                    _whrpop = popu='TS' and popuny=1,
                    _intrt  = ads.gentrt,
                    _whrtrt = ,
                    _trt    = atrsort atrslbl,
                    _inind  = ads.indhba1c,
                    _whrind = ,
                    _dep    = ept,
                    _time   = tptno tptnodc,
                    _trf    = epttrf epttrfdc,
                    _whrcf  = anatyp ? "LOCF" and tptno>0 and epttrf=0,
                    _whrres = (anatyp ? "-IR" or anatyp ? "-AD") and tptno>0 and epttrf=0,
                    _whrdis = anatyp ? "-AD"  and tptno>0 and epttrf=0,
                    _whrah  = 0,
                    _se     = Y,
                    _ci     = N,
                    _p1090  = N,
                    _p3366  = N,
                    _alpha  = 0.05,
                    _dp     = 1,
                    _anades = ,
                    _trfdes = ,
                    _patient= Patient,
                    _whrout = ,
                    _ft1    = ,
                    _ft2    = ,
                    _ft3    = ,
                    _ft4    = ,
                    _ft5    = ,
                    _ft6    = ,
                    _ft7    = ,
                    _ft8    = ,
                    _rtf    = Y,
                    _inbas  = ,
                    _outno  = ,
                    _debug  = N );
```

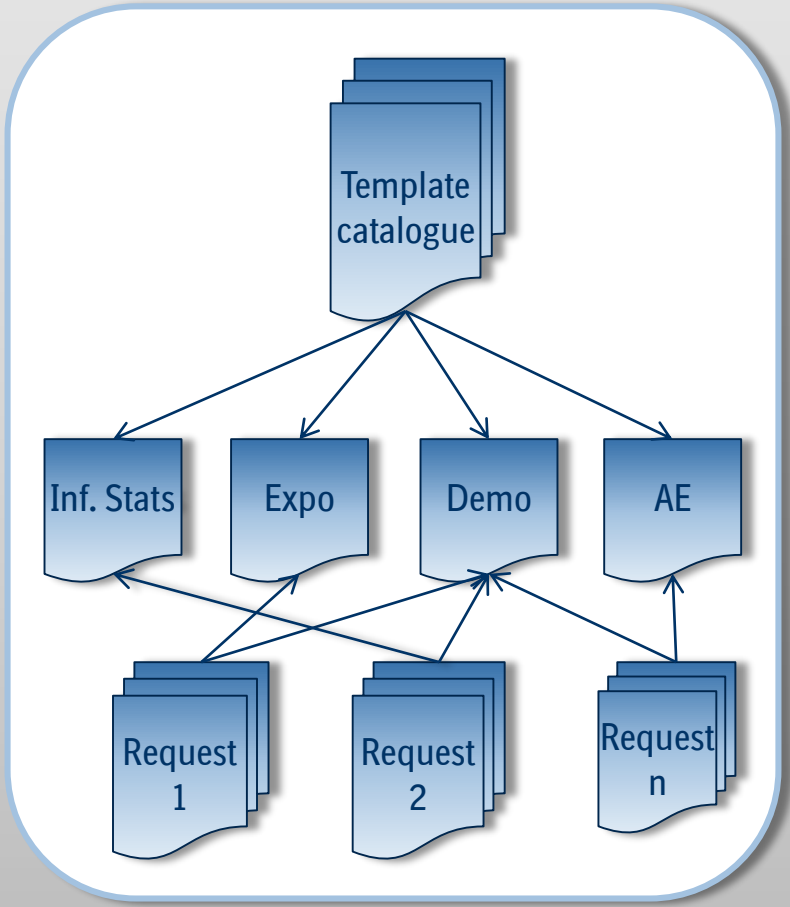
Title:	Descriptive statistics for time difference between the date of last available haematocrit (%) OC-A
Program (link):	ru0455\labtdesc1001.sas
Output #:	3
Output (link):	ru0455\labtdesc1001-3.lst
Macro call:	%tempatdesc1(_WHRPOP = popu='TS' and popuny=1 ,_WHRTRT = analno=2080010 and atrsort in ('B00100' 'C00100' 'C00250' 'E00035') ,_ININD = inlab_c ,_WHRAH = 0 ,_WHRCF = anatyp ? "LOCF" and tptno>0 and epttrf=0 ,_WHRDIS = anatyp ? "-AD" and tptno>0 and epttrf=0 ,_WHRRES = (anatyp ? "-IR" or anatyp ? "-AD") and tptno>0 and epttrf=0 ,_OUTNO = 3 ,_WHROUT = BY4ORD in (1.3 1.5 3.5) ,_FT1 = Baseline data of Haematocrit not considered.)

Ad hoc request

- central TFL macro library

In Progress

- Pros;
  - Programs are of high quality, double programmed several times on different data;
  - Easy to use due to standard parameters;
  - Implementation can be supported by standard tools;
  - Validation can be streamlined by categorization of effort, e.g. double programming, consistency check, source code review ;
  - Less time for discussions on layout needed;

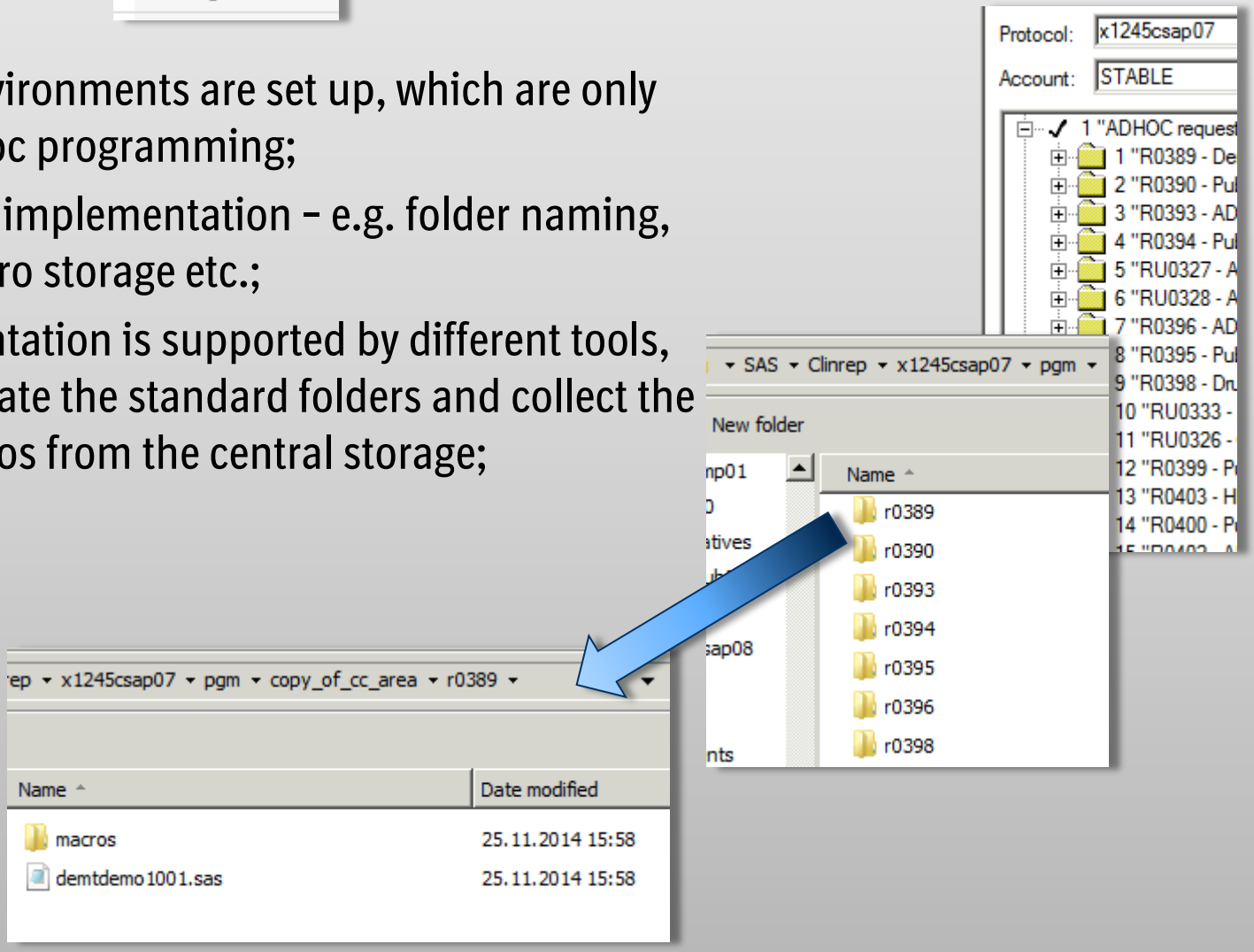


Ad hoc request

- implementation

In Progress

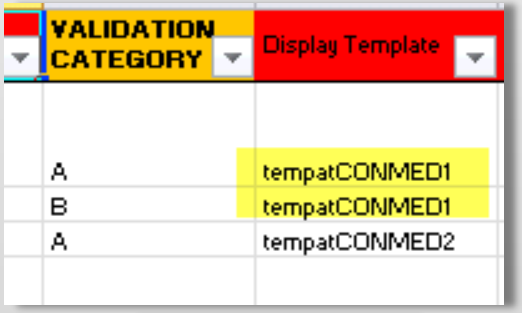
- Dedicated environments are set up, which are only used for ad hoc programming;
- Standards for implementation – e.g. folder naming, template macro storage etc.;
- The implementation is supported by different tools, which e.g. create the standard folders and collect the relevant macros from the central storage;

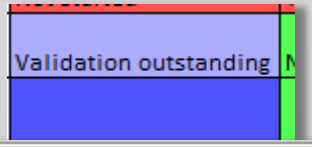


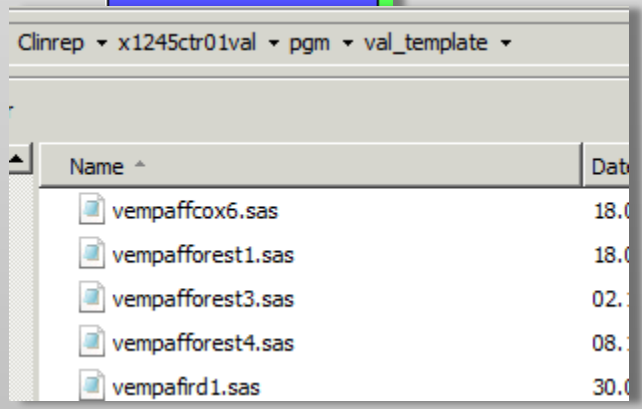
Ad hoc request - validation

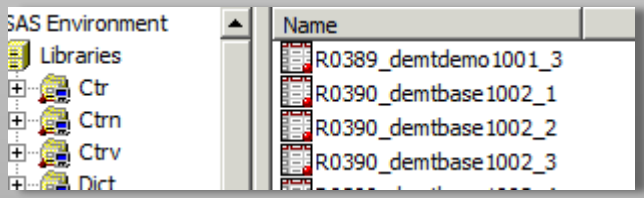
Validation ongoing

- Standard setup – double program at least each template once (category A), for remaining outputs source code review and consistency checks are performed;
- On demand the process can be accelerated by source code review only;
- Validation starts after draft approval by the statistician (standard);
- Like for output template programs a large number of validation programs is available;
- Validation programs follow the same naming convention like first programs, so the connection between matching programs can be easily made;
- Double programming is supported by result datasets stored in a standard library;





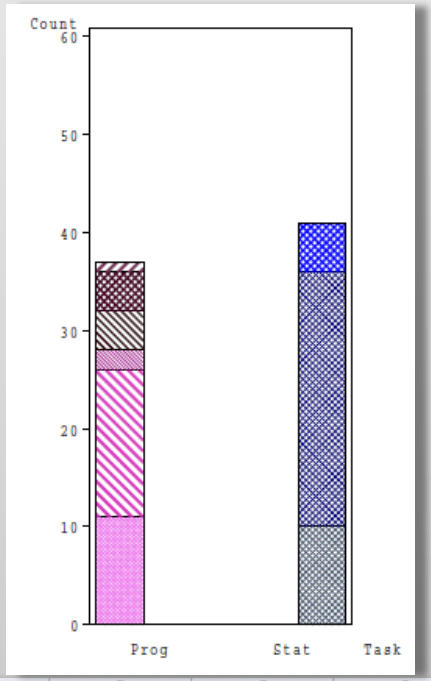


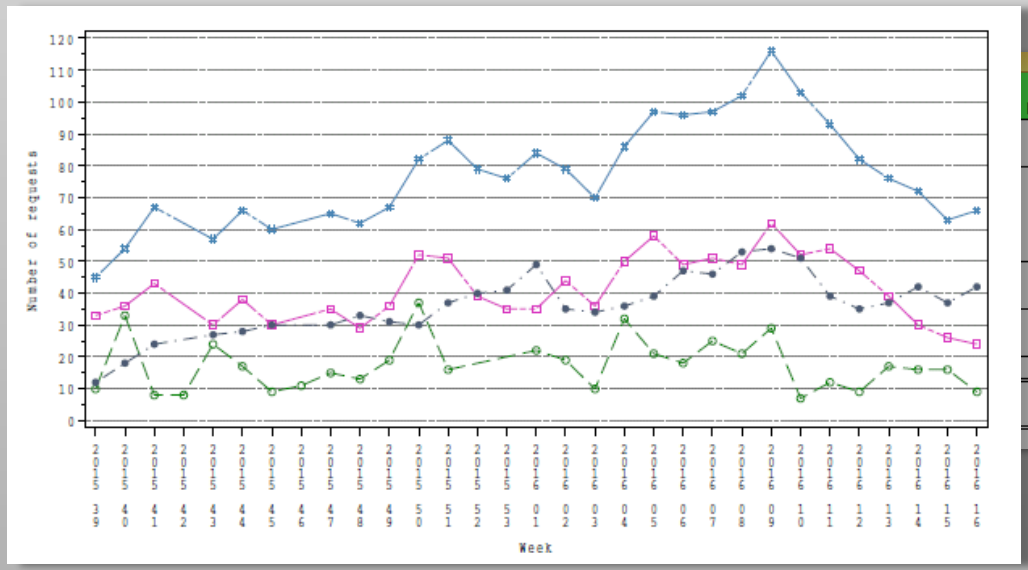


Ad hoc request

- oversight

- Standard setup supports implementation of tracking tools based on sharepoint lists;
- Most ad hoc programming is covered by a dedicated project team;
- Centralized tracking more easily shows, what's going on, how is the status, do we need prioritization or additional ressources;
- You learn to know, what is feasible to achieve with your standard ad hoc team ;



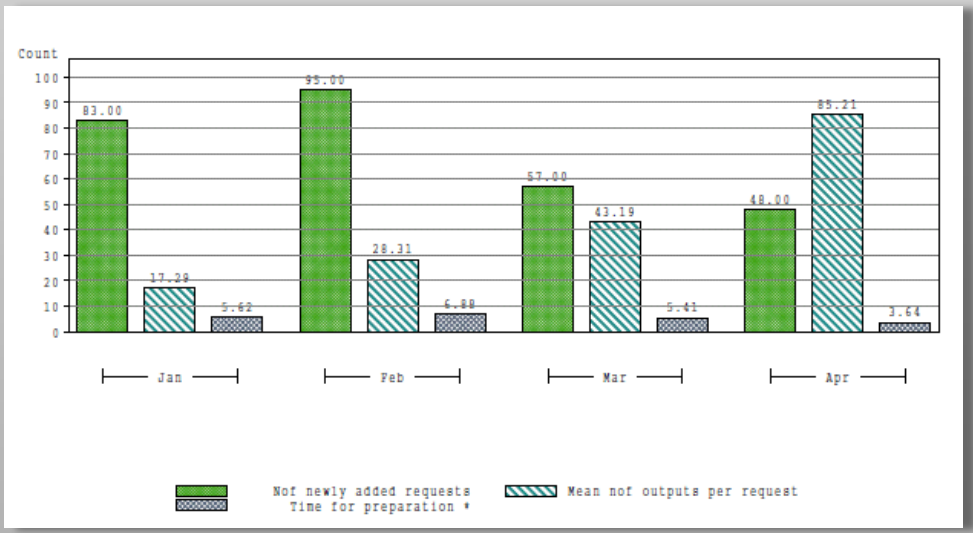


A	B	C	D	E	F	
PRIOSOF	timeline	Status	Draft specification	SHORTCONT	Request ID	
	2016-04-20	In Progress	No	Adcom	RU0865	
1	2016-04-20	In Progress	No	Other	RU0862	
1	2016-04-20	Draft sent out	No	Other	RU0867	
1	2016-04-20	Not Started	No	Authority	RU0868	
2	2016-04-20	Draft sent out	No	Adcom	RU0852	
1	2016-04-24	In Progress	No	Other	RU0866	

Ad hoc request

- conclusion

- In PCA there are experts for each an every task, this leads to quick turnarounds, but you need to think of backup solutions to avoid bottlenecks due to unavailability of programmers;
- A standardized setup helps to bring in new programmers with less effort, starting with the more easy tasks like output implementation;
- Implementing PCA successfully needs frontloading, but it will help you in more stressful times later on;
- Experienced programmers can create several hundred outputs in less than one day;



Backup slides

Boehringer Ingelheim

Ad hoc request

- some numbers

Boehringer Ingelheim

- Over 1400 ad hoc requests in four years
- Standard wise 20 to 30 requests a month, up to ~ 90 within a month;
- Standard setup – dedicated programmers for ad hoc work, but they are not limited, too, escalation model available;
- ~ 70 ADS – from high performer (treatment, population) to low runners (MTT);
- ~ 300 different templates for all type of analyses;
- ~ 95 – 98 % of requests can be covered with PCA, this reduces the pressure and frees capacity for the remaining 2 %;