

## Protokoll

NGFN Qualitätsmanagement-Workshop, 05.11.03, 13.00-18.00 Uhr, Bonn

**Teilnehmer:** Frauke Behrens, RZPD  
Thomas Bettecken, Genotypisierung  
Florian Burckhardt, GEM  
Helmut Fuchs, Mausklinik  
Thomas Häupl, Klinik Rheuma, Infektion  
Wolfgang Huber, Datenanalyse/Statistik  
Anja Hügel, Projektmanagement NGFN  
Claus Hultschig, Microarrays  
Ludger Klein-Hitpass, Expressionsanalyse (Affymetrix)  
Andreas Koch, SNP-Genotypisierung  
Chris Lawerenz, Datenanalyse  
Michael Majores, Vertretung Albert Becker Microarray  
Ulrike Mrwa, Qualitätsmanagementbeauftragte PT-DLR  
Reinhold Schäfer, Expressionsanalyse  
Martin Strödicke, Proteinanalyse  
Holger Sülthmann, Microarray  
Uta Strasser, Projektmanagement NGFN  
Dieter Weichenhan, Expression Profiling  
Stefan Wiemann, cDNA Generierung  
Thomas Wienker, GEM

Um im NGFN das wissenschaftliche Potential der erheblichen und z.T. in systematischen Ansätzen generierten Daten in vollem Umfang bereichsübergreifend, effizient und langfristig erschließen zu können, bedarf es einer bestmöglichen Vergleichbarkeit der Ergebnisse und der generierten Daten. Dies zu erreichen erfordert Maßnahmen im Bereich des Qualitätsmanagements (QM).

Ziel eines QM im Rahmen des NGFN sollte die Gewährleistung eines durchgängig hohen Qualitätsniveaus für die im NGFN generierten Ressourcen und Daten und ihre Vergleichbarkeit über das gesamte Forschungsnetzwerk hinweg sein. Neben Definition und Erhalt von Qualitätsstandards sollte auch deren kontinuierliche Verbesserung Ziel des QM sein.

In einem ersten Workshop mit den benannten Qualitätsbeauftragten im NGFN und zusätzlichen NGFN Mitgliedern mit QM-Erfahrungen wurden Erwartungen, Maßnahmen und Bedenken gegenüber einem Qualitätsmanagement im NGFN erörtert.

Bei allen Anwesenden bestand Einigkeit darüber, dass das Thema Qualitätsmanagement für das NGFN in Zukunft von zunehmender Bedeutung sein wird. Allerdings wurde in der Eingangsdiskussion auch klar, dass eine bis ins Detail festgelegte Standardisierung für alle etablierten Methoden im NGFN nicht zu verwirklichen ist und auch im Hinblick auf konkrete Arbeitsabläufe im Labor nicht sinnvoll erscheint. Eine gewisse Flexibilität muss im Sinne der Wissenschaft gewährleistet bleiben.

Dagegen sollte die Aufstellung von einem Minimum an Standards und eine klare Definition von Schnittstellen für Material- und Datenübergabe angestrebt werden und als „Spielregeln“ für die Anwendung im NGFN festgelegt werden, zu deren Einhaltung sich alle Partner im Netzwerk verpflichten.

Neben einer Definition von Schnittstellen beim Austausch von Material und Information, wurde auch die Transparenz der eingesetzten Protokolle zur Aufarbeitung gefordert. Für die Endanwender im NGFN ist die Aufarbeitung oft entscheidend, daher ist die Mitteilung des verwendeten Protokolls notwendig.

Beim Datenmanagement und bei der Datenanalyse sind die Übertragbarkeit der Datenformate und die Nachvollziehbarkeit der Analyse die wichtigsten Kriterien. Dem Aspekt der Übertragbarkeit kann durch den Einsatz von Standards Rechnung getragen werden. Für die Nachvollziehbarkeit ist die Dokumentation folgender Fragen erforderlich:

- Was ist das biologische/medizinische System?
- Wie operiert das Assay und das Datenaufnahmesystem?
- Welche Primärdaten wurden erhoben?
- Wie wurden die daraus gezogenen Schlussfolgerungen erhalten? (Audit-Trail der statistischen und bioinformatischen Analyse)

Bei der Erhebung klinischer Daten muss die Vergleichbarkeit durch einen Minimalsatz an erhobenen Patientendaten im NGFN gewährleistet sein, die krankheitsspezifisch modifiziert sein können und inhaltlich ggf. mit Standards in den medizinischen Kompetenznetzen abgeglichen werden.

Für klinische Studien wie auch für Experimentserien muss statistisch belegt werden, dass das Design und der Stichprobenumfang für die Ziehung relevanter Schlussfolgerungen geeignet sind.

Wesentlich für die Zuführung klinischer Daten in ein System und auch die Weitergabe von klinischen Informationen ist die Gewährung der Datensicherung und des Datenschutzes (s. auch Vorgaben der Ethikkommissionen).

Von den Teilnehmern vorgeschlagen wurde auch die Einrichtung einer QM-Internet-Plattform für das NGFN. Auf dieser Plattform wären die Ansprechpartner für die einzelnen methodischen Bereiche zu finden und evtl. die eingesetzten Protokolle einzusehen. Weiter könnte hier die Möglichkeit geboten werden wissenschaftliche Fragestellungen im NGFN zu diskutieren und dazugehörige Daten zu vergleichen. Es wurde der Vorschlag gemacht, ein E-mail-Forum einzurichten.

Anhand einer Auflistung verschiedener Themenbereiche und den dazugehörigen im NGFN etablierten Methoden, wurden die für das NGFN wichtigsten QM-relevanten Themen identifiziert. Diese Themen sollen nun in weiteren Arbeitstreffen jeweils in kleineren, inhaltlich mit allen notwendigen Kompetenzen bestückten Gruppen bearbeitet werden.

Ziel ist hierbei die Ausarbeitung von Minimalstandards in dem jeweiligen Bereich, die eine Nachvollziehbarkeit der Daten und Interpretation gewährleisten sollen.

Die Ergebnisse der Arbeitsgruppen sollten in einem weiteren Schritt in einem Katalog zusammengestellt werden. Zur Einhaltung der in diesem Katalog formulierten Standards und Schnittstellen sollen sich die Antragsteller aus NGFN1 für die Folgeanträge NGFN2 freiwillig verpflichten. Dieser Katalog könnte auch Element in allen Neuanträgen aus dem NGFN1 sein. Frist für die Antragstellung NGFN2 ist der 30. Januar 2004.

Die u.g. jeweiligen Arbeitsgruppen können jederzeit um weitere interessierte Teilnehmer erweitert werden. Vor allem sollte jede Gruppe, falls noch nicht integriert, einen Bioinformatiker zu ihren Treffen hinzuziehen.

Die Arbeitstreffen werden von den einzelnen Arbeitsgruppen selbst organisiert. Informationen (Ort, Zeit) über die stattfindenden Treffen sollten aber in jedem Falle an das PM weitergegeben werden. Wo nötig kann hier auch gerne Unterstützung seitens des PM angeboten werden.

Für folgende Themenbereiche wurden auf dem Workshop Arbeitsgruppen zusammengestellt:

#### Klinik

1. Patientenrekrutierung
2. klinische Dokumentation
3. Probengewinnung und -lagerung

**Arbeitsgruppe:** Bettecken, Burkhardt, Häupl, Lawerenz, Weichenhan, Wienker

#### Datengenerierung, Techniken

1. Microarrays

**Arbeitsgruppe:** Becker, Behrenz, Huber, Hultschig, Klein-Hitpass, Lawerenz, Sültmann

2. Genotypisierung

**Arbeitsgruppe:** Bettecken, Koch, Saar, Wienker, andere

3. Sequenzierung

**Arbeitsgruppe:** Blöcker, Nürnberg, Yaspo, andere

Keine Teilnehmer des Workshops, das PM wird diese Arbeitsgruppe informieren

4. cDNA library Produktion, Cloning, Ressourcen

- Definition generierter Ressourcen aus dem NGFN (RZPD, cDNA Konsortium, andere)
- Definition von Qualitätsstandards für diese Ressourcen
- Definition von notwendigen Informationen zur Weitergabe der Ressourcen

**Arbeitsgruppe:** Behrenz, Wiemann

5. Funktionsassays, Zellkulturen

- Definition von Standards für eingesetzter Ressourcen in Assays (Klone , ORFs, siehe auch 4.)
- Definition von Standards für Assay-Systeme (Auswahl zellulärer Systeme, Relevanz für die Bio-Medizinische Fragestellung, etc.)
- Definition von Standards für das Design der Experimente und die Daten-Analyse (Bioinformatik, Statistik)
- Definition von Standards für die Weitergabe von Primärdaten und deren Interpretation

**Arbeitsgruppe:** Huber, Schäfer, Wiemann

6. Proteinanalyse

- Definition von Standards für eingesetzte Ressourcen in den Experimenten (Klone, Proteine, ORF, Proteinreinheit, Mass Spec Fingerprinting, etc.)
- Definition von Standards für Analyse Algorithmen, das Design der Experimente und die Daten-Analyse (Bioinformatik, Statistik)

**Arbeitsgruppe:** Huber, Korf, Strödicke, Wiemann

#### Datenmanagement (auch allen o.g. Themen zugehörig)

1. LIMS - Nutzen und Zertifizierung von Qualitätsmerkmalen im Bereich Laborinformationssysteme
2. Nomenklatur für Material und Daten, gemischte Nomenklatur zu Fachbereichen
3. File formats for data exchange (e.g. XML DTDs)
4. LSID - Nutzung des life science identifier in einem System
5. MAGE-ML - Nutzen von standardisierten Schnittstellen zu allen Datenbanken wie auch zu LIMS-Systemen entsprechend der MGED-Kriterien (Microarray Gene Expression Data Society)
6. Ontologien (MGED-Ontologien, weitere Ontologien mit internationalem Standard, NGFN-Ontologien (spezifische, für einzelne Netze, für bestimmte Forschungsschwerpunkte

bestehende controlled vocabularies): systematische Erfassung und zur Verfügung stellen in einem Datenbanksystem, Einbau von core und extended Ontologien

7. Autorisierung und Authentifizierung: Anmeldung, Zuweisen von Datenbereichen zu privaten Bereichen / öffentlichen Bereichen
8. Gewährung des Datenschutzes und Datensicherheit

#### Datenanalyse (auch allen o.g. Themen zugehörig)

- Definition von Standards für Design der Experimente und Daten-Analyse: Statistische Signifikanz, „Ausreißer“-Erkennung, Relevanz der biologischen und medizinischen Interpretation
- Reproduzierbarkeit von Messungen, Positiv- und Negativkontrollen, Kalibrierung
- Unabhängige Überprüfung
- Reproduzierbarkeit der Datenanalyse (Bereitstellung von nachvollziehbaren audit-trails)

Die genannten Diskussionspunkte zum Datenmanagement werden an die beiden Bioinformatik task forces zur Auswertung von molekularen bzw. patientenbezogenen Daten weitergeleitet. Es soll abgeklärt werden, inwieweit o.g. Punkte schon in den task forces bearbeitet wurden oder noch werden bzw. ob die Themen zur Bearbeitung übernommen werden können.

Es wäre sinnvoll, wenn Vertreter der Bioinformatik aus diesem Workshop (Lawrenz, Huber) die o.g. Punkte an die Bioinformatik task forces weiterkommunizieren würden.

In den o.g. Arbeitsgruppen sollten dennoch, unabhängig von den Bioinformatik task forces, Qualitätsmanagement-Aspekte zu Datenmanagement und der Datenanalyse spezifisch für die Plattformen und krankheitsorientierten Netze diskutiert werden. Daher ist es wichtig, zu allen Treffen einen Bioinformatiker/Statistiker hinzuzuziehen.