Aus dem Institut für Pharmakologie und Toxikologie
Geschäftsführender Direktor: Prof. Dr. Frank Czubayko
des Fachbereichs Medizin der Philipps-Universität Marburg

Entwicklung eines systematischen Verfahrens zur Untersuchung von indikationsübergreifenden Signalwegen und Zielmolekülen für pharmazeutische Forschung und Entwicklung

Inaugural-Dissertation zur Erlangung des Doktorgrades der Humanbiologie
(Dr. rer. physiol.)
dem Fachbereich Medizin der Philipps-Universität Marburg
vorgelegt von

Johanna Fehr aus Friedberg/Hessen
Marburg, 2009

Angenommen von Fachbereich Medizin der Philipps-Universität Marburg am 16.11.2009

Gedruckt mit Genehmigung des Fachbereichs

Dekan: Prof. Dr. Matthias Rothmund

Referent: Prof. Dr. Frank Czubayko

1. Korreferent: Prof. Dr. Gerhard Klebe

Für meine Eltern

Inhaltsverzeichnis

1 ZUSAMMENFASSUNG 1

2 SUMMARY 3

3 EINFÜHRUNG 5

3.1 GRUNDGRÖßEN VON PHARMAZEUTISCHER FORSCHUNG UND ENTWICKLUNG 5

3.1.1 Substanzen und Medikamente 5

3.1.2 Indikationen, Indikationsgruppen und therapeutische Gebiete 6

3.1.3 Zielmoleküle und Signalwege 7

3.1.4 Forschungsprojekte 8

3.2 ABLAUF DES FORSCHUNGS- UND ENTWICKLUNGSPROZESSES UND AKTUELLE HERAUSFORDERUNGEN 9

3.2.1 Forschungs- und Entwicklungsprozess von Medikamenten 9

3.2.2 Aktuelle Herausforderungen für pharmazeutische Forschung und Entwicklung 11

3.3 INDIKATIONSÜBERGREIFENDE ZUSAMMENHÄNGE 13

3.3.1 Fallbeispiel: Indikationsübergreifende Funktion der Phosphodiesesterase 5 13

3.3.2 Fallbeispiel: Indikationsübergreifende Funktion von VEGF-A 17

3.4 ANSÄTZE ZUR UNTERSTÜTZUNG VON FORSCHUNG UND ENTWICKLUNG 20

3.4.1 Ansätze zur Unterstützung der Auswahl von geeigneten Zielmolekülen 20

3.4.2 Ansätze zur Integration systembiologischer Daten in Forschung und Entwicklung 21

4 ZIELSETZUNG DER ARBEIT 23

5 MATERIALIEN 25

6 METHODEN 27

6.1 THEORETISCHE GRUNDLAGEN DER DATENBANKMETHODIK 27

6.1.1 Grundlagen des Entity-Relationship-Modell 27

6.1.2 Grundlagen relationaler Datenbanken 28

6.1.2.1 Konzept und Struktur 28

6.1.2.2 Relationale Algebra 29

6.2 ROHDATEN 31

6.2.1 Pharmaprojects 31

6.2.1.1 Ursprung der Daten 31

6.2.1.2 Aufbereitung der Rohdaten 32

6.2.2 Kyoto Encyclopedia of Genes and Genomes (KEGG) 34

6.2.2.1 Ursprung der Daten 34

6.2.2.2 Aufbereitung der Rohdaten 34
6.3 GRUNDSTRUKTUR DER DATENBANK 36
6.3.1 Inputtabellen *36*
6.3.1.1 Therapeutische Gebiete, Indikationsgruppen und Indikationen („TAs and indications“) 36
6.3.1.2 Indikationen und pharmazeutisch aktive Substanzen („Indications and drugs“) 36
6.3.1.3 Pharmazeutisch aktive Substanzen und Zielmoleküle („Drugs and targets“) 38
6.3.1.4 Zielmoleküle und Signalwege („Targets and pathways”) 39
6.3.1.5 Pharmazeutisch aktive Substanzen und Firmen („Drugs and companies“) 40
6.3.2 Beziehungen zwischen den Inputtabellen *41*
6.3.3 Verwendete Abfragetypen und Makros *42*
6.4 ANALYSEN 44
6.4.1 Vernetzungsmatrix der therapeutischen Gebiete *46*
6.4.2 Vernetzungsgrad therapeutischer Gebiete *48*
6.4.3 Vernetzungsmatrix der Indikationsgruppen *49*
6.4.4 Vernetzungsanalyse ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen *50*
6.4.5 Gewichtung der Vernetzung von therapeutischen Gebieten *51*
6.4.5.1 Gewichtung ohne Durchschnittsbildung 52
6.4.5.2 Gewichtung mit einfacher Durchschnittsbildung 55
6.4.5.3 Gewichtung mit doppelter Durchschnittsbildung 57
6.4.5.4 Gewichtung mit doppelter Durchschnittsbildung und Multiplikator 57
6.4.5.5 Normierte Vernetzungsmatrix 60
6.4.6 Gewichtung der Vernetzung von Indikationsgruppen *60*
6.4.6.1 Gewichtung ohne Durchschnittsbildung 61
6.4.6.2 Gewichtung mit einfacher Durchschnittsbildung 64
6.4.6.3 Gewichtung mit doppelter Durchschnittsbildung 65
6.4.6.4 Gewichtung mit doppelter Durchschnittsbildung und Multiplikator 66
6.4.6.5 Normierte Vernetzungsmatrix 68
6.4.7 Gewichtung der Vernetzung ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen *68*
6.4.8 Signalwegranglisten *69*
6.4.8.1 Signalwegranglisten für Verbindungen zwischen therapeutischen Gebieten 70
6.4.8.2 Signalwegranglisten für Verbindungen zwischen Indikationsgruppen 71
6.4.8.3 Anteilige Verbindungsstärke 72
6.4.9 Identifikation von gemeinsamen Zielmolekülen *73*
6.4.10 Untersuchungen zur Charakterisierung der Datenbasis *76*
6.4.10.1 Eigenschaft der therapeutischen Gebiete 76
6.4.10.2 Eigenschaften der Signalwege 77
6.4.10.3 Eigenschaften der Zielmoleküle 79
6.4.10.4 Eigenschaften der Substanzen 80
6.5 STATISTISCHE METHODEN 81

6.6 BEWERTUNGSMATRIX UND BEWERTUNGSMETHODEN 83
6.6.1 Biologische Evidenz für indikationsübergreifende Zielmoleküle 84
6.6.2 Pharmazeutische Machbarkeit 87
6.6.3 Wettbewerbsintensität 88
6.6.4 Strategische Eignung 89

7 ERGEBNISSE 91

7.1 ENTWICKLUNG EINER METHODIK ZUR IDENTIFIZIERUNG VON INDIKATIONSÜBERGREIFENDEN SIGNALWEGEN 91
7.1.1 Grundstruktur der relationalen Datenbank 94
7.1.2 Allgemeine Statistik 96
7.1.2.1 Verteilung der Forschungsprojekte über therapeutische Gebiete 96
7.1.2.2 Verteilung der Forschungsprojekte über Entwicklungsphasen 97
7.1.2.3 Forschungsaktivität nach therapeutischen Gebieten 98
7.2 ANALYSE DER VERBINDUNGEN VON THERAPEUTISCHEN GEBIETEN DURCH GEMEINSAME SIGNALWEGE 100
7.2.1 Vernetzungsmatrix von therapeutischen Gebieten durch gemeinsame Signalwege 100
7.2.2 Vernetzungseigenschaften von Signalwegen und Zielmolekülen 105
7.2.2.1 Vernetzungsfähigkeit eines Signalwegs in Abhängigkeit seiner Größe 105
7.2.2.2 Vernetzungseigenschaften von Zielmolekülen 110
7.2.3 Gewichtung 111
7.2.3.1 Logik der Gewichtung 111
7.2.3.2 Variationen der Gewichtungslogik 113
7.2.3.3 Gewichtete Vernetzungsmatrizen 119
7.2.3.4 Variation des Gewichtungsfokus 123
7.3 ANALYSE DER VERBINDUNGEN VON INDIKATIONSGRUPPEN DURCH GEMEINSAME SIGNALWEGE 127
7.3.1 Verbindungspunkt Onkologie mit Bewegungsapparat 128
7.3.2 Verbindungspunkt Zentralnervensystem mit Herz-Kreislauf-System 131
7.3.3 Verbindungspunkt Zentralnervensystem mit Transplantationsmedizin und Immunsystem 133
7.3.4 Verbindungspunkt Bewegungsapparat mit Herz-Kreislauf-System 135
7.3.5 Verbindungspunkt Onkologie mit Herz-Kreislauf-System 137
7.4 IDENTIFIKATION UND PRIORISIERUNG VON GEMEINSAMEN SIGNALWEGEN 140
7.4.1 Methodik und Anwendung 140
7.4.2 Identifikation und Priorisierung von Signalwegen in der Verbindung von therapeutischen Gebieten: Beispiel Hämatologie und Magen-Darm-System 142
7.4.3 Identifikation und Priorisierung von Signalwegen in der Verbindung von Indikationsgruppen: Beispiel Lungenkrebs und rheumatoide Arthritis 144

7.5 IDENTIFIKATION VON MÖGLICHEN ZIELMOLEKÜLKANDIDATEN INNERHALB EINES SIGNALWEGS 147
7.5.1 Zielsetzung des vierten Analyseschritts 147
7.5.2 Apoptose als Verbindung zwischen Lungenkrebs und rheumatoider Arthritis 148
7.6 PRIORISIERUNG VON ZIELMOLEKÜLKANDIDATEN 151
7.6.1 Zielsetzung der Priorisierung 151
7.6.2 Priorisierungslogik 153
7.6.3 Exemplarische Priorisierung von fünf Molekülen für die Verbindung von Lungenkrebs und rheumatoider Arthritis durch den Apoptose-Signalweg 156
7.6.3.1 Detaillierte Bewertung der biologischen Evidenz für indikationsübergreifende Funktion 157
7.6.3.2 Detaillierte Bewertung der pharmazeutischen Machbarkeit 172
7.6.3.3 Detaillierte Bewertung der Wettbewerbsintensität 180
7.6.3.4 Gesamtbewertung 183

8 DISKUSSION 187

8.1 DISKUSSION DER ANALYSESCHRITTE AM BEISPIEL AUSGEWÄHLTER ERGEBNISSE 187
8.1.1 Kombinationen von therapeutischen Gebieten mit erhöhter Erfolgswahrscheinlichkeit für indikationsübergreifende Medikamentenentwicklung 189
8.1.2 Kombinationen von Indikationsgruppen mit erhöhter Erfolgswahrscheinlichkeit für indikationsübergreifende Medikamentenentwicklung 192
8.1.3 Identifizierung und Priorisierung von gemeinsamen Signalwegen 195
8.1.4 Identifizierung von möglichen Forschungskandidaten für Zielmoleküle mit indikationsübergreifender Funktion 197
8.1.5 Priorisierung von Kandidaten für zukünftig indikationsübergreifende Zielmoleküle 199
8.2 ANWENDUNG DER METHODIK ZUR IDENTIFIZIERUNG VON INDIKATIONSÜBERGREIFENDEN ZUSAMMENHÄNGEN FÜR PHARMAZEUTISCHE FORSCHUNG UND ENTWICKLUNG 201
8.2.1 Besondere Aspekte der Methodik 201
8.2.1.1 Berücksichtigung indirekter Verbindungen durch gemeinsame Signalwege 201
8.2.1.2 Lernen von Anderen durch Berücksichtigung bekannter Forschungs- und Entwicklungsprojekte 202
8.2.1.3 Trennung von biologischen und marktseitigen Zusammenhängen 203
8.2.1.4 Anwendung auf unterschiedlichen Detailebenen 204
8.2.1.5 Automatisierung zur Vereinfachung der Analyse großer Datenmengen 204
8.2.2 Herausforderungen und Grenzen der Methodik 205
8.2.2.1 Pflege der Datenbank 205
8.2.2.2 Restriktionen der Rohdaten 205
8.2.2.3 Grundeigenschaften von therapeutischen Gebieten, Indikationsgruppen und Signalwegen 206
8.2.2.4 Abbildung von Signalwegen 207

8.3 CHANCEN UND HERAUSFORDERUNGEN VON INDIKATIONSÜBERGREIFENDEN ZUSAMMENHÄNGEN FÜR DIE PHARMAZEUTISCHE INDUSTRIE 208
8.3.1 Indikationsübergreifende Zusammenhänge als Chance für Forschung und Entwicklung 208
8.3.2 Indikationsübergreifende Zusammenhänge und Nebenwirkungsrisiko 209
8.3.3 Spezifische Herausforderungen indikationsübergreifender Substanzen 210
8.3.4 Kommunikation zwischen Forschungsgruppen als zentraler Erfolgsfaktor 212
8.4 FAZIT UND AUSBLICK 213

9 LITERATURVERZEICHNIS 217

10 ANHANG 237

10.1 DETAILLIERTE GRAPHISCHE DARSTELLUNG DER MICROSOFT ACCESS ABFRAGEN 237
10.1.1 Abfragen zur Vernetzungsmatrix der therapeutischen Gebiete 237
10.1.2 Abfragen zur Vernetzungsmatrix der Indikationsgruppen 239
10.1.3 Abfragen zur Vernetzungsanalyse ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen 243
10.1.4 Abfragen zur Gewichtung der Vernetzung von therapeutischen Gebieten 244
10.1.4.1 Abfragen zur Gewichtung ohne Durchschnittsbildung 244
10.1.4.2 Abfragen zur Gewichtung mit einfacher Durchschnittsbildung 250
10.1.4.3 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung 252
10.1.4.4 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung und Multiplikator 253
10.1.5 Abfragen zur Gewichtung der Vernetzung von Indikationsgruppen 258
10.1.5.1 Abfragen zur Gewichtung ohne Durchschnittsbildung 258
10.1.5.2 Abfragen zur Gewichtung mit einfacher Durchschnittsbildung 264
10.1.5.3 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung 266
10.1.5.4 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung und Mulitplikator 267
10.1.6 Abfragen zu Signalwegranglisten 273
10.1.6.1 Abfrage für die Signalwegranglisten für Verbindungen zwischen therapeutischen Gebieten 273
10.1.6.2 Abfrage für die Signalwegranglisten für Verbindungen zwischen Indikationsgruppen 275
10.1.7 Abfragen zur Identifikation von gemeinsamen Zielmolekülen 277
10.1.8 Abfragen zur Charakterisierung der Datenbasis 285
10.1.8.1 Eigenschaften der therapeutischen Gebiete 285
10.1.8.2 Eigenschaften der Signalwege 286
10.1.8.3 Eigenschaften der Zielmoleküle 288
10.1.8.4 Eigenschaften der Substanzen 290
10.2 IDENTIFIKATION UND PRIORISIERUNG VON SIGNALWEGEN IN DER VERBINDUNG VON INDIKATIONSGRUPPEN (WEITERE ANALYSEN) 292
10.2.1.1 Signalwegpriorisierung zwischen Brustkrebs und rheumatoider Arthritis 292
10.2.1.2 Signalwegpriorisierung zwischen Alzheimer und Demenz und Thrombose 294
10.2.1.3 Signalwegpriorisierung zwischen Schizophrenie und Bluthochdruck 296

10.2.1.4 Signalwegpriorisierung zwischen Alzheimer und Demenz und generalisierten Immunerkrankungen 297
10.2.1.5 Signalwegpriorisierung zwischen multipler Sklerose und generalisierten Immunerkrankungen 299
10.2.1.6 Signalwegpriorisierung zwischen Thrombose und rheumatoider Arthritis 300
10.2.1.7 Signalwegpriorisierung zwischen Lungenkrebs und Thrombose 303
10.3 ABKÜRZUNGSVERZEICHNIS 305
10.4 VERZEICHNIS DER AKADEMISCHEN LEHRER 307
10.5 DANKSAGUNG 308

1 Zusammenfassung

Medikamente, die indikationsübergreifend zur Therapie von mehreren Krankheiten genutzt werden können, sind meist nicht nur von medizinischem Nutzen für die Patienten, sondern auch von wirtschaftlichem Nutzen für das entwickelnde Unternehmen. Obwohl immer mehr Wissen über biologische Signalwege und funktionelle Zusammenhänge von Zielmolekülen verfügbar ist, ist die Entwicklung dieser indikationsübergreifenden Medikamente meist von Zufällen abhängig. Ziel der vorliegenden Arbeit war daher die Entwicklung eines systematischen Verfahrens zur Identifikation von indikationsübergreifenden Signalwegen und Zielmolekülen.

Zu diesem Zweck wurde eine Access-basierte relationale Datenbank aufgebaut, die systematisch und automatisiert Verbindungen von therapeutischen Gebieten und Indikationsgruppen durch indikationsübergreifende Signalwege und Zielmoleküle untersucht. Eine angeschlossene Priorisierung ermöglicht, konkrete Zielmolekülkandidaten zu ermitteln, die eine erhöhte Chance für indikationsübergreifende Bedeutung haben. Die Datenbank berücksichtigt dabei erstmals neben direkten Verbindungen zweier therapeutischer Gebiete durch ein gemeinsames Zielmolekül auch indirekte Verbindungen über Signalwege, die kein gemeinsames Zielmolekül haben. Mit Hilfe verschiedener Gewichtungsformen können unterschiedlich starke Verbindungen differenziert sowie der Schwerpunkt der Analyse variiert werden. Im Einzelnen können die folgenden Aspekte untersucht werden:

1. **Verbindung von therapeutischen Gebieten durch gemeinsame Signalwege**

   Der Grad der Verbindung eines therapeutischen Gebiets mit einem anderen kann durch Berücksichtigung von gemeinsamen Signalwegen und Zielmolekülen bestimmt werden. Die Verbindungen können entsprechend ihrer biologischen Stärke und ihrer kommerziellen Bedeutung gewichtet werden.

2. **Verbindung von Indikationsgruppen durch gemeinsame Signalwege**

   Innerhalb von bereits bestehenden Forschungsschwerpunkten können Kombinationen von Indikationsgruppen identifiziert werden, die besonders eng miteinander verbunden sind und damit eine hohe Chance für indikationsübergreifende Projekte bieten.

3. **Identifikation und Priorisierung von gemeinsamen Signalwegen**

   Die Datenbank erlaubt es, die wichtigsten Signalwege einer Verbindung von zwei Indikationsgruppen zu identifizieren. Dies ermöglicht die Spezialisierung von Methoden und Versuchsansätzen auf diejenigen Signalwege, die mit der größten Wahrscheinlichkeit für beide Indikationen von Bedeutung sind.

4. **Identifikation von möglichen Zielmolekülkandidaten**

   Ist bereits eine Ausrichtung auf Indikationsgruppen und Signalwege erfolgt, können alle bereits bekannten indikationsübergreifenden Zielmoleküle sowie Kandidaten für zukünftig indikationsübergreifende Zielmoleküle ermittelt werden. Kandidaten für zukünftig indikationsübergreifende Zielmoleküle werden bisher nur für eine der beiden Indikationsgruppen erforscht bzw. genutzt. Aufgrund ihrer Position im Signalweg haben sie aber eine gute Chance, in Zukunft auch für die andere Indikationsgruppe von Bedeutung zu sein.

5. **Priorisierung von Zielmolekülkandidaten**

   Eine Bewertung der Zielmolekülkandidaten in den vier Bereichen „biologische Evidenz für indikationsübergreifende Funktion“, „pharmazeutische Machbarkeit“, „Wettbewerbsintensität“ und „strategische Eignung“ reduziert die Empfehlung von Zielmolekülkandidaten auf die am besten geeigneten.

Das in dieser Arbeit entwickelte Instrument bietet damit eine neue Möglichkeit, systematisch die Forschung und Entwicklung an indikationsübergreifenden Zielmolekülen und Signalwegen zu unterstützen. Zur erfolgreichen Entwicklung von indikationsübergreifenden Medikamenten ist darüber hinaus vor allem intensive Kommunikation und Zusammenarbeit zwischen den einzelnen Forschungs- und Entwicklungsgruppen wichtig.

2 Summary

Drugs that can be used to treat more than one disease are not only a great benefit for patients but have also proven to be an economical advantage for developing pharmaceutical companies. Although a great deal of information is available about signaling pathways and functional relationships between drug targets, so far the development of these cross-indication drugs has been mostly a product of chance.

The goal of this work was to establish a systematic methodology for identifying drug targets and signaling pathways that are common in more than one disease (common targets/common pathways).

An Access-based database was built that automatically and systematically analyzes connections of therapeutic areas and indication groups via common targets and pathways. An adjacent prioritization allows naming drug targets that have an increased chance to be important in more than one indication. The database allows identifying both direct connections between therapeutic areas via common targets as well as indirect connections via common signaling pathways that do not share common targets.

In brief, it is possible to analyze the following:

1. **Connections of therapeutic areas via common pathways**

   Shared biological pathways interconnect different therapeutic areas. The degree of their connectivity can be visualized in a matrix. Depending on the focus of the analysis it is possible to weigh the connections according to their biological or commercial strength.

2. **Connections between indication groups via common pathways**

   Most pharmaceutical companies already have established research focus areas. The database therefore enables the identification of the closest connections of indication groups within two predefined therapeutic areas. These indication groups have the highest chance to be targetable by cross-indication projects.

3. Identification and prioritization of common pathways

Using the database it is possible to identify the pathways that are most important for the connection of two indication groups. This knowledge supports specialization of research tools and capabilities along those pathways that are most likely to encompass common targets and common physiological functions.

4. Identification of common target candidates

For a strategic focus on indication groups and pathways, it is possible to create a list of both known common targets as well as novel common target candidates. Known common targets are already in research and development for both of the selected indication groups. Common target candidates are part of the common pathway, however have only been pharmaceutically addressed for one of the indication groups. Due to their role in the common pathway, these targets have an increased likelihood to be of use for both indication groups in the future.

5. Prioritization of common target candidates

Common target candidates can be evaluated in four categories: “biological evidence of a cross-indication function”, “feasibility of drug development”, “competitive environment” and “strategic fit”. This allows the separation of more promising candidates from less optimal ones.

Pharmaceutical companies and academic research groups could use this novel methodology to systematically support research and development of common targets and pathways.

To ultimately succeed with the development of cross-indication drugs, intensive communication and cooperation across therapeutic areas are most critical.

3 Einführung

Diese Doktorarbeit befasst sich mit der Entwicklung von Medikamenten durch pharmazeutische Forschung und Entwicklung. Dabei werden vor allem biologische Zusammenhänge, die in mehr als einer Krankheit eine Rolle spielen, untersucht.

Im folgenden Einführungskapitel sollen für die Arbeit wichtige Zusammenhänge, die Motivation der Beschäftigung mit dem Thema sowie der aktuelle Forschungsstand erläutert werden.

3.1 Grundgrößen von pharmazeutischer Forschung und Entwicklung

Die Entwicklung eines Medikaments ist ein komplexer Prozess, in dem viele verschiedene Größen eine Rolle spielen. Im Folgenden sollen diejenigen definiert werden, die für die vorliegende Arbeit von außerordentlicher Bedeutung sind.

3.1.1 Substanzen und Medikamente

Der Begriff des Medikaments kommt vom lateinischen Wort *medicamentum*, Heilmittel. Ein Medikament ist ein Arzneimittel, das in bestimmter Dosierung zur Heilung, Vorbeugung oder Linderung einer Krankheit dient. Medikamente werden von pharmazeutischen Unternehmen entwickelt und auf den Markt gebracht. Im Sprachgebrauch wird ein Medikament daher in der Regel mit dem Namen bezeichnet, der ihm vom entwickelnden bzw. vertreibenden Unternehmen gegeben wird (z.B. Aspirin® von Bayer HealthCare). Der Begriff Substanz beschreibt den Wirkstoff eines Medikaments (z.B. Acetylsalicylsäure für Aspirin®). Der Substanzname kann dabei an die chemische Struktur der Substanz angelegt oder ein Kunstwort sein. Eine Substanz, die unter Patentschutz steht, wird mit Ausnahme von bestimmten Lizenzvereinbarungen, nur unter einem Medikamentennamen vertrieben. Substanzen, die keinen Patentschutz mehr

haben, können in verschiedenen Medikamenten vorkommen (z.B. Acetylsalicylsäure in Acesal® von Nycomed).

Generell können Substanzen zu verschiedenen Molekülklassen gehören, wobei in der Regel zwischen kleinen, chemisch synthetisierbaren Substanzen („small-molecule drugs“ oder „small molecules“) und biologischen Substanzen („biological drugs“) unterschieden wird (Overington u. a., 2006). Zur Gruppe der biologischen Substanzen gehören neben den Antikörpern auch Nukleinsäuren und therapeutische Viren.

3.1.2 Indikationen, Indikationsgruppen und therapeutische Gebiete

Der Begriff der Indikation steht im medizinischen Kontext für den Einsatz einer bestimmten medizinischen Maßnahme bei einem bestimmten Krankheitsbild und leitet sich von lateinischen Wort *indicare*, anzeigen, ab. Im Zusammenhang mit Medikamenten versteht man unter der Indikation eines Medikaments das Krankheitsbild, für dessen Behandlung das Medikament eine Zulassung hat. Bereits vor Eintritt in die klinische Entwicklung muss die Indikation einer neuen Substanz festgelegt werden, da dies für die Auswahl der Patienten, die in die klinischen Studien ab Phase II eingeschlossen werden, von Bedeutung ist und von den regulatorischen Gremien verlangt wird (FDA, 2004). Jeder Substanz in der Entwicklung ist daher eine eindeutige Indikation zugeordnet (z.B. Entwicklung einer Substanz für die Behandlung von nicht-kleinzelligem Lungenkarzinom).

Für die Analysen der vorliegenden Arbeit wurden verwandte Indikationen zu einer Indikationsgruppe zusammengefasst. Ein Beispiel für eine solche Zusammenfassung ist die Indikationsgruppe Lungenkrebs, die die Indikationen kleinzelliges Lungenkarzinom, nicht-kleinzelliges Lungenkarzinom und nicht weiter spezifizierte Formen von Lungenkrebs („lung cancer, general“) enthält. Die Einführung der Größe der Indikationsgruppen ist deshalb sinnvoll, weil sich die Pathophysiologie der zusammengefassten Krankheiten häufig ähnelt. Eine Untersuchung von biologischen Zusammenhängen zwischen einzelnen Indikationen ist aufgrund der sehr kleinen Einheit und der großen Anzahl von unterschiedlichen Indikationen nur schwer möglich. Darüber hinaus sind Zusammenhänge zwischen verwandten Indikationen (z.B. verschiedenen Formen

von Lungenkrebs) *a priori* naheliegend. Durch Zusammenfassung der verwandten Indikationen in Indikationsgruppen wird eine Untersuchung zwischen unterschiedlichen Bereichen vereinfacht.

In der Praxis der pharmazeutischen Industrie werden verwandte Indikationen in sogenannten therapeutischen Gebieten zusammengefasst. Therapeutische Gebiete können als Übergruppe der Indikationsgruppen aufgefasst werden. Die Indikationsgruppe Lungenkrebs beispielsweise gehört zum therapeutischen Gebiet Onkologie. Die Definition der therapeutischen Gebiete ist damit sowohl durch die biologischen Zusammenhänge als auch durch die medizinische Praxis bestimmt. Während das therapeutische Gebiet Onkologie vor allem in den Arbeitsbereich von Internisten bzw. Onkologen fällt, kann beispielsweise das Gebiet Urogenitalsystem den Gynäkologen und Urologen zugeordnet werden. Diese praktische Ausrichtung an Ärztegruppen hat ihren Ursprung in der historisch gewachsenen Organisationsstruktur von pharmazeutischen Unternehmen. In den meisten Firmen sind den Vertriebsorganisationen bestimmten Forschungs- und Entwicklungsgruppen zugeordnet, die gemeinsam innerhalb eines therapeutischen Gebiets arbeiten.

3.1.3 Zielmoleküle und Signalwege

Eine pharmazeutisch aktive Substanz erreicht ihre Wirkung, in dem sie ein bestimmtes Zielmolekül („drug target“) angreift. Grundsätzlich können einzelne Proteine, Nukleinsäuren und andere spezifische Zellstrukturen (z.B. Ribosomen, Mikrotubuli) als Zielmoleküle dienen. Nicht bei allen pharmazeutisch aktiven Substanzen sind der genaue Wirkmechanismus und das Zielmolekül bekannt. Der Großteil der Substanzen greift jedoch an Proteinen an (Overington u. a., 2006). In Bezug auf die Anzahl der bekannten und bewiesenen Zielmoleküle gibt es verschiedene Meinungen. Erste Untersuchungen Ende der 90er Jahre konnten damals knapp 500 Zielmoleküle identifizieren (Drews, 1996; Drews und Ryser, 1997). Andere Veröffentlichungen konnten 2003 nur 273 Proteine zählen, die durch Medikamente angegriffen werden. Ein Versuch, einen Überblick über die Anzahl der verfügbaren Zielmoleküle zu bekommen, ist die „Therapeutic Targets Database“ (Chen u. a., 2002), die noch im Jahr 1996 ca. 500 Zielmoleküle zählte

und 2006 bereits 1535 Zielmoleküle enthielt (Zheng u. a., 2006). Zusammenfassend lässt sich sagen, dass es nur eine verhältnismäßig kleine Zahl von anerkannten Protein-Zielmolekülen gibt, diese jedoch stetig wächst (Overington u. a., 2006; Zheng u. a., 2006). Zielmoleküle können sowohl zum menschlichen Organismus gehören, als auch – zum Beispiel bei einer Infektion – zum Organismus des Erregers, wobei mit über 80 % der Großteil der Konsensus-Zielmoleküle menschlichen Ursprungs ist (Overington u. a., 2006).

Moleküle sind innerhalb der Zelle in funktionellen Gruppen organisiert, die in ihrer Gesamtheit für die physiologischen Funktionen verantwortlich sind. Diese funktionellen Gruppen werden als Signalwege („molecular signalling pathways“) bezeichnet und ihre Funktion wird in der Regel durch bestimmte extrazelluläre Botenstoffe getriggert, die intrazelluläre Reaktionskaskaden anstoßen (Fishman und Porter, 2005). Es existiert keine standardisierte Definition von Signalwegen, so dass deren Anzahl je nach Ansatz (z.B. ein Signalweg pro Zelloberflächenrezeptor oder ein Signalweg pro Zelltyp und Funktion) variieren kann. In jedem Fall ist die Zahl der Signalwege aber kleiner als die der Zielmoleküle (Fishman und Porter, 2005). Eine mögliche Quelle für umfangreiche Informationen über Signalwege bietet die Datenbank der Kyoto Encyclopedia of Genes and Genomes (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa Laboratories, 2008; Kanehisa u. a., 2008), die auch in der vorliegenden Arbeit verwendet wurde.

3.1.4 Forschungsprojekte

Der Begriff des Forschungsprojekts wird im Rahmen dieser Arbeit für jede präklinische oder klinische Entwicklung (vergleiche Abschnitt 3.2) einer Substanz für eine bestimmte Indikation verwendet. Ein Forschungsprojekt beschreibt daher eine eindeutige Kombination von Substanz und Indikation. Je nachdem, wie weit ein Forschungsprojekt bereits durch den Prozess der pharmazeutischen Entwicklung fortgeschritten ist, hat es einen bestimmten Entwicklungsstatus (zum Beispiel „Phase I“). Zur Vereinfachung werden auch diejenigen Projekte, die bereits mit der Zulassung des Medikaments in der Indikation abgeschlossen sind, als Forschungsprojekte bezeichnet.

3.2 Ablauf des Forschungs- und Entwicklungsprozesses und aktuelle Herausforderungen

3.2.1 Forschungs- und Entwicklungsprozess von Medikamenten

Die Entwicklung eines neuen Medikaments ist ein vergleichsweise langwieriger und teurer Prozess. Es wird davon ausgegangen, dass von der Synthese einer ersten Substanz bis zur Zulassung des neuen Medikaments durchschnittlich neun bis zwölf Jahre vergehen (Dickson und Gagnon, 2004). Zusätzlich dauert es im besten Fall ein Jahr, meist jedoch mehrere Jahre, das Zielmolekül auszuwählen und zu validieren, für das die Substanz entwickelt werden soll (Fishman und Porter, 2005). Dickson und Gagnon schätzten die Kosten für die Entwicklung einer neuen Medikaments im Jahr 2004 auf ca. 800 Mio. US $ (Dickson und Gagnon, 2004).

Grundsätzlich kann der Entwicklungsprozess in vier Phasen geteilt werden (Dickson und Gagnon, 2004). Die erste Phase entspricht der eigentlichen Forschung und wird auch als Präklinik bezeichnet. In der Präklinik wird im ersten Schritt ein Zielmolekül ausgewählt, das – zum Beispiel aufgrund einer theoretischen Forschungshypothese – einen möglichen Ansatzpunkt für eine medikamentöse Therapie bietet. Im nächsten Schritt wird dieses Zielmolekül experimentell validiert. Kann eine Funktion festgestellt werden, wird ein *in vitro* Versuchsaufbau entwickelt, mit dem es möglich ist, Substanzen zu suchen, die dieses Zielmolekül angreifen und seine Funktion beeinflussen. Mit technischer Unterstützung werden dann in dem *in vitro* Versuchsaufbau viele Substanzen getestet, um eine geeignete Substanz zu identifizieren, die das Zielmolekül angreift und seine Funktion in gewünschter Weise beeinflusst („high-throughput screening“). Ist eine mögliche Leitsubstanz („lead“) identifiziert, muss die Wirkung der Substanz erneut im Experiment bewiesen werden. Dieser Schritt wird auch als „proof-of-principle“ bezeichnet. Je nach Beschaffenheit und Funktion der Leitsubstanz wird ein Optimierungsschritt eingefügt, bei dem zum Beispiel die chemischen Charakteristika einer Substanz verändert werden (Sams-Dodd, 2006). Wenn eine Leitsubstanz feststeht, müssen deren toxikologische und

pharmakologische Eigenschaften in mindestens zwei Spezies von Labortieren gezeigt werden. Kann die Sicherheit der Leitsubstanz bewiesen werden, wird vom Hersteller ein Antrag an die regulatorische Behörde zum Eintritt des neuen möglichen Medikaments („investigational new drug“, IND) in die klinische Entwicklung gestellt (Dickson und Gagnon, 2004). Regulatorische Behörden, die vor für die Zulassung von neuen Medikamenten zuständig sind, sind unter anderem die europäische Arzneimittelagentur (European Medicines Agency, EMEA) und die amerikanische Food and Drug Administration (FDA). Hat die zuständige Behörde keine Einwände, kann mit der klinischen Entwicklung begonnen werden. In der Regel werden parallel zu den klinischen Studien noch Langzeit-Tierversuche durchgeführt. Sollten sich hier toxische Effekte zeigen, werden die klinischen Studien unmittelbar abgebrochen (Dickson und Gagnon, 2004). Die Dauer der präklinischen Phase wird vor allem durch den Zeitraum der Zielmolekülauswahl bestimmt. Dauert diese im besten Fall ein Jahr, kann nach ca. sechs Jahren mit einem Eintritt in die klinische Entwicklung gerechnet werden (Fishman und Porter, 2005).

Die zweite Phase der Medikamentenentwicklung beinhaltet die klinische Entwicklung des Medikaments für den Menschen und dauert durchschnittlich fünf Jahre (Dickson und Gagnon, 2004). Die klinische Entwicklung ist in drei Phasen aufgeteilt. In der Phase I wird an einer kleinen Anzahl von gesunden Probanden überprüft, ob die Leitsubstanz beim Menschen toxische Effekte hat und welche Dosierungen verträglich sind. Ist das Medikament verträglich, kann in Phase II die Wirksamkeit überprüft werden. Dazu wird eine größere Gruppe von Probanden benötigt, die an der Krankheit leiden, die durch das neue Medikament behandelt werden soll. Bestätigt sich die Wirksamkeit und treten weiterhin keine bedenklichen Nebenwirkungen auf, geht die Leitsubstanz in die Phase III. Diese Phase schließt eine noch größere Gruppe von Probanden ein, die unter der Erkrankung leiden, und kann unterschiedliche Dosierungen der Substanz testen. Die klinische Phase III Studie ist besonders wichtig, um die Wirksamkeit in einer größeren Gruppe zu zeigen. Häufig werden auch Nebenwirkungen erst in der größeren Probandengruppe auffällig. Kann auch diese klinische Studie erfolgreich abgeschlossen werden, gilt die Sicherheit und die Wirksamkeit der Substanz als

bestätigt und das Unternehmen kann einen Antrag zur Zulassung eines neuen Medikaments stellen („new drug applicaton“, NDA) (Dickson und Gagnon, 2004).

Bereits während der klinischen Phase gibt es zahlreiche Interaktionspunkte der regulatorischen Behörde mit dem pharmazeutischen Unternehmen (FDA, 2004). Die eigentliche Prüfung des Zulassungsantrags nach Phase III kann von wenigen Monaten bis mehreren Jahren dauern. In diesem Zeitraum bereitet das pharmazeutische Unternehmen in der Regel die Markteinführung vor. Erfolgt die Zulassung, wird das Medikament auf den Markt gebracht. Insbesondere in den ersten Jahren schließt sich eine intensive Beobachtung möglicher Nebenwirkungen an, die zum Teil erst jetzt aufgrund der höheren Patientenanzahl auffallen („post-marketing-surveillance“) (Dickson und Gagnon, 2004; FDA, 2004).

3.2.2 Aktuelle Herausforderungen für pharmazeutische Forschung und Entwicklung

Der Erfolg pharmazeutischer Forschung und Entwicklung ist von verschiedenen Größen abhängig. Neben der wissenschaftlichen Unsicherheit des Medikamentenerfolgs beeinflussen auch wirtschaftliche und regulatorische Faktoren die Medikamentenentwicklung (Dickson und Gagnon, 2004).

Geringe Produktivität trotz steigender Kosten

Um neue Medikamente zu entwickeln, investiert die pharmazeutische Industrie jährlich mehrstellige Milliardenbeträge (Araujo u. a., 2007). Alleine von 1993 bis 2004 haben sich die Kosten für Forschung und Entwicklung dabei um 147 % gesteigert – die Anzahl der Anträge für Medikamentenzulassungen bei der amerikanischen Food and Drug Administration (FDA) stiegen in selben Zeitraum aber nur um 38 % (Lynch u. a., 2004; Araujo u. a., 2007). Andere Veröffentlichungen sprechen sogar von einem 30jährigen Produktivitätsrückgang in pharmazeutischer Forschung und Entwicklung (Couzin, 2002; Butcher, 2005). Diese Situation beunruhigt nicht nur in die pharmazeutische Industrie, sondern auch regulatorische Behörden wie die FDA (FDA, 2004) und zahlreiche

wissenschaftliche Autoren, die von einer Enttäuschung durch die Medikamentenentwicklung sprechen (Butcher, 2005; Sams-Dodd, 2006; Araujo u. a., 2007; Hellerstein, 2008a; Schrattenholz und Soskic, 2008).

Schlechte Erfolgsaussichten für neue Zielmoleküle

Die Anzahl von bekannten, pharmazeutisch angreifbaren Zielmolekülen steigt zwar insgesamt an (Overington u. a., 2006), jedoch geschieht dies ausgesprochen langsam. Der Trend für die Zulassung von Medikamenten gegen neue Zielmoleküle, dies sind Moleküle, die bisher noch nicht durch ein Medikament angegriffen werden, ist entmutigend. Zambrowicz und Sands zeigten 2003, dass pro Jahr im Durchschnitt nur ca. vier neue Zielmoleküle durch neu zugelassene Medikamente adressiert werden (Zambrowicz und Sands, 2003). Andere Publikationen sprechen von einer Zulassung von ein bis drei Medikamenten pro Jahr für die gesamte weltweite Industrie gegen neue Zielmoleküle (Butcher, 2005; Araujo u. a., 2007).

Herausforderung zur Integration großer Datenmengen

Während die Entwicklung von Medikamenten früher primär durch Beobachtungen bestimmt war (Drews, 2000; Hellerstein, 2008a), sind in den vergangenen Jahrzehnten große Fortschritte in der Verfügbarkeit von biologischen Daten gemacht worden. Die vollständige Identifizierung des Genoms sowie detaillierte Informationen über das Proteom und wichtige Signalwege im Organismus führen zu großen Datenmengen, deren Integration in Forschung und Entwicklung eine Herausforderung ist. Es sind unter anderem neue computergesteuerte und mathematische Konzepte nötig sind (Nicholson und Wilson, 2003; Nikolsky u. a., 2005; Schrattenholz und Soskic, 2008). Dennoch wird es letztlich aufgrund der Komplexität und des nichtlinearen Verhaltens biologischer Systeme unmöglich sein, diese vollständig zu modellieren (Ekins u. a., 2007).

3.3 Indikationsübergreifende Zusammenhänge

Bei steigenden Kosten und geringen Chancen, erfolgreich Medikamente gegen neue Zielmoleküle zu entwickeln, liegt der Ansatz nahe, bereits bekannte Zielmoleküle maximal auszunutzen. In diesem Kontext müssen indikationsübergreifende Zusammenhänge betrachtet werden. Unter einem indikationsübergreifenden Zielmolekül versteht man ein Zielmolekül, das für mehr als eine Krankheit therapeutisch genutzt werden kann. Ein indikationsübergreifender Signalweg ist ein Signalweg, der in mehreren Krankheitsbildern eine Rolle spielt und therapeutisch adressiert werden kann. Das indikationsübergreifende Zielmolekül ist damit der direkte Angriffspunkt einer möglichen Substanz, während es sich bei dem indikationsübergreifenden Signalweg um den funktionellen Zusammenhang mehrerer Zielmoleküle handelt. Die Nutzung eines indikationsübergreifenden Zielmoleküls ist für die Industrie daher unmittelbar durch Testen einer existierenden Leitsubstanz oder Entwicklung einer neuen Substanz möglich. Im Gegensatz dazu ist ein indikationsübergreifender Signalweg eher als Hinweis auf eine mögliche Gruppe von Zielmolekülen, gegen die Substanzen entwickelt werden können, zu sehen.

Historisch gesehen, wurde die Entdeckung von indikationsübergreifender Funktion einzelner Zielmoleküle häufig durch Zufallsbeobachtungen ausgelöst (Kling, 1998; Hellerstein, 2008b). In anderen Fällen wurden indikationsübergreifende Signalwege – wie Angiogenese – als Ansatzpunkte für die Suche nach indikationsübergreifender Zielmoleküle genutzt (Ferrara u. a., 2006; Folkman, 2007). Unabhängig von der Art der Entdeckung der indikationsübergreifenden Zielmoleküle waren diese sowohl von medizinischem Nutzen für die Patienten als auch von wirtschaftlichem Nutzen für die entwickelnde Firma. Zwei Fallbeispiele erläutern dies im Folgenden.

3.3.1 Fallbeispiel: Indikationsübergreifende Funktion der Phosphodiesesterase 5

Der Phosphodiesterase 5 (PDE 5) kommt eine wichtige Funktion im Zusammenhang mit der Signaltransduktion durch Stickstoffmonoxid (NO) zu.

Das Enzym beendet die durch NO über cyclisches Guanosinmonophosphat (cGMP) vermittelte Entspannung der glatten Muskulatur durch Abbau von cGMP. Eine Inhibition der Phosphodiesterase führt daher durch Akkumulation von cGMP zu einer Verstärkung des Relaxationseffekts (Omori und Kotera, 2007).

Nachdem 1977 die erste Hypothese über eine mögliche pharmakologische Nutzung einer selektiven PDE-Inhibition veröffentlicht wurde (Weiss und Hait, 1977), wurde im Jahr 1986 ein Projektteam bei der Firma Pfizer gebildet, das einen Inhibitor der PDE zur Behandlung von Bluthochdruck entwickeln sollte (Kling, 1998; Ghofrani u. a., 2006).

Im Jahr 1988 musste die Indikation Bluthochdruck jedoch aufgegeben werden, da sich die initiale Forschungshypothese zur Wirkung eines PDE-Inhibitors in der Niere zur Blutdrucksenkung nicht bestätigen ließ (Kling, 1998). Die starke Expression des Enzym-Isotyps PDE 5 in der Gefäßmuskulatur sowie die Möglichkeit mit neuen Inhibitionsuntersuchungen gezielt nach selektiven PDE 5-Inhibitoren zu suchen, führte im Jahr 1989 zur Entwicklung der Leitsubstanz UK-92,480, die später Sildenafil genannt wurde. Indikation für die neue Substanz war nun Angina Pectoris (Kling, 1998; Ghofrani u. a., 2006). Mehrere Phase I Studien zeigten moderate Nebenwirkungen von Sildenafil – unter anderem das Auftreten von penilen Erektionen .(Kling, 1998; Walker u. a., 1999; Ghofrani u. a., 2006). Nach Phase II Studien, die keine befriedigende Wirkung von Sildenafil bei Angina pectoris zeigen konnten, und den problematischen Wechselwirkungen mit Nitraten zur Behandlung von Angina pectoris, musste 1993 auch diese Indikation verworfen werden (Webb u. a., 1999; Ghofrani u. a., 2006).

Die Rettung von Sildenafil wurde Anfang der 90er Jahre durch neue Forschungsergebnisse eingeleitet, die einen Zusammenhang von NO und cGMP mit Störungen der penilen Erektion (erektile Dysfunktion) zeigten (Ignarro u. a., 1990; Azadzoi u. a., 1992; Bush u. a., 1992). Während Störungen der penilen Erektion bis zu den 80er Jahren primär als psychosomatische Erkrankung gesehen worden waren, wurde nun diskutiert, dass eine Störung in der NO-cGMP-vermittelten Vasodilatation des arteriellen Gefäßgeflechts im Penis (Corpus Cavernosum) Ursache der erektilen Dysfunktion sein könnte. Tatsächlich ist heute

bekannt, dass eine normale penile Erektion durch sexuelle Stimuation über nicht-adrenerge, nicht-cholinerge Nervenfasern des Parasympathikus vermittelt wird, die NO als Neurotransmitter im Corpus Cavernosum verwenden. Zusätzliche Stimulation über cholinerge Fasern regt die Bildung von weiterem Stickstoffmonoxid durch die endotheliale NO-Synthase an. NO führt über cGMP zu einer Relaxation der Gefäße des Corpus Cavernosum und damit sowohl zu erhöhtem Bluteinstrom als auch zur Kompression der ableitenden Venen (McVary, 2007). Diese Pathophysiologie lieferte eine Erklärung für das Auftreten der Nebenwirkung Erektion unter Sildenafil Behandlung und bot eine ideale Wirkgrundlage für Sildenafil zur Therapie von erektiler Dysfunktion. Eine Erektion wird unter Sildenafil nur dann ausgelöst, wenn der initiale NO-liefernde Stimulus durch sexuelle Stimulation vorhanden ist. Erste klinische Studien für Sildenafil bei erektiler Dysfunktion wurden von Pfizer 1993 initiiert (Boolell u. a., 1996a; Boolell u. a., 1996b) und ein großes Phase II/Phase III-Programm schloss sich 1994 und 1995 an. In allen Studien konnten eine dosisabhängige Verstärkung der Erektion in 70 % aller Anwendungen bei guter Verträglichkeit gezeigt werden (Goldstein u. a., 1998; Dinsmore u. a., 1999; Montorsi u. a., 1999). Das klinische Programm war sogar so erfolgreich, dass es nach Abschluss der Studie 1997 zu Widerstand bei der Rückgabe der unverbrauchten Tabletten kam (Kling, 1998). 1998 wurde Sildenafil unter dem Handelsnamen Viagra® zur Behandlung von erektiler Dysfunktion von der FDA und der EMEA zugelassen. Innerhalb eines Jahres erreichte Viagra® mehr als eine Milliarde US $ Umsatz (Pfizer Inc., 2003) und wuchs bis zur Markteinführung der Wettbewerbsmedikamente Cialis® (Eli Lilly) und Levitra® (Bayer HealthCare/Glaxo Smith Kline) im Jahr 2003 jährlich um ca. 20 % (Pfizer Inc., 2003; Pfizer Inc., 2004).

Nach erfolgreichem Markteintritt der Substanz Sildenafil für die Behandlung erektiler Dysfunktion 1998, konnte Pfizer in den folgenden Jahren von neuen Forschungserkenntnissen profitieren und dieselbe Substanz im Jahr 2005 als Revatio® zur Behandlung von pulmonaler Hypertension (Bluthochdruck im Lungenkreislauf) auf dem Markt bringen. Nachdem im Jahr 1991 erste Hinweise für eine Rolle von cGMP bei pulmonaler Hypertension veröffentlicht worden waren (Haynes u. a., 1991), konnten in den folgenden Jahren verschiedene

experimentelle Modelle eine vasodilatatorische und damit blutdrucksenkende Wirkung von PDE 5-Inhibitoren in den Lungengefäßen zeigen (Braner u. a., 1993; Ichinose u. a., 1995; Thusu u. a., 1995; Cohen u. a., 1996). Heute ist bekannt, dass PDE 5 neben der hohen Konzentration im Corpus Cavernosum auch in der Lunge stark exprimiert ist (Bohle u. a., 2000; German u. a., 2000; Corbin u. a., 2005). In der Lunge findet eine lokale Anpassung der Durchblutung an die Ventilation durch einen NO-vermittelten Mechanismus statt. Bei Dehnung der Alveolen durch Ventilation kommt es zur NO-Bildung und über cGMP zur Vasodilatation (Grimminger u. a., 1995; Ide u. a., 1999; Ghofrani u. a., 2004a). Dieser physiologische Mechanismus erlaubt die optimale Aufnahme von Sauerstoff ins Blut. Erste klinische Studien mit Sildenafil bei pulmonaler Hypertension konnten 1998 bis 2000 eine dosisabhängig Absenkung des pulmonalen Blutdrucks feststellen (Ghofrani u. a., 2006). Nachdem Effektivität und Toleranz von chronischer Behandlung mit Sildenafil gezeigt werden konnten (Kothari und Duggal, 2002; Sastry u. a., 2002; Sastry u. a., 2004), führte Pfizer von 2002 bis 2005 die sogenannte SUPER-1 Studie („Sildenafil Use in Pulmonary Hypertension“) durch (Galie u. a., 2005). 2005 wurde Sildenafil als Revatio® zur Behandlung von pulmonaler Hypertension zugelassen. Im Jahr 2007 stellte Revatio® mit ca. 200 Mio. US$ 10 % der gesamten Sildenafil-Umsätze von Pfizer (Pfizer Inc., 2008a).

Die Entwicklung des indikationsübergreifenden Zielmoleküls PDE 5 endet jedoch nicht mit der Markteinführung von Revatio®. Zahlreiche Veröffentlichungen beschreiben klinische Evidenz der PDE 5-Inhibition durch Sildenafil bei der Behandlung des Raynaud-Syndroms (Lichtenstein, 2003; Fries u. a., 2005; Gore und Silver, 2005; Kamata u. a., 2005). Für die Behandlung von Herzinsuffizienz wird Sildenafil ebenso diskutiert (Katz u. a., 2000; Maurice u. a., 2003; Hassan und Ketat, 2005; Patel und Katz, 2005; Takimoto u. a., 2005; Wilkins u. a., 2005) wie für eine mögliche Reduktion der Höhenkrankheit (Ghofrani u. a., 2004b; Perimenis, 2005; Richalet u. a., 2005; Hsu u. a., 2006; Fagenholz u. a., 2007) und chronischer obstruktiver Lungenerkrankungen (Barbera u. a., 2003; Voelkel und Cool, 2003; Naeije, 2005; Alp u. a., 2006). Die Notwendigkeit von PDE 5 für kognitive Fähigkeiten beim Hund (Inoha u. a., 2002) und bei der Ratte (Buerk u.

a., 2003) sowie die Beobachtung, dass Sildenafil im Tiermodell die Angio- und Neurogenese nach einem Schlaganfall fördert (Zhang u. a., 2002; Zhang u. a., 2003), eröffneten die Diskussion für eine mögliche Bedeutung des Wirkstoffs bei cerebralen Ischämien (Rosengarten u. a., 2006). In wie weit sich das Indikationsspektrum von Sildenafil in den kommenden Jahren noch verbreitern wird, bleibt abzuwarten.

3.3.2 Fallbeispiel: Indikationsübergreifende Funktion von VEGF-A

Ein anderes Beispiel für ein indikationsübergreifendes Zielmolekül ist der Wachstumsfaktor VEGF-A („Vascular Endothelial Growth Factor A“). In seiner Funktion als Gefäßwachstumsstimulator löst VEGF-A primär über den Rezeptor VEGFR2 eine intrazelluläre Reaktionskaskade aus, die über verschiedene molekulare Effekte den Erhalt des bestehenden Endothels und die Neubildung von Gefäßen fördert (Ellis und Hicklin, 2008). In diesem Zusammenhang ist VEGF-A nicht nur für die Embryonalentwicklung und die Wundheilung notwendig, sondern spielt auch eine wichtige Rolle für das Wachstum und den Erhalt von Gefäßen bei Krebserkrankungen.

Historisch gesehen wurde im Jahr 1971 erstmals die Bedeutung der Angiogenese für Tumorwachstum und Metastasierung beschrieben (Folkman, 1971). Im weiteren Verlauf wurde die Erforschung und die pharmazeutische Nutzung von VEGF-A stark durch die Firma Genentech geprägt. Es vergingen 18 Jahre bis die Arbeitsgruppe von Dr. Ferrara bei Genentech den chemischen Faktor, der für die Wirkung auf das Endothel verantwortlich ist, als VEGF beschrieb (Ferrara und Henzel, 1989). Zusätzlich war schon 1993 ein Faktor identifiziert worden, der die Permeabilität von Gefäßen beeinflusst, und von dem sich später herausstellen sollte, dass er mit VEGF identisch ist (Senger u. a., 1983).

Anfang der 90er Jahre wurde von Genentech der erste Anti-VEGF-Antikörper A4.6.1 entwickelt und dessen inhibitorische Wirkung auf Tumorzellen bewiesen (Kim u. a., 1993). In den folgenden Jahren wurde durch gerichtete Mutagenese ein humanisierter Antikörper namens Bevacizumab erstellt (Presta u. a., 1997), für dem im Januar 1997 ein Antrag zur Zulassung in die klinische Entwicklung gestellt wurde. Im April desselben Jahres begannen die Phase I Studien, die die

Verträglichkeit von Bevacizumab zeigen konnten (Gordon u. a., 2001; Margolin u. a., 2001). Aufgrund der vielversprechenden Ergebnisse startete Genentech 1998 ein großes Phase II Programm mit fünf Studien zur Untersuchung der Wirkung von Bevacizumab als Monotherapie und Kombitherapie in verschiedenen Krebserkrankungen (Ferrara u. a., 2004). Die Studien konnten eine deutliche Erhöhung des progressionsfreien Überlebenszeitraums bei Nierenzellkarzinomen (Yang u. a., 2003) und Kolorektalkarzinomen (Kabbinavar u. a., 2003) zeigen und mündeten in mehreren Phase III Studien. Im Februar 2004 bekam Bevacizumab die Erstzulassung als Avastin® in den USA zur Behandlung von metastasierendem Kolorektalkarzinom in Kombination mit Chemotherapie (Genentech Inc., 2008c).

Genentech profitierte von der frühen Erkenntnis der übergreifenden Funktion von Anti-Angiogenese-Therapie bei Krebserkrankungen und seinem breit aufgestellten klinischen Entwicklungsprogramm. Von 2005 bis 2008 folgten weitere Zulassungen von Avastin® in zusätzlichen onkologischen Indikationen in und außerhalb der USA (Ferrara u. a., 2007). Roche übernahm außerhalb der USA den Vertrieb von Avastin® und trug maßgeblich zur Erweiterung des Indikationsspektrums bei (Hoffmann-La Roche, 2008a; Hoffmann-La Roche, 2008b). Die kontinuierlichen Neuzulassungen konnten den Umsatz, den Genentech und Roche weltweit mit Avastin® erreichen konnten, von 2004 bis 2007 jährlich um über 80 % steigen lassen, so dass die Avastin®-Umsätze 2007 mehr als vier Milliarden US $ erreichten (Hoffmann-La Roche, 2005; Hoffmann-La Roche, 2006; Hoffmann-La Roche, 2007; Genentech Inc., 2008a; Hoffmann-La Roche, 2008c).

Neben der wichtigen Rolle von Angiogenese für Krebserkrankungen konnten Forschungsergebnisse der späten 90er Jahre auch eine Rolle von VEGF-A bei anderen Krankheiten feststellen, die durch Gefäßneubildung geprägt sind. Insbesondere Erkrankungen des Auges wie diabetische Retinopathie standen dabei im Vordergrund, da es hierfür Tiermodelle gab (Aiello u. a., 1994; Malecaze u. a., 1994; Miller u. a., 1994; Alon u. a., 1995). Trotz Fehlens eines Tiermodells entschloss sich Genentech 1997 nach Abschluss der Humanisierung des Avastin® Antikörpers in Kooperation mit der Business Unit Novartis Ophthalmics die Anti-VEGF-Therapie bei neovaskulärer altersabhängiger

Makuladegeneration (AMD) zu testen (Ferrara u. a., 2006; Genentech Inc., 2008b). Motivation hierfür war laut Ferrara vor allem der hohe medizinische Bedarf (Ferrara u. a., 2006). Bei der neovaskulären AMD kommt es durch Ansammlung von extrazellulärem Material (sogenannten Drusen) in der Netzhaut zu einer Schädigung des Pigmentepithels mit folgender inflammatorischer Reaktion und VEGF-vermittelter Gefäßneubildung. Dieser Vorgang führt in der Regel zur Erblindung (Zarbin, 2004; de Jong, 2006; Jager u. a., 2008).

1997 bis 1999 entwickelte Genentech ein dem Bevacizumab verwandtes humanisiertes, rekombinantes Antikörperfragment (Fab-Fragment), das Ranibizumab genannt wurde. Dieses Fab-Fragment ist deutlich kleiner als der Antikörper Bevacizumab und hat eine etwa 100fach höhere Bindungsaffinität zu VEGF-A (Ferrara u. a., 2006). Nach Einreichen des Antrags zur Zulassung zur klinischen Prüfung im Oktober 1999 folgten die Phase I und Phase II Studien bis 2006 und bestätigten die Wirkung von Ranibizumab bei Patienten mit neovaskulärer AMD (Rosenfeld u. a., 2005; Heier u. a., 2006; Rosenfeld u. a., 2006b). Die zwei großen Phase III Studien MARINA (Rosenfeld u. a., 2006a) und ANCHOR (Brown u. a., 2006; Kaiser u. a., 2007) führten schließlich zur Zulassung von Ranibizumab als Lucentis® zur Injektion in den Glaskörper für die Behandlung von neovaskulärer AMD in den USA im Juni 2006 und in der EU im Januar 2007 (Ferrara u. a., 2006).

Lucentis® wird in den USA und in Mexiko von Genentech vertrieben; Novartis hat die Marketingrechte für andere Länder (Genentech Inc., 2008b). Obwohl die Therapie mit Lucentis® sehr kostspielig ist, gilt sie als beste Behandlung für neovaskuläre AMD (Jager u. a., 2008). Die Behandlungsalternative mit dem RNA-Aptamer Macugen® zeigte einen Abfall der Wirksamkeit nach einem Jahr (Gragoudas u. a., 2004) und die ANCHOR-Studie konnte die Überlegenheit von Lucentis® über den Photosensitizer Visudyne® klinisch belegen (Brown u. a., 2006). Lucentis® konnte nach Markteinführung in 2006 bereits 2007 alleine in den USA einen Umsatz von 825 Mio. US $ generieren (Genentech Inc., 2008a).

3.4 Ansätze zur Unterstützung von Forschung und Entwicklung

In Zusammenhang mit den im Abschnitt 3.2.2 beschriebenen Herausforderungen für pharmazeutische Forschung und Entwicklung gibt es verschiedene konzeptionelle und computerunterstützte Unterstützungsansätze. Im Folgenden ein Überblick über die verfügbaren Ansätze gegeben werden.

3.4.1 Ansätze zur Unterstützung der Auswahl von geeigneten Zielmolekülen

Die Auswahl eines geeigneten Zielmoleküls für ein neues Forschungsprojekt ist von immenser Bedeutung. Neben den biologischen und medizinischen Zusammenhängen spielen unter anderem auch ökonomische, legale, technische, und politische Einflussfaktoren eine Rolle (Knowles und Gromo, 2003). Dennoch gibt es nur verhältnismäßig wenige Werkzeuge, die die Auswahl von geeigneten Zielmolekülen unterstützen. Beinahne alle haben ihren Schwerpunkt auf der Identifikation von vollständig neuen Zielmolekülen.

Basierend auf den Eigenschaften bereits bekannter Zielmoleküle werden typische Charakteristika von therapeutisch nutzbaren Zielmolekülen identifiziert (Overington u. a., 2006; Zheng u. a., 2006; Xu u. a., 2007). Aus diesen können dann Regeln definiert werden, die als Richtlinie bei der Erschließung von neuen Zielmolekülen hilfreich sein sollen. Overington und andere betonen vor allem die Zugehörigkeit eines möglichen neuen Zielmoleküls zu bestimmten Proteinfamilien, die Zelloberflächenlage des Moleküls und das Vorkommen von bestimmten Bindedomänen, die typischerweise von Medikamenten erkannt werden („drug-binding domains“) (Overington u. a., 2006). Xu und andere führen in diesem Zusammenhang den Begriff der „drug-target-likeness“ ein, mit dem sie indikationsunabhängig beschreiben, welche Moleküle über typische Charakteristika erfolgreicher Zielmoleküle verfügen (Xu u. a., 2007).

Andere computergestützte Methoden verwenden vor allem Ähnlichkeiten zu bereits bekannten Zielmolekülen, um neue mögliche Zielmoleküle zu identifizieren (Wang u. a., 2004; Russ und Lampel, 2005). Dabei wird unter

anderem auf das Vorkommen von „drug-binding domains“ an den Zielmolekülen geachtet, das durch Sequenzvergleiche bestimmt werden kann (Kramer und Cohen, 2004; Russ und Lampel, 2005).

3.4.2 Ansätze zur Integration systembiologischer Daten in Forschung und Entwicklung

Um die Fülle von verfügbaren systemischen Daten für pharmazeutische Forschung und Entwicklung nutzen zu können, sind computergesteuerte Methoden unerlässlich (Nicholson und Wilson, 2003; Nikolsky u. a., 2005; Schrattenholz und Soskic, 2008).

Es existiert eine Fülle von unterschiedlichen Ansätzen, die in drei Kategorien eingeteilt werden können. Die erste Kategorie enthält *in silico* Modellierungen zur Voraussage von Metabolismus und Toxikologie neuer Leitsubstanzen basierend auf bekannten Zusammenhängen von chemischer Struktur und Aktivität. In die zweite Gruppe fallen kumulierte „high-throughput-screening“-Daten, mit denen ein Profil der Wirkung einer Substanz auf eine bestimmte Zellpopulation oder Körperflüssigkeit modelliert werden kann. Im Zusammenhang mit der vorliegenden Arbeit ist vor allem die dritte Kategorie von Bedeutung, die Datenbanken über zelluläre Funktionen und Reaktionen enthält (Bugrim u. a., 2004).

Gemeinsamer Aspekt der systembiologischen Datenbanken der dritten Kategorie ist die Abbildung der Zusammenhänge als Interaktionsnetzwerk. Je nach Inhalt und Zielsetzung der Anwendung stehen drei Untergruppen von Datenbanken zur Verfügung (Nikolsky u. a., 2005). Die erste Untergruppe enthält Datenbanken, die Protein-Protein-Interaktionen abgespeichert haben. Diese gehen sowohl auf manuell überarbeitete Textsuchen als auch auf experimentelle Daten aus Yeast-2-Hybrid Versuchen und Ko-Immunopräzipitationen zurück (Nikolsky u. a., 2005). Neben zahlreichen öffentlich verfügbaren Datenbanken akademischen Ursprungs (Bader u. a., 2003; Peri u. a., 2004; Salwinski u. a., 2004) gibt es auch einige kommerzielle Anbieter, die Informationen über Protein-Protein-Interaktionen in ihre Softwarepakete zur Auswertung von experimentellen Daten integriert haben. Beispiele hierfür sind unter anderem die MetaCore™ Datenbank (GeneGo

bioinformatics software Inc., 2008), die PathArt™ Datenbank (Jubilant Biosys, 2008), die Pathway Analysis™ Datenbank (Ingenuity Systems Inc., 2008) und die ResNet™ Datenbank (Ariadne Inc., 2008).

Die zweite Untergruppe von systembiologischen Datenbanken enthält Informationen über Signalwege und physiologische Prozesse. Auch hier sind einige große Datenbanken akademischen Ursprungs und daher öffentlich verfügbar: die Kyoto Encyclopedia of Genes and Genomes (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa Laboratories, 2008; Kanehisa u. a., 2008), die Gene Ontology Datenbank (Ashburner u. a., 2000) und die Datenbank Reactome (Joshi-Tope u. a., 2005). Darüber hinaus existieren kommerzielle Angebote wie zum Beispiel Protein Lounge™ (Protein Lounge, 2008).

In der dritten Untergruppe sind vor allem komplette Softwarepakete enthalten, die zur Analyse von großen Datenmengen im Zusammenhang mit Protein-Protein-Interaktionen und Signalwegen gedacht sind. Diese, in der Regel kommerziell erhältlichen Produkte, sollen zum Beispiel die Auswertung von Microarray-Daten erleichtern und durch Integration von systembiologischen Informationen auch minimale Veränderungen in der Genexpression identifizieren können (Nikolsky u. a., 2005; Draghici u. a., 2007; Wang u. a., 2008). Softwarepakete werden unter anderem von den Firmen GeneGo (GeneGo bioinformatics software Inc., 2008), Ingenuity (Ingenuity Systems Inc., 2008), Jubilant Biosys (Jubilant Biosys, 2008) und Ariadne (Ariadne Inc., 2008) angeboten.

Zusammenfassend kann festgestellt werden, dass es zahlreiche Ansätze zur Integration systembiologischer Informationen in die Forschung und Entwicklung gibt. Schwerpunkt dieser Werkzeuge ist jedoch meist die Unterstützung der Auswertung von großen Datenmengen aus „high throughput screening“ oder Microarrays. Keines der verfügbaren Werkzeuge bietet die Möglichkeit, gezielt indikationsübergreifende Zielmoleküle oder Signalwege zu identifizieren.

4 Zielsetzung der Arbeit

Um dem Innovationsdruck im Bereich Forschung und Entwicklung (F&E) der pharmazeutischen Industrie in Zukunft gerecht zu werden, wird die Erforschung und Nutzung von gebietsübergreifenden Zielmolekülen von außerordentlicher Bedeutung sein. In diesem Zusammenhang sollte sowohl das zunehmende Wissen über biologische Signalwege als auch die wachsende Erfahrung der pharmazeutischen Industrie genutzt werden.

Ziel der vorliegenden Arbeit ist die Entwicklung eines systematischen Verfahrens zur Identifikation von indikationsübergreifenden Signalwegen und Zielmolekülen. Dieses könnte sowohl von pharmazeutischen Unternehmen als auch akademischen Forschungsgruppen genutzt werden, F&E-Aktivitäten gezielt auf indikationsübergreifende Signalwege und Zielmoleküle auszurichten.

In standardisierter und automatisierter Weise sollen konkrete Empfehlungen für die Ausrichtung von F&E-Aktivitäten ermittelt werden. Über die direkten Verbindungen zwischen therapeutischen Gebieten durch gemeinsame indikationsübergreifende Zielmoleküle hinaus, sollen in dieser Arbeit erstmals auch indirekte Verbindungen durch gemeinsame funktionelle Signalwege berücksichtigt werden.

Um hierbei die Ausrichtung von Forschung und Entwicklung in unterschiedlichen Spezialisierungsgraden zu ermöglichen (z.B. Neuausrichtung therapeutischer Gebiete vs. Ausrichtung innerhalb bestehender Schwerpunktindikationen und Signalwege), muss das Verfahren die folgenden Fragen beantworten können:

1. Gibt es Kombinationen von **therapeutischen Gebieten**, bei denen die Erfolgswahrscheinlichkeit indikationsübergreifende Medikamente zu entwickeln, besonders groß ist?
2. Welche **Indikationsgruppen** eignen sich zur Fokussierung von F&E-Aktivitäten in ausgewählten therapeutischen Gebieten, wenn die Chance für indikationsübergreifende Medikamentenentwicklung erhöht werden soll?

3. Welche **Signalwege** sind für die Verbindung von ausgewählten Indikationsgruppen besonders wichtig und könnten damit einen interessanten F&E-Schwerpunkt darstellen?
4. Welche **Zielmoleküle** könnten Forschungskandidaten mit einer erhöhten Chance für die Entwicklung von indikationsübergreifenden Medikamenten sein?
5. Mit welchen Kriterien kann eine Auswahl von **Zielmolekülkandidaten** für Forschung und Entwicklung **priorisiert** werden?

5 Materialien

Rohdaten/Programme	*Herkunft*
KEGG: Kyoto Encyclopedia of Genes and Genomes	Kanehisa Laboratories, Kyoto (Japan) (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa Laboratories, 2008; Kanehisa u. a., 2008)
Pharmaprojects - Pharmaceutical Research & Development Pipeline Intelligence	Informa Healthcare, London (Großbritannien) (Informa UK Ltd, 2008)
Microsoft® Office Access 2003	Microsoft Corporation, Redmond (USA)
Microsoft® Office Excel 2003	Microsoft Corporation, Redmond (USA)
GraphPad Prism® 5.0	GraphPad Software, Inc., La Jolla (USA)

6 Methoden

6.1 Theoretische Grundlagen der Datenbankmethodik

Die in der dieser Arbeit aufgebaute Datenbank orientiert sich an den theoretischen Grundlagen zu relationalen Datenbanken von Edgar Codd (Codd, 1970) sowie dem Entity-Relationship-Modell (Chen, 1975). Im Folgenden werden diese Grundlagen in Kürze angesprochen und in den Kontext zur in der Doktorarbeit verwendeten Datenbank gestellt.

6.1.1 Grundlagen des Entity-Relationship-Modell

Das Entity-Relationship-Modell (ERM) beschreibt ein Modell zur Abbildung der realen Welt in der Datenmodellierung und geht auf Peter Chen zurück (Chen, 1975). Im historischen Zusammenhang erfolgte die Entwicklung des ERM nach der Veröffentlichung der theoretischen Grundlagen zu relationalen Datenbanken durch Edgar Codd (Codd, 1970). Beim Aufbau einer Datenbank ist es zweckmäßig, das Modell als ersten Schritt zu betrachten.

Der Zielsetzung des ERM ist die Abbildung einer realen Situation mit Hilfe von Parametern aus bestimmten Kategorien. Diese kategorische Darstellungsweise soll dazu dienen, unterschiedliche Daten auf die gleiche Weise darzustellen und damit die Datenmodellierung zu standardisieren und zu vereinfachen.

Der ersten Baustein des ERM sind die Objekte („Entities“), die reale Äquivalente (z.B. Personen, Zielmoleküle) haben und durch Attribute charakterisiert sind. Ein charakterisierendes Attribut kann je nach Objekt ganz verschiedene Ausprägungen haben (z.B. Name für eine Person, Gen Identifizierungsnummer für ein Zielmolekül). Verschiedene Objekte sind durch Beziehungen („Relationships”) untereinander verbunden, wobei die Art der Beziehung variieren kann. In diesem Zusammenhang ist besonders die Kardinalität einer Beziehung von Bedeutung. Diese beschreibt, ob ein Objekt eine eindeutige Beziehung zu einem anderen Objekt hat (z. B. gehört eine Indikation eindeutig in ein

therapeutisches Gebiet) oder eine mehrdeutige Beziehung (z.B. kann ein Zielmolekül Teil von verschiedenen Signalwegen sein). Vereinfacht formuliert soll eine reale Situation nach dem Entity-Relationship-Modell vollständig in Objekte, Attribute und Beziehungen zerlegt werden, um sie für die Datenmodellierung zugängig zu machen. Diese Zerlegung kann durch ERM-Diagramme unterstützt werden, die die einzelnen Strukturen in standardisierten geometrischen Formen und Linien darstellen. Es existiert eine Vielzahl von unterschiedlichen Notierungen dieser Diagramme (Smith und Smith, 1977; Storey, 2005). Abb. 6.1 zeigt die Datenstruktur der vorliegenden Arbeit vereinfacht in der sogenannten Chen-Notation (Chen, 1975).

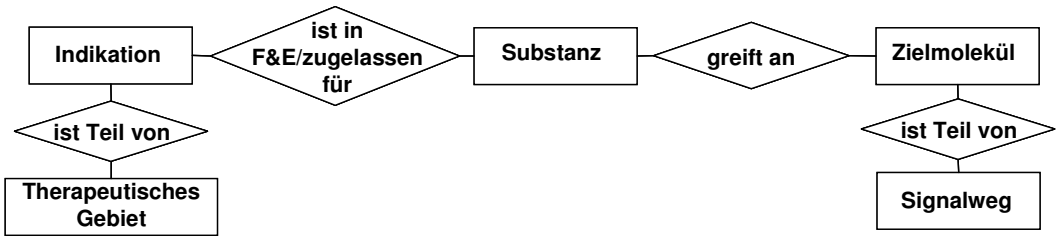


Abb. 6.1: **Der Datenkontext dieser Arbeit kann als Entity-Relationship-Modell in der Chen-Notation dargestellt werden**
Die zum Aufbau der Datenbank notwendigen Informationen wurden in Objekte (dargestellt als Rechtecke) und Beziehungen der Objekte untereinander (dargestellt als Rauten) zerlegt.
Die Abbildung macht deutlich, dass zwischen allen Objekten der Datenbank Beziehungen bestehen, die indirekt Verbindungen zwischen allen Objekten schaffen.
Unterschiedliche Kardinaliäten der Beziehungen können in der Chen-Notation nicht unmittelbar abgelesen werden. So ist zum Beispiel das Objekt Indikation eindeutig ein Teil eines bestimmten therapeutischen Gebiets während ein Zielmolekül Teil mehrere Signalwege sein kann.

6.1.2 Grundlagen relationaler Datenbanken

6.1.2.1 Konzept und Struktur

Das Konzept der relationalen Datenbanken wurde Ende der 60er Jahre von Edgar Codd entwickelt und dient der Speicherung und Analyse von großen Datenstrukturen (Codd, 1970). Grundprinzip der relationalen Datenbank ist, dass bestimmte Objekte mit ihren Attributen in einzelnen Tabellen („Records“) gespeichert werden und die Tabellen über Beziehungen der Objekte miteinander verbunden sind.

Dies hat den Vorteil, dass verschiedene Informationen zu bestimmten Objekten, die eine unterschiedliche Datenstruktur haben, in logischem Zusammenhang

miteinander abgebildet werden können (z.B. Mitarbeiter einer Abteilung vs. Projekte einer Abteilung oder Medikament, das ein bestimmtes Zielmolekül adressiert vs. Signalwege, in denen ein Zielmolekül relevant ist). Ohne die Verwendung von separaten Tabellen und der Verbindung der Tabellen durch Beziehungen wäre eine Darstellung dieser Datenstruktur nur sehr begrenzt möglich und mit Redundanzen verbunden (z.B. müssten pro Mitarbeiter einer Abteilung alle Projekte gelistet werden).

Für die Struktur einer relationalen Datenbank gilt, dass die Attribute der Objekte so in den Spalten der jeweiligen Tabelle gespeichert werden müssen, dass jede Zeile der Tabelle als Datentupel eindeutig charakterisiert ist. Jede Tabelle muss dabei aus mindestens zwei Spalten bestehen. Dasjenige Attribut, welches ein Tupel eindeutig identifiziert, wird als „Primary key“ oder Primärschlüssel markiert. Eindeutige Schlüssel können sich auch aus mehreren Attributen zusammensetzen (z.B. Name eines Medikaments und Entwicklungsphase). Die korrekte Verteilung von Schlüsseln und die Verbindung von Objekten über gemeinsame Eigenschaften (sogenannte Fremdschlüssel) werden als referentielle Integrität bezeichnet.

Um Redundanzen sowie mögliche Inkonsistenzen beim Aktualisieren einzelner Datenpunkte zu vermeiden, sollte eine relationale Datenbank normalisiert werden (Codd, 1970; Codd, 1971; Delobel, 1978). Vereinfacht ausgedrückt beschreibt die Normalisierung eine schrittweise Zerlegung der Relationen in einzelne Tabellen. Grundsätzlich sind Zerlegungen bis zur sogenannten vierten Normalform möglich (Fagin, 1977). Je weiter eine relationale Datenbank allerdings normalisiert wird, desto höher wird auch die Zahl der Inputtabellen. Dadurch verlängert sich die Rechenzeit. Aus Gründen der Performance und der Übersichtlichkeit wurde daher im Rahmen der vorliegenden Arbeit bewusst auf Normalisierungen über die erste Normalform hinaus verzichtet.

6.1.2.2 Relationale Algebra

Innerhalb einer relationalen Datenbank, in der alle Informationen in Form von Tabellen gespeichert sind, können Daten mit Hilfe von Abfragen gesucht werden.

Dabei werden die Beziehungen der Daten mit Hilfe von Operationen der sogenannten relationalen Algebra berücksichtigt (Codd, 1970).

Die folgenden Operationen sind Begriffe der relationalen Algebra:

- **Projektion:** Streichung einzelner Spalten einer Tabelle; bei Entstehung von Mehrfachtupel können diese zu einem Tupel zusammengefasst werden (auch sogenannte Attributsbeschränkung)
- **Restriktion:** Extraktion einzelner Tupel aus einer Tabelle durch Vergleich mit einer Selektionsbedingung innerhalb einer Tabelle (auch Selektion)
- **Verbund:** Zusammenfügen von mehreren Tabellen über gemeinsame Attribute (Spalten)
- **Division:** Begrenzung auf diejenigen Tupel einer Tabelle, die mit allen Tupeln einer weiteren Tabelle in Verbindung stehen
- **Durchschnitt:** Begrenzung auf alle Tupel, die in zwei Tabellen vorkommen
- **Vereinigung:** Zusammenlegung von Tupeln aus separaten Tabellen in einer gemeinsamen Tabelle
- **Differenz:** Entfernen von Tupeln aus einer Tabelle, sofern diese auch in einer anderen vorhanden sind
- **Umbenennung:** Veränderung von Attributsnamen (Spaltennamen)

Darüber hinaus sind durch Kombination weitere Operationen wie die Differenzbildung, die Umbenennung von Spalten und die Bildung eines Kreuzprodukts aus zwei Tablelen möglich. Grundsätzlich gilt, dass alle Operationen der relationalen Algebra manuell durchgeführt werden könnten, allerdings ist dies aufgrund der großen Datenmengen praktisch nicht möglich.

6.2 Rohdaten

Zum Aufbau der Datenbank wurden zweierlei externe Rohdaten verwendet. Zum einen Informationen über Medikamente, die sich bereits auf dem Markt befinden, bzw. momentan in der Entwicklung sind und zum anderen biologische Daten, die die molekulare Zusammensetzung von Signalwegen beschreiben.

Um Inkonsistenzen zu vermeiden, wurden die englischen Bezeichnungen aus den Rohdaten für den Aufbau der Datenbank übernommen. Die Struktur (Tabellen, Beziehungen) sowie die funktionellen Elemente (Abfragen, Makros) der Datenbank sind daher in englischer Sprache angelegt und werden ebenso im Folgenden aufgeführt.

6.2.1 Pharmaprojects

6.2.1.1 Ursprung der Daten

Pharmaprojects ist ein Unternehmen der Informa Healthcare UK Gruppe (Informa UK Ltd, 2008), das darauf spezialisiert ist, pharmazeutische Forschungs- und Entwicklungspipelines zu verfolgen und öffentlich verfügbare Informationen hierzu zusammenzuführen. Pharmaprojects verfügt über eine umfangreiche Datenbank für öffentlich bekannte Medikamente bzw. öffentlich bekannte Substanzen in Forschung und Entwicklung. Der für den Aufbau der Datenbank verwendete Ausschnitt der Pharmaprojectsdaten wurde im Rahmen der Lizenz der Firma McKinsey&Company, Inc. in Absprache mit Pharmaprojects bezogen und von McKinsey&Company, Inc. freundlicherweise für die vorliegende Arbeit zur Verfügung gestellt. Entsprechend der Lizenzbedingungen und in Absprache mit Pharmaprojects und McKinsey&Company, Inc. werden in dieser Arbeit keine Rohdaten der Pharmaprojects Datenbank preisgegeben. Verwendete Beispiele stellen stets Informationen dar, die auch ohne Zugang zur Pharmaprojects Datenbank öffentlich verfügbar sind.

6.2.1.2 Aufbereitung der Rohdaten

Für den Aufbau der Datenbank wurden ausschließlich Daten über Substanzen verwendet, die sich in präklinischer Forschung oder klinischer Entwicklung befinden oder bereits auf dem Markt sind. Substanzen mit abgebrochener oder unterbrochener Entwicklung wurden ausgeschlossen, um die Analyse nicht durch negative Datenpunkte zu beeinflussen.

Für den Aufbau der Datenbank wurden ausgewählte Felder der Rohdaten aus dem Microsoft Excel Format aufbereitet und in Microsoft Access importiert. Eine Aufbereitung der Daten war notwendig, da in dem Datenauszug der Pharmaprojects Datenbank mehrere Informationen (z.B. mehrere Indikationen für eine Substanz „Generic Name“) nicht durch verschiedene Zeilen, sondern durch einen Zeilenumbruch innerhalb derselben Zelle getrennt waren (siehe Abb. 6.2). Dieser Zeilenumbruch innerhalb derselben Zelle wird in Microsoft Excel als sogenanntes „soft enter“ bezeichnet und mit dem Kürzel „char10“ beschrieben.

Generic name	Any indication	Any indication status
Substance A	Indication 1 Indication 2 Indication 3 Indication 4	Launched Launched Launched Launched

Abb. 6.2: Beispiel der Datenstruktur des Auszugs aus den Pharmaprojects Rohdaten vor der Datenaufbereitung
Die Tabelle zeigt einen Ausschnitt der Rohdaten für die Substanz Promestriene, wie sie in Microsoft Excel vor der Datenaufbereitung vorlag. In der ersten Spalte ist der Name der Substanz vermerkt; diese Spalte trägt den Namen „Generic name“, da es sich bei dem Substanznamen in der Regel um eine generische Bezeichnung der chemikalischen Substanz handelt. In der zweiten Spalte „Any indication“ stehen alle Indikationen, für die sich die Substanz in der Entwicklung befindet. In der dritten Splate „Any indication status“ ist der Entwicklungsstatus der Substanz in der jeweiligen Indikation vermerkt.
Inhaltlich kann abgelesen werden, dass für die Substanz „Substance A“ im verwendeten Datensatz vier verschiedene Indikationen vermerkt sind. In allen Indikationen wurde „Substance A“ bereits auf den Markt gebracht.
Strukturell zeigt das Beispiel, dass alle Indikationen mit ihrem dazugehörigen Entwicklungsstatus in einer gemeinsamen Zelle standen und nur durch Zeilenumbruch („soft enter“) getrennt waren.

Eine Trennung dieser Daten war wichtig, da jede Indikation einer Substanz zur Bedeutung der Substanz für das therapeutische Gebiet, zu dem die Indikation gehört, beiträgt. Da weder Microsoft Excel noch Microsoft Access bei Rechenoperationen Zeilenumbrüche innerhalb einer Zelle unterscheiden können, war ein Aufbereitungsschritt notwendig. Im Rahmen des Datenaufbereitung und aller folgenden Schritte wurden die Originalbezeichnungen der Tabellenfelder aus

den Rohdaten beibehalten. Die von der Aufbereitung betroffenen Felder (z.B. „Generic name“ und „Any indication“) wurden in ein separates Exceldokument überführt und mit Hilfe eines Makros getrennt. Das Makro teilte bei Erkennen des „char10“ die Zelle und fügte eine neue Zeile ein, in die es die gleichbleibende Information kopierte. Dabei wurden zusätzliche Informationen, wie zum Beispiel der Status einer bestimmten Indikation, eindeutig in der gleichen Zeile der entsprechenden Indikation zugeordnet (siehe Abb. 6.3).

Generic name	Any indication	Any indication status
Substance A	Indication 1	Launched
Substance A	Indication 2	Launched
Substance A	Indication 3	Launched
Substance A	Indication 4	Launched

Abb. 6.3: Beispiel des Auszugs aus den Pharmaprojects Rohdaten nach Datenaufbereitung
Die Abbildung zeigt den Zustand des in Abb. 6.2 darstellten Rohdatenausschnitts nach Durchführung der Datenaufbereitung. Im Gegensatz zur Struktur aus Abb. 6.2 sind alle vier Indikationen der Substanz „Substance A“ mit ihrem jeweiligen Entwicklungsstatus jetzt in getrennten Zeilen gespeichert.

Auf diese Weise wurden aus den Pharmaprojects Rohdaten die weiteren Tabellenfelder „Any Indication“/“Any Indication Status“, „Any target name“/“Any Gene ID“ in Abhängigkeit des Tabellenfelds „Generic Name“ getrennt. Die aufbereiteten Rohdaten wurden als Inputtabellen in Microsoft Access importiert und entsprechend den in ihnen enthaltenen Rohdaten „Indications and drugs“ und „Drugs and targets“ genannt.

Um darüber hinaus Informationen über die Eigentumssituation einer Substanz zur Verfügung zu haben, wurde eine dritte Rohdatentabelle in Access aus Pharmaprojects Tabellenfeldern erstellt. Die Tabelle „Drugs and companies“ listet alle Substanzen mit den dazugehörigen Firmen und erklärt, ob es sich bei der Firma um den Entwickler („Originator“) oder einen Lizenznehmer der Substanz handelt. Eine Trennung der dazu benötigten Pharmaprojects Felder „Licensee“/Licensee Status“ erfolgte dabei in Abhängigkeit von „Generic Name“ mit dem oben beschriebenen Makro.

Neben der Aufbereitung der Rohdaten und dem Import der resultierenden Inputtabellen in Access wurde in einem weiteren Arbeitsschritt eine Zuordnung aller in Pharmaprojects enthaltener Indikationen zu therapeutischen Gebieten und Indikationsgruppen durchgeführt. Die Zuordnung erfolgte in Anlehnung an das

„Anatomical Therapeutic Chemical Classification System“ der World Health Organisation (WHO Collaborating Centre for Drug Statistics Methodology, 2008) und verwendete die Information des Pharmaprojects Feldes „Primary Therapy Code“. Weitere Zuordnungen, insbesondere zu funktionellen Indikationsgruppen, wurden durch Diskussion mit Experten der pharmazeutischen Industrie verifiziert. Aufgrund des unterschiedlichen Charakters der Gruppe „Other“, die neben Diagnostika und Mittel zum Gebrauch während Operationen auch alle nicht genauer spezifizierten Indikationen enthält, wurde diese Gruppe von der Analyse ausgeschlossen. Wegen seiner sehr geringen Größe wurde das Gebiet der Erkrankungen im Ohr- und Mundbereich („Ear and Mouth“) ebenfalls nicht analysiert.

6.2.2 Kyoto Encyclopedia of Genes and Genomes (KEGG)

6.2.2.1 Ursprung der Daten

Mit zunehmender Anerkennung der Bedeutung von Signalwegen für Biologie und Medizin gibt es eine Vielzahl von universitären und kommerziellen Initiativen zur Definition von Signalwegen und ihrer molekularen Zusammensetzung. Die in der vorliegenden Arbeit verwendete Quelle ist die Kyoto Encyclopedia of Genes and Genomes (KEGG). KEGG ist ein 1995 gestartetes Projekt des Kanehisa Laboratory im Kyoto University Bioinformatics Center und des Human Genome Centers der University of Tokyo (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa u. a., 2008). Ziel des Projekts ist eine elektronische Abbildung der biologischen Zusammenhänge in den Zellen und der Biosphäre. Im Rahmen dessen bietet KEGG online eine Vielzahl von Datenbanken – unter anderem auch eine Signalweg-Datenbank – an (Kanehisa Laboratories, 2008). Die Datenbank ist für akademische Zwecke frei zugänglich.

6.2.2.2 Aufbereitung der Rohdaten

Für die Analysen der vorliegenden Arbeit wurden ausschließlich funktionale humane Signalwege verwendet. Metabolische Signalwege wurden ausgeschlossen, da ihre Bestandteile häufig chemische Intermediate sind und

daher nicht eindeutig durch eine Gen-Identitätsnummer („Gene ID“) in Pharmaprojects zurückverfolgt werden können.

Im ersten Schritt der Datenaufbereitung wurden die Rohdaten aller verwendeten Signalwege im xml-Format (Bray u. a., 2006) von der KEGG Website heruntergeladen und in Microsoft Excel importiert. Dann wurden die für die Analyse relevanten Felder „title“ (Name des Signalwegs), „name2“ (Identitätsnummer des Signalwegbestandteils) und „type“ (Art des Signalwegbestandteils) ausgewählt und in ein neues Excel Arbeitsblatt kopiert. Im Anschluss wurden mit Hilfe eines Filters in Excel alle Einträge ausgewählt, die ein humanes Gen repräsentieren (Auswahl des „type“: „gene“). Um die Identitätsnummer („Gene ID“) der Gene mit der von Pharmaprojects vergleichbar zu machen, wurde im nächsten Schritt bei allen Einträgen die Kodierung „hsa:“ vor der „Gene ID“ gelöscht. Zu diesem Zweck wurde die „find and replace“-Funktion von Microsoft Excel genutzt.

Ähnlich der oben beschriebenen Problematik der Rohdatenbeschaffenheit von Pharmaprojects lagen auch bei den KEGG Rohdaten Mehrfacheinträge pro Zelle vor. Dabei waren mehrere „Gene IDs“ lediglich durch ein Leerzeichen getrennt. Analog des für die Pharmaprojects Daten beschriebenen Prinzips wurden mit Hilfe eines Makros die Mehrfacheinträge „name2“ in Abhängigkeit des Signalweges („title“) getrennt. Die daraus resultierende Tabelle wurde in einem weiteren Filterschritt mit Hilfe der Excel Funktion „copy unique entries only“ von Doppeleinträgen bereinigt und als Inputtabelle „Targets and pathways“ in Microsoft Access importiert.

6.3 Grundstruktur der Datenbank

6.3.1 Inputtabellen

Zum Aufbau der Datenbank wurden fünf Inputtabellen erstellt und in Microsoft Access importiert. Aufgrund der englischen Sprache der Rohdaten wurden den Inputtabellen ebenfalls englische Namen gegeben, die im Weiteren zur Benennung der Tabellen beibehalten werden.

6.3.1.1 Therapeutische Gebiete, Indikationsgruppen und Indikationen („TAs and indications“)

Diese erste Inputtabelle hinterlegt die Zuordnung von Indikationen zu Indikationsgruppen und therapeutischen Gebieten („TAs“). Sie ist durch die folgenden Felder gekennzeichnet:

- Name des therapeutischen Gebiets („TA name”)
- Name der Indikation („Indication”)
- Name der Indikationsgruppe („Indication group name“)

Alle Felder wurden in Microsoft Access als Text formatiert und das Feld „Indication“ wurde zum Primärschlüssel bestimmt.

Insgesamt wurden 748 verschiedene Indikationen zu 91 Indikationsgruppen und 16 therapeutischen Gebieten zugeordnet. Wie oben erläutert, wurden die therapeutischen Gebiete „Ear and Mouth“ sowie „Other“ aufgrund ihrer Heterogenität bzw. Größe von der Analyse ausgeschlossen, so dass die Inputtabelle „TAs and indications“ 707 Indikationen in 84 Indikationsgruppen und 14 therapeutischen Gebieten enthält.

6.3.1.2 Indikationen und pharmazeutisch aktive Substanzen („Indications and drugs“)

Die zweite Inputtabelle listet alle in der Analyse verwendeten pharmazeutisch aktiven Substanzen (zugelassene Medikamente oder Substanzen in Forschung und Entwicklung) und die Indikationen, für die diese Substanzen erforscht werden

bzw. zugelassen sind sowie den Entwicklungsstatus des jeweiligen Forschungsprojekts. Die Substanzen werden analog zu den Pharmaprojects Rohdaten mit einem generischen Namen bezeichnet („Generic name“), der entweder dem Arbeitsnamen des Herstellers entspricht oder eine Bezeichnung der chemischen Substanz sein kann. Mögliche Synonyme, zum Beispiel bei auslizensierten Medikamenten, wurden nicht in die Datenbank integriert, um keine Vervielfältigung derselben chemischen Substanz zu erhalten. Im Fall einer Lizenz wurde die Substanz für den Lizenznehmer unter dem gleichen Namen vermerkt wie für den Entwickler. Substanzen, die in Pharmaprojects als eigenständig entwickelt vermerkt sind, d.h. durch einen eigenen „Originator“-Eintrag charakterisiert sind, wurden auch in der vorliegenden Datenbank als eigenständige Einträge übernommen.

Durch Aufbereitung der Rohdaten war es möglich, jede Indikation einer Substanz mit ihrem entsprechenden Status als separaten Eintrag zu verwenden. Eine Substanz, die in mehreren Indikationen entwickelt wird, wurde daher auch jeweils einmal in der entsprechenden Indikation aufgenommen.

Die Tabelle „Indications and drugs“ enthält die folgenden Felder:

- Name der Substanz („Generic name“)
- Name der Indikation („Any indication”)
- Status der Entwicklung („Any indication status“)

Alle Felder wurden in Microsoft Access als Text formatiert. Die Felder „Generic Name“ und „Any indication“ wurden zum Primärschlüssel bestimmt. Insgesamt gingen auf diese Weise 11393 Substanzen, die im Durchschnitt für je ca. 1,6 Indikationen entwickelt werden, in die Untersuchungen ein. Dies führte zu insgesamt 18061 unabhängigen Einträgen von Substanz, Indikation und Status. Wie in Abb. 6.4 gezeigt ist, befindet sich die Mehrheit aller Projekte in der präklinischen Forschung. Ein weiterer großer Anteil ist bereits auf dem Markt als Medikament eingeführt. Obwohl bei Gewinnung der Rohdaten alle Substanzen ausgeschlossen wurden, die keine aktive Indikation haben, gibt es für einige Substanzen einzelne Indikationen, deren Entwicklung eingestellt wurde oder nicht

weiter spezifiziert ist. Diese Indikationen sind in der Datenbank enthalten, da die Substanzen in einer anderen Indikation weiterhin aktiv sind und somit nicht komplett von der Untersuchung ausgeschlossen wurden. Diese inaktiven Indikationen wurden im weiteren Verlauf der Untersuchungen durch ein spezielles Gewicht von der Analyse ausgeschlossen (vergleiche Gewichtung Abschnitt 6.4.5)

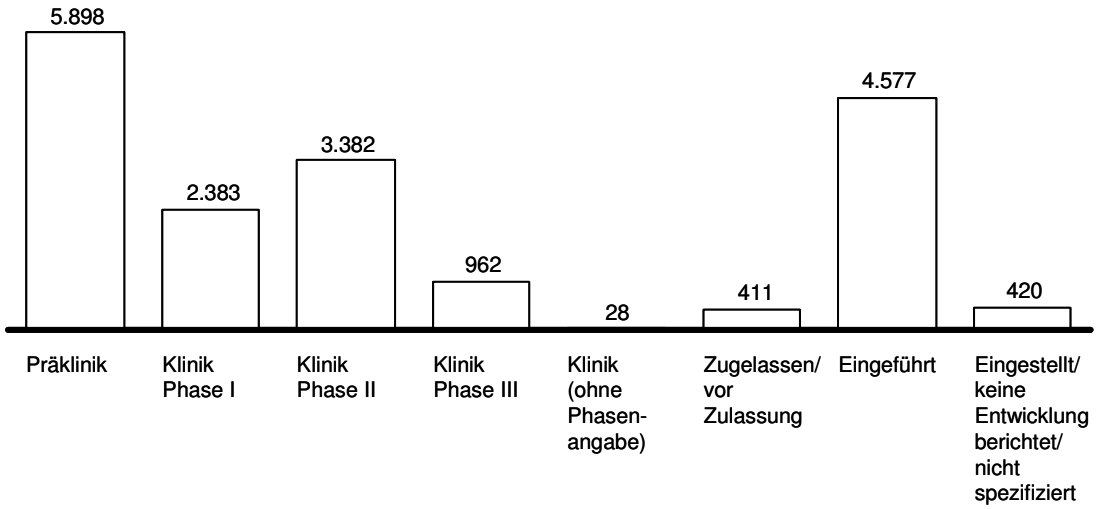


Abb. 6.4: **Die Forschungsprojekte verteilen sich über alle Entwicklungsphasen**
Alle Forschungsprojekte (d.h. die Entwicklung einer pharmazeutisch aktiven Substanz für eine bestimmte Indikation), die in der Inputtabelle „Indications and drugs“ vermerkt sind und in die Analysen eingehen, wurden nach ihrem jeweiligen Forschungs- und Entwicklungsstatus sortiert.
Die Abbildung zeigt, dass sich die Mehrheit der Projekte in der Präklinik („Präklinik“) befindet, oder bereits als Medikament eingeführt wurde („Eingeführt“). 411 Projekte haben die klinische Phase III erfolgreich verlassen, sind aber noch vor bzw. im Zulassungsprozess („Zugelassen/vor Zulassung“). 420 Projekte sind entweder eingestellt worden, nicht weiter spezifiziert oder über sie sind keine Entwicklungsberichte verfügbar („Eingestellt/keine Entwicklung berichtet/nicht spezifiziert“). Diese Projekte werden im weiteren Verlauf von der Untersuchung ausgeschlossen. Die restlichen Projekte verteilen sich auf die klinischen Phasen I-III („Klinik Phase I-III“). Für eine sehr kleine Anzahl von Projekten ist es nicht möglich zwischen den klinischen Phasen zu unterscheiden („Klinik (ohne Phasenangabe)“).

6.3.1.3 Pharmazeutisch aktive Substanzen und Zielmoleküle („Drugs and targets“)

Diese Tabelle enthält die Zuordnung der pharmazeutisch aktiven Substanzen zu den von ihnen adressierten Zielmolekülen sowie die „Gene ID“ der Zielmoleküle. Substanzen, die mehrere Zielmoleküle adressieren, wurden für jedes Zielmolekül separat in die Datenbank aufgenommen.

Die Tabelle „Drugs and targets“ besteht aus folgenden Feldern:

- Name der Substanz („Generic name“)
- Name des Zielmoleküls („Any target name“)
- ID des Zielmoleküls („Any gene ID“)

Alle Felder wurden als Text formatiert; „Generic name“ und „Any gene ID“ wurden als Primärschlüssel definiert. Die meisten der 11393 Substanzen adressieren nur ein Zielmolekül. Eine geringe Anzahl von Substanzen hat jedoch mehrere Zielmoleküle, so dass für die Analyse 13597 unabhängige Kombinationen von Substanz und Zielmolekül zur Verfügung stehen. Für die Analysen der vorliegenden Arbeit wurden nur diejenigen Einträge verwendet, bei denen das Zielmolekül in den Rohdaten eindeutig mit einer „Gene ID“ beschrieben ist. Abb. 6.5 zeigt, dass dies für ca. 60% aller Einträge der Fall ist.

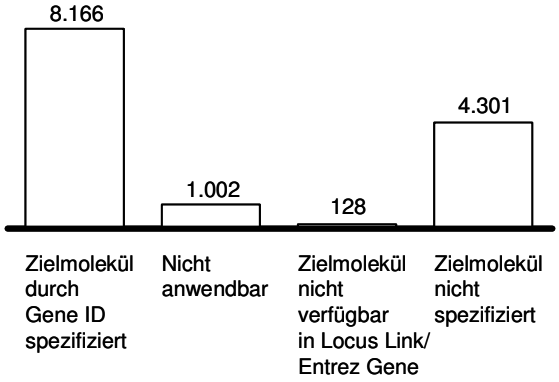


Abb. 6.5: **60% der Kombinationen von Substanz und Zielmolekül sind eindeutig durch eine „Gene ID“ beschrieben**
Für alle Kombinationen aus Substanz und Zielmolekül, die in der Inputtabelle „Drugs and targets“ aufgeführt sind, wurden untersucht, ob eine Spezifizierung der Zielmoleküle durch eine „Gene ID“ vorhanden ist.
Die Abbildung zeigt, dass 60% der Kombinationen von Substanz und Zielmolekül (8166) durch ein eindeutiges Zielmolekül mit dazugehöriger „Gene ID“ beschrieben sind. Bei 1002 Einträgen war entweder für das Zielmolekül keine „Gene ID“ anwendbar, da es sich z.B. um ein chemisches Intermediat handelt oder weil in den Rohdaten anstelle des Zielmoleküls und/oder der „Gene ID“ „not applicable“ vermerkt wurde. Bei 128 Einträgen sind zwar die Zielmoleküle bestimmt, jedoch konnte in der Gendatenbank LocusLink (Maglott u. a., 2000) bzw. der Nachfolgerdatenbank Entrez Gene (Maglott u. a., 2005) keine „Gene ID“ gefunden werden. Bei den letzten 4301 Einträgen war das Zielmolekül der Substanz entweder unbekannt oder nicht angegeben. Für die Analyse wurden nur diejenigen Kombinationen verwendet, bei denen das Zielmolekül angegeben und eindeutig über eine „Gene ID“ spezifizierbar ist.

6.3.1.4 Zielmoleküle und Signalwege („Targets and pathways”)

Diese Tabelle definiert die Zuordnung der Zielmoleküle zu den Signalwegen. Um später eine eindeutige Zuordnung der Zielmoleküle zwischen den einzelnen

Tabellen zu erlauben, wurden in dieser Tabelle die Gen-Identitätsnummern („Gene ID“ oder „Target ID“) zur Beschreibung der Zielmoleküle verwendet. Die Tabelle enthält die folgenden Felder:

- Name des Signalwegs (“Pathway”)
- ID des Zielmoleküls („Target ID“)

Beide Felder wurden als Text und als Primärschlüssel formatiert. Insgesamt wurden 56 humane funktionale Signalwege mit insgesamt 5087 Zielmolekülen in die Datenbank eingespeist.

6.3.1.5 Pharmazeutisch aktive Substanzen und Firmen („Drugs and companies“)

In dieser Tabelle ist die Eigentumssituation einer pharmazeutisch aktiven Substanz hinterlegt. Es wurden für jede pharmazeutisch aktive Substanz sowohl der ursprüngliche Entwickler („Orginator“) als auch mögliche Lizenznehmer eingetragen. Zusätzlich dazu wurde in „Drugs and companies“ der am weitesten fortgeschrittene Entwicklungsstatus der Substanz bei den einzelnen Firmen vermerkt. Das bedeutet, dass eine Substanz, die bei dem Entwickler eventuell bereits in einer Indikation auf den Markt verfügbar ist, für den „Originator“ als „launched“ vermerkt wurde, jedoch für einen Lizenznehmer gegebenenfalls ein früherer klinischer oder präklinischer Status gespeichert wurde. Der in der Inputtabelle vermerkte Status stimmt nicht zwangsweise mit allen Status der einzelnen Indikationen bei einer Firma überein, da lediglich pro Firma der am weitesten fortgeschrittene Status vermerkt wurde.

Die Tabelle hat die folgenden Felder:

- Name der Substanz („Generic name“)
- Name des Entwicklers bzw. Lizenznehmers („Company“)
- Am weitesten fortgeschrittener Status der Substanz in der Firma („Status“)
- Eigentumssituation, d.h. ist die Firma Entwickler oder Lizenznehmer („Origin“)

Alle Felder wurden als Text formatiert; es wurde kein Primärschlüssel vergeben.

6.3.2 Beziehungen zwischen den Inputtabellen

Grundprinzip der Analysen zu indikationsübergreifenden Zielmolekülen und Signalwegen stellt die Zuordnung von Signalwegen zu therapeutischen Gebieten mit Hilfe der adressierten Zielmoleküle dar. Zu diesem Zweck wurden die Inputtabellen in Microsoft Access durch Beziehungen verbunden. Eine Beziehung zwischen einem Feld einer Tabelle und einem Feld gleichen Inhalts (z.B. Name der Substanz „Generic name“) einer anderen Tabelle verknüpft Tabellen logisch. Diese logische Verknüpfung ist notwendig, um unterschiedliche Attribute eines Eintrags aus unterschiedlichen Tabellen zusammenzuführen (z.B. Indikationen einer Substanz und Zielmoleküle derselben Substanz). Dabei erlaubt die Beziehung auch Felder als inhaltlich gleich oder gleichartig zu definieren, die auf Grund der Bezeichnung ihrer Inputdaten unterschiedliche Namen tragen (z.B. „Target ID“ und „Gene ID“). Die Beziehungen zwischen den Inputtabellen sind in Abb. 6.6 dargestellt. Zur Abbildung der Tabellen und ihrer Beziehungen wurden dieselben Symbole wie in Microsoft Access verwendet.

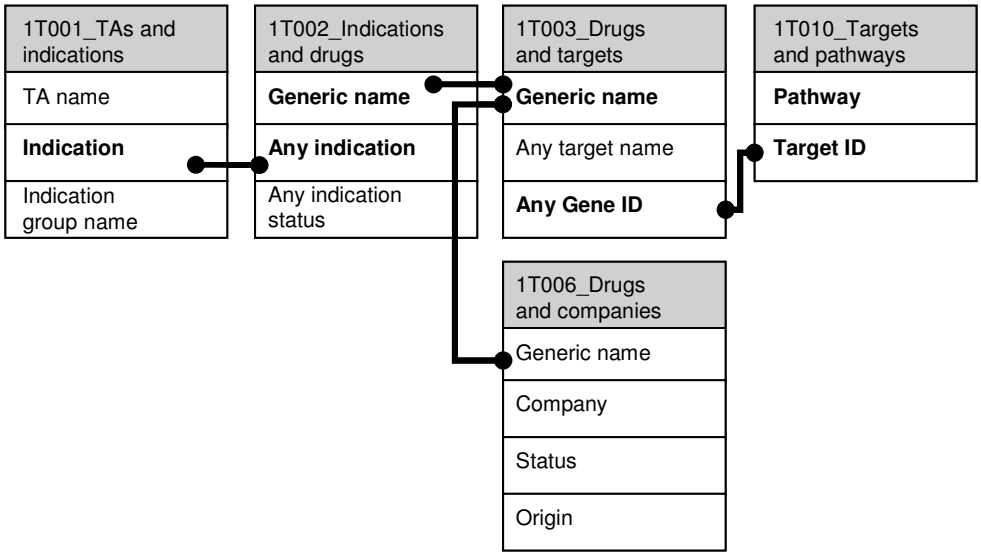


Abb. 6.6: **Die Grundstruktur der Datenbank besteht aus den fünf Inputtabellen und ihren Beziehungen untereinander**
Die Abbildung zeigt die fünf Inputtabellen und die Verbindungen der gleichartigen Felder untereinander. Jede Inputtabelle hat einen Titel, der aus einem Zahlen-/Buchstabencode sowie dem Namen der Tabelle besteht. Während die Zahlencodierung die Tabelle eindeutig kennzeichnet, dient der Name dazu, den Inhalt zu erklären. Alle Inputtabellen tragen eine 1 als erste Ziffer ihrer Codierung. Der Tabellentitel wird im grau unterlegten Rechteck gezeigt; die Felder der Tabelle sind in den darunterliegenden weißen Rechtecken aufgelistet. Fett gedruckte Felder sind Primärschüssel der jeweiligen Tabelle. Die Beziehungen zwischen den Feldern sind durch schwarze Linien gekennzeichnet.

Darüber hinaus wurden zahlreiche Beziehungen zwischen den einzelnen Tabellen in Abfragen definiert. Diese Beziehungen werden im Rahmen der Abfragen erläutert.

6.3.3 Verwendete Abfragetypen und Makros

Zur Analyse der Daten wurden in Microsoft Access verschiedene Abfragetypen („queries“) verwendet. Wie die Inputtabellen wurden die Abfragen auf Englisch verfasst und werden im Folgenden mit ihren englischen Namen benannt.

- **„Append queries“ und „Delete queries“:** Der größte Teil der Datenbankabfragen sind als sogenannte „Append“-Abfragen angelegt. Diese schreiben Daten in eine vorher definierte Intermediärtabelle. Vorteil dieser Abfrageart ist, dass sich die Intermediärtabellen jederzeit zwischen den Abfragen ansehen lassen, und somit einzelne Rechenschritte verfolgt werden können. Damit die Tabellen jedoch vor dem Durchlauf einer erneuten Abfrage gelöscht werden, wurde für jede „Append“-Tabelle eine „Delete“-Abfrage erstellt, die die Inhalte der Intermediärtabelle löscht, aber ihre Struktur erhält.

- **„Make-Table queries“:** Die letzten Abfragen von Analyseabfolgen wurden durch Abfragen abgebildet, die selbstständig eine Tabelle erzeugen. Alle so erzeugten Tabellen sind Outputtabellen. Wird eine Analyse erneut durchgerechnet, so überschreibt die „Make-Table query“ die alte Outputtabelle.

- **„Select Queries“:** Für einzelne Analysen wurden „Select“-Abfragen verwendet, die jeweils das Ergebnis der Abfrage anzeigen, jedoch nicht in Form einer Tabelle speichern. Der Vorteil der „Select Queries“ ist, dass nicht für jede Abfrage eine Intermediärtabelle notwendig ist, jedoch sind einzelne Rechenschritte weniger gut nachverfolgbar.

- **„Crosstab Queries“:** Zur Erstellung der Industriematrix auf Ebene von therapeutischen Gebieten und Indikationsgruppen wurden Abfragen verwendet, die Kreuztabellen erstellen. Dies ermöglicht die Darstellung der Vernetzung in einer Matrix.

Alle Analysen wurden in Microsoft Access so angelegt, dass sie manuell Schritt für Schritt einzeln berechnet werden können. Dies hat den Vorteil, dass maximale Flexibilität durch späteres Einführen von Zwischenschritten und Variationen möglich war, sowie die Rechenabläufe in Einzelschritten kontrollierbar sind.

Um den Ablauf von standardisierten Analysen (z.B. Vernetzungsmatrix von Indikationsgruppen, Identifikation von Signalwegranglisten etc.) in der Durchführung zu vereinfachen, wurden Makros verwendet. Die Makros definieren in Microsoft Access einen festgelegten Ablauf von Abfragen bzw. anderen Funktionen (z.B. Öffnen von Tabellen). Jedes Makro wurde durch genaue Beschreibung seiner Funktion und der dazugehörigen Outputtabelle definiert. Da die Verwendung von Makros mögliche Anwendungsfehler (z.B. Vergessen des Löschens einer Intermediärtabelle) vermeidet, wird die Verwendung der Makros für alle standardisierten Funktionen empfohlen.

6.4 Analysen

Im folgenden Abschnitt werden die in der Datenbank aufgebauten Analysen detailliert erklärt. Um die Erklärung verständlich zu illustrieren, wurde für jede Abfrage eine Abbildung analog zur Ansicht in Microsoft Access erstellt (vergleiche Abb. 6.7).

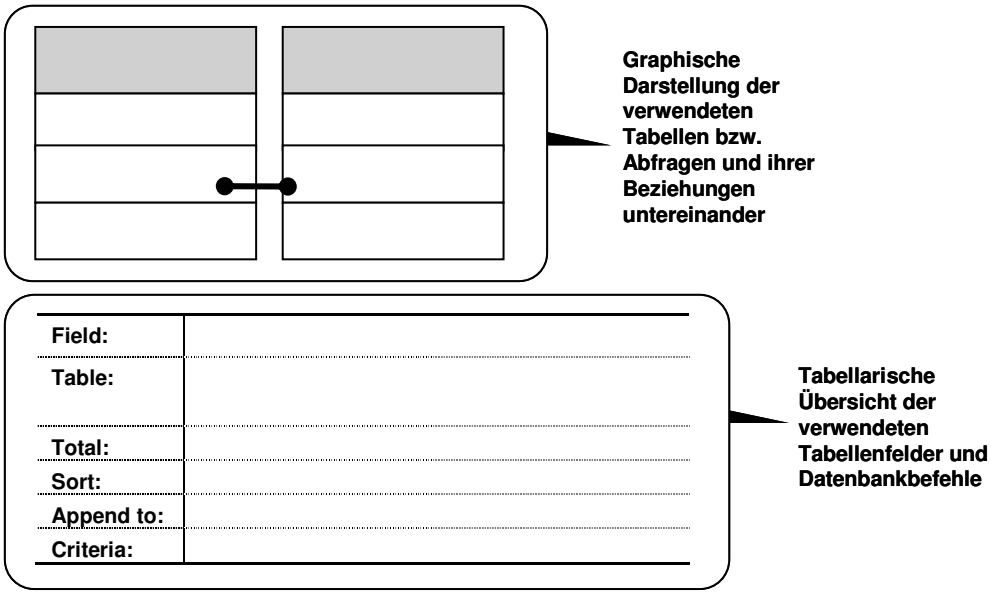


Abb. 6.7: **Alle Abbildungen im Methodenteil sind nach demselben Schema konzipiert** Diese Abbildung zeigt das Konzept aller folgenden Abbildungen von Microsoft Access Abfragen. Der obere Teil der Abbildung zeigt die verwendeten Tabellen sowie die Beziehungen der Tabellen untereinander (gekennzeichnet durch schwarze Verbindungslinie). Im unteren Teil können die verwendeten Tabellenfelder und mögliche Datebankbefehle tabellarisch gelistet werden. Die Struktur der oben dargestellten Tabelle ist typisch für eine „Append"-Abfrage. Die Darstellung ist an die Originalansicht in Microsoft Access angelehnt.

Die Abbildungen der Abfragen bestehen immer aus zwei Teilen:

- Graphische Darstellung der verwendeten Tabellen bzw. Abfragen sowie ihrer Beziehungen untereinander:
    - Jede Tabelle bzw. Abfrage, aus der Daten verwendet werden, wird einzeln abgebildet. Im grauen Rechteck findet sich der Tabellentitel; in den weißen Feldern stehen die Tabellenfelder. Fett gedruckte Tabellenfelder sind Primärschlüssel
    - Verbindungen zwischen gleichartigen Feldern sind durch schwarze Linien gekennzeichnet

- Tabellarische Übersicht der verwendeten Tabellenfelder und Datenbankbefehle. In Abb. 6.7 werden die Felder für eine „Append“-Abfrage gezeigt:
    - „Field“: Hier wird der Name des Felds, das in die Intermediärtabelle übernommen werden soll, eingetragen.
    - „Table“: Hier steht der Titel der Tabelle, aus der das Feld genommen werden soll.
    - „Total“: Hier können Rechenoperationen eingetragen werden. Häufig genutzte Operationen sind „Group by“ (Gruppieren von gleichartigen Zeilen), „Sum“ (Summieren), „Count“ (Zählen der Zeilen) oder „Avg“ (Durchschnittsbildung).
    - „Sort“: Hier kann eine Sortierung der Spalte (z.B. „Ascending“) eingetragen werden.
    - „Append to“: Hier steht der Name des Feldes der Intermediärtabellen, in das die Information geschrieben werden soll.
    - „Criteria“: Hier können mögliche Selektionskriterien eingetragen werden.

  Für andere Abfragetypen (z.B. „Make Table queries“ oder „Select queries“) tragen die Tabellenzeilen bei gleichem Funktionsprinzip etwas andere Namen. Die Erläuterung für diese Abfragetypen findet sich in der Bildunterschrift an der jeweiligen Stelle.

Alle Abfragen und Tabellen wurden in Microsoft Access analog zur Benennung der Inputtabellen mit einem Buchstaben-/Zahlencode sowie einem kurzen Namen betitelt. Abfragen beginnen stets mit einem Q für „Query“; Tabellen haben ein T für „Table“ als ersten Buchstaben ihres Codes. Die Darstellung der Abfragen lehnt sich an die Originalansicht in Microsoft Access an und verwendet daher englische Begriffe. Damit die Lesbarkeit des folgenden Kapitels nicht durch die hohe Anzahl von Abbildungen gestört wird, wird im Folgenden nur die erste

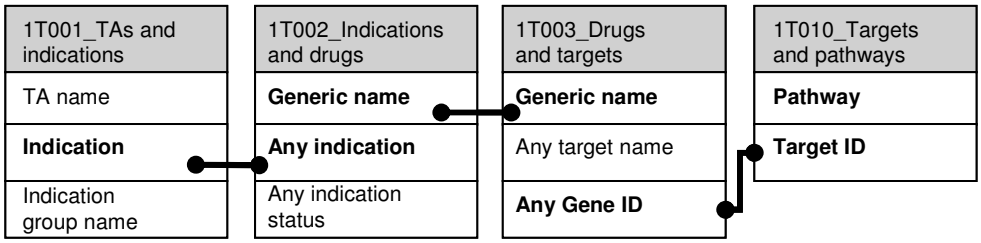
Abfrage detailliert dargestellt. Alle anderen Abbildungen können im Anhang der Arbeit nachgeschlagen werden.

6.4.1 Vernetzungsmatrix der therapeutischen Gebiete

Zur Analyse der Vernetzung der therapeutischen Gebiete durch gemeinsame Signalwege wurden alle verbindenden Signalwege für jede Verbindung von zwei therapeutischen Gebieten ermittelt. Ein Signalweg wurde dann als Verbindung gewertet, wenn es für beide therapeutische Gebiete mindestens eine aktive Substanz gibt, die ein Zielmolekül aus dem Signalweg adressiert.

Alle Analysen wurden in Microsoft Access als „Append"-Abfragen formuliert, die zuvor definierte Intermediärtabellen befüllen. Basierend auf diesen Intermediärtabellen wurde dann der jeweils folgende Abfrageschritt formuliert.

Erster Schritt der Analyse war die Zuordnung der Signalwege zu den therapeutischen Gebieten, in denen sie durch ihre Zielmoleküle von Bedeutung sind. Grundprinzip der Abfrage ist die Tatsache, dass die Beziehungen zwischen den einzelnen Tabellen eine Verbindung eines therapeutischen Gebiets („TA name") zu bestimmten Signalwegen („Pathway") ermöglichen. Der Verbindungsweg wird dabei über die Substanzen („Generic name") ermöglicht. Eine Substanz kann über die Indikationen, für die sie sich in Forschung und Entwicklung befindet, einem bestimmten therapeutischen Gebiet zugeordnet werden. Gleichermaßen kann die Substanz über das von ihr adressierte Zielmolekül auch einem Signalweg zugeordnet werden. Indirekt kann so die Verbindung zwischen dem therapeutischen Gebiet und dem Signalweg hergestellt werden. Abb. 6.8 zeigt die Abfrage der Zuordnung der Signalwege zu therapeutischen Gebieten. Die durch die Abfrage befüllte Intermediärtabelle listet pro therapeutisches Gebiet alle Signalwege, zu denen eine Verbindung hergestellt werden kann. Damit ein Signalweg für ein therapeutisches Gebiet gelistet wird, ist im Extremfall eine einzelne im Gebiet aktive Substanz mit einem Zielmolekül im entsprechenden Signalweg ausreichend.



Field:	TA name	Pathway
Table:	1T001_TAs and indications	1T010_Targets and pathways
Total:	Group by	Group by
Sort:		
Append to:	TA name	Pathway
Criteria:		

Abb. 6.8: **„Append"-Abfrage ordnet SIgnalwege zu therapeutischen Gebieten zu**
Diese Abfrage verwendet die vier Inputtabellen „1T001_TAs and indications", „1T002_Indications and drugs", „1T003_Drugs and targets" und „1T010_Targets and pathways" zur Zuordnung von therapeutischen Gebieten („TA name") und Signalwegen („Pathway"). Die Verbindung wird durch die Beziehungen gleichartiger Felder zwischen den Inputtabellen gewährleistet.
Die neue Intermediärtabelle enthält die Spalten „TA name" und „Pathway" und gruppiert die Informationen so, dass jede Kombination aus therapeutischem Gebiet und Signalweg nur einmal angezeigt wird.

Im nächsten Schritt wurden alle Signalwege gelistet, die für ein Paar von therapeutischen Gebieten relevant sind. Für diese Abfrage wurde zunächst ein kartesisches Produkt aller therapeutischen Gebiete mit den dazugehörigen Signalwegen erzeugt. Mit Hilfe von logischen Bedingungen, die in Access durch das Kürzel „Expr" gekennzeichnet sind, wurden aber nur diejenigen Ergebnisse ausgewählt, bei denen die therapeutischen Gebiete verschieden („Expr1"), der Signalweg jedoch übereinstimmend („Expr2") war. Der Aufbau der Abfrage mit den logischen Bedingungen ist Abb. 10.1 im Anhang zu dargestellt; die resultierende Intermediärtabelle enthält die erste Form der Matrix der therapeutischen Gebiete.

Im folgenden Schritt wurden die Anzahl der gemeinsamen Signalwege zwischen zwei therapeutischen Gebieten durch eine Zählfunktion ermittelt (siehe Abb. 10.2) und als Kreuztabelle dargestellt (siehe Abb. 10.3). Zur Dokumentation wurde noch eine weitere Abfrage angeschlossen, welche eine permanente Outputtabelle produziert („make-table-query", nicht gezeigt).

6.4.2 Vernetzungsgrad therapeutischer Gebiete

Wie im Abschnitt 6.4.1 dargestellt wurde, wurde mit Hilfe der Vernetzungsanalyse die Stärke der Verbindung zwischen zwei therapeutischen Gebieten beschrieben. Generell gilt, dass eine Aussage über die Anzahl der verbindenden Signalwege immer nur für die Verbindung von zwei bestimmten therapeutischen Gebieten gemacht werden kann. Als Hilfsmaß zur Auswertung wurde jedoch ein Parameter eingeführt, der die Vernetzung eines therapeutischen Gebiets insgesamt, d.h. mit allen 13 anderen therapeutischen Gebieten beschreibt. Dieser wurde Vernetzungsgrad genannt und wie folgt berechnet (Beispiel für ein therapeutisches Gebiet namens TA_1):

$$Vernetzungsgrad\ TA_1 = \sum_{n=2}^{14} (Anzahl\ Signalwegverbindungen\ TA_1 - TA_n) \qquad (6.1)$$

Die absolute Zahl des Vernetzungsgrads hat keine Relevanz. Der Vernetzungsgrad ist nur als relativer Wert hilfreich und beschreibt dann die tendenziellen Unterschiede in der Vernetzung verschiedener therapeutischer Gebiete. Der relative Vernetzungsgrad hat Werte von 0 bis 1 und entsteht aus dem Vernetzungsgrad durch Identifikation des maximalen Vernetzungsgrads im Vergleich aller therapeutischen Gebiete und Division durch dieses Maximum. Das therapeutische Gebiet, das insgesamt am stärksten vernetzt ist bekommt daher den Wert 1.

$$Relativer\ Vernetzungsgrad\ \ TA_1 = \frac{Vernetzungsgrad\ TA_1}{Maximaler\ Vernetzungsgrad\ aller\ TAs} \qquad (6.2)$$

Der relative Vernetzungsgrad wurde auch nach der Gewichtung (siehe Abschnitt 6.4.5) genutzt, um zu ermitteln, in wie weit die Gewichtung zur Differenzierung der Vernetzungsmatrix beigetragen hat. Zu diesem Zweck wurde der relative gewichtete Vernetzungsgrad analog zur obigen Formeln (6.1) und (6.2) aus den einzelnen gewichteten Verbindungswerten ermittelt und mit dem relativen Vernetzungsgrad vor Gewichtung verglichen.

6.4.3 Vernetzungsmatrix der Indikationsgruppen

Analog zur Vernetzungsanalyse von therapeutischen Gebieten wurde eine Vernetzungsanalyse auf Ebene von Indikationsgruppen durchgeführt. Endprodukt dieser Analysen war eine 77 x 77 Matrix, die die Anzahl der gemeinsamen Signalwege aller Indikationsgruppen untereinander darstellt.

In der ersten Abfrage für diese Analyse wurden alle für eine Indikationsgruppe relevanten Signalwege aufgelistet. Analog zu der Abfrage für therapeutische Gebiete nutzte die in Abb. 10.4 dargestellte Abfrage die Beziehungen zwischen den Inputtabellen „Indications and drugs“, „Drugs and targets“, „TAs and indications“ sowie „Targets and pathways“. Im Gegensatz zur Analyse für therapeutische Gebiete (vergleiche Abb. 6.8) wurde neben dem therapeutischen Gebiet zusätzlich die Unterteilung in Indikationsgruppen in die Abfrage übernommen.

In nächsten Schritt wurden – wie schon für therapeutische Gebiete beschrieben – alle Signalwege aufgelistet, die für ein Paar von Indikationsgruppen verbindend sind. Auch für die Indikationsgruppen wurde zunächst ein kartesisches Produkt aller Indikationsgruppen mit den entsprechenden Signalwegen erzeugt. Die logischen Bedingungen in der Abfrage für die Indikationsgruppen entsprechen konzeptionell denen der Abfrage für die therapeutischen Gebiete und stellen sicher, dass nur diejenigen Tupel dargestellt werden, bei denen die Indikationsgruppen verschieden sind, der Signalweg jedoch derselbe ist (siehe Abb. 10.5.).

In nächsten Analyseschritt (dargestellt in Abb. 10.6) wurde die Anzahl der gemeinsamen Signalwege für die Indikationsgruppenpaare ermittelt. Diesen Zweck erfüllte eine Zählfunktion, die die Anzahl der Signalwegeinträge für jedes Paar von Indikationsgruppen aus der Indikationsgruppenmatrix auslas. Zusätzlich zum Namen der Indikationsgruppen („Sub TA name“) wurden die Namen der therapeutischen Gebiete in die Tabelle übernommen („TA name“).

Nach Ermittlung der Anzahl der gemeinsamen Signalwege wurde auch auf Ebene der Indikationsgruppen eine Matrix mit Hilfe einer Kreuztabelle erstellt (siehe

Abb. 10.7). Zur Dokumentation wurde aus der Kreuztabelle durch eine weitere Abfrage eine Outputtabelle generiert („make-table-query“, nicht gezeigt).

6.4.4 Vernetzungsanalyse ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen

Im Gegensatz zur Vernetzungsanalyse aller Indikationsgruppen wurden bei der hier beschriebenen Analyse gezielt einzelne therapeutische Gebiete ausgewählt. Alle Ergebnisse, die durch die Auswahlanalyse gewonnen werden können, sind prinzipiell bereits in der Gesamtvernetzungsanalyse (vergleiche Abschnitt 6.4.3) für die entsprechenden therapeutischen Gebiete enthalten. Dennoch war es notwendig eine Abfragelogik für die ausgewählte Gebiete zu erstellen, da die Gesamtmatrix durch ihre enorme Größe sehr lange Rechenzeit benötigt und damit für die detaillierte Analyse einzelner therapeutischer Gebiete auf Ebene ihrer Indikationsgruppen ungeeignet ist.

Grundsätzlich wurden die Abfragen für die Auswahlanalyse analog der Abfragen der Gesamtanalyse (vergleiche Abschnitt 6.4.3) angelegt. Der erste Analyseschritt listet alle zu den Indikationsgruppen gehörenden Signalwege und ist übereinstimmend mit der in Abb. 10.4 dargestellten Abfrage. Wichtigster Unterschied gegenüber der Gesamtanalyse war das Einbeziehen eine „Parameter“-Abfrage in Microsoft Access im Rahmen der nächsten Abfrage zur Erstellung der Vernetzungsmatrix (siehe Abb. 10.8). Eine „Parameter“-Abfrage führt im Programmablauf von Access zum Öffnen von einem Eingabefenster, in das ein Auswahlkriterium (hier der Name eines therapeutischen Gebiets) eingegeben werden kann. Die „Parameter“-Abfrage aus Abb. 10.8 ermöglichte damit eine benutzerfreundliche Veränderung der für die Analyse ausgewählten therapeutischen Gebiete und berechnete das kartesische Produkt nur für die eingegebenen Gebiete.

Alle folgenden Schritte (Anzahl der gemeinsamen Signalwege, Kreuztabelle, Outputtabelle) verlaufen wie im Abschnitt 6.4.3 beschrieben. Damit sowohl die gesamte Matrix aller Indikationsgebiete als auch die Matrix ausgewählter therapeutischer Gebiete in der Datenbank enthalten bleiben können, wurden die

Abfragen bzw. Tabellen separat angelegt und in der Datenbank durch den Namenszusatz „Select“ gekennzeichnet.

6.4.5 Gewichtung der Vernetzung von therapeutischen Gebieten

Um unterschiedlich starke Verbindungen zwischen therapeutischen Gebieten differenzieren zu können, wurden verschiedene Ansätze zur Gewichtung der einzelnen Einträge in der Datenbank implementiert.

Grundsätzlich wird die Stärke einer Verbindung zwischen zwei therapeutischen Gebieten durch die Summe der Stärke der einzelnen Signalwegverbindungen bestimmt. Die Stärke einer einzelnen Signalwegverbindung zwischen zwei therapeutischen Gebieten ist wiederum von zwei Grundparametern abhängig:

- Anzahl der relevanten Projekte; d.h. Projekte in Entwicklung/auf dem Markt, die mit ihren Substanzen Zielmoleküle im entsprechenden Signalweg adressieren und deren Indikationen in die untersuchten therapeutischen Gebiete fallen
- Entwicklungsstatus der relevanten Projekte

Dabei gilt, dass jede Substanz, die eine Indikation in einem der beiden therapeutischen Gebiete und ein Zielmolekül aus dem entsprechenden Signalweg hat, als „Stütze“ der Signalwegverbindung berücksichtigt wird. Eine Substanz, die nicht nur Indikationen innerhalb eines der beiden therapeutischen Gebiete hat, sondern bereits heute indikationsübergreifend für beide therapeutischen Gebieten entwickelt wird, wird somit bei der Gewichtung auch doppelt berücksichtigt. Dies ist sinnvoll, um der besonderen Bedeutung von bereits bestehenden direkten Verbindungen zwischen zwei therapeutischen Gebieten durch indikationsübergreifende Zielmoleküle gerecht zu werden.

Alle Gewichtungsanalysen wurden grundsätzlich nach diesem Prinzip aufgebaut, jedoch wurden – je nach inhaltlicher Ausrichtung der Analyse – Variationen in die Berechnungslogik eingeführt. Die Abfragen der Gewichtung wurden in der Datenbank mit dem Namen „Weighting Base“ versehen. Je nach Variation der Berechnungslogik wurde der Name mit dem Kürzel „noav“ für „no average“ bei

der Gewichtung ohne Durchschnittsbildung bzw. „fullav“ für „full average“ bei der Gewichtung mit doppelter Durchschnittsbildung ergänzt.

6.4.5.1 Gewichtung ohne Durchschnittsbildung

Zunächst wurde eine Inputtabelle erstellt, die die unterschiedlichen Entwicklungsstatus der Forschungsprojekte mit einem Gewichtungsfaktor versieht. Tab. 6.1 zeigt eine mögliche Zuordnung dieses Gewichtungsfaktors („Status value“). Diese – auch als konservative Gewichtung bezeichnete – Zuordnung verleiht denjenigen Projekten mehr Gewicht, die in ihrer Entwicklung weiter fortgeschritten sind. Das führt bei Applikation des Gewichtungsfaktors dazu, dass Verbindungen, die durch bereits weit fortgeschrittene Projekte belegt sind, stärker erscheinen als Verbindungen mit Projekten in frühen Entwicklungsphasen. Je nach Analyseausrichtung wurde der Gewichtungsfaktor durch Verändern der Inputtabelle variiert.

Any indication status	Status value
Launched	3
Registered	3
Pre-registration	2
Phase III Clinical Trial	2
Phase II Clinical Trial	2
Phase I Clinical Trial	1
Clinical Trial (unspecified phase)	1
Preclinical	1
No Development Reported	0
Suspended	0
Unspecified	0

Tab. 6.1: Der Gewichtungsfaktor richtet sich nach dem Entwicklungsstaus einer Indikation
Jedem möglichen Entwicklungsstatus eines Forschungsprojekts („Any indication status“) wurde ein Gewichtungsfaktor („Status value“) zugeordnet. Diese Gewichtung wurde in Form einer Inputtabelle in der Datenbank gespeichert und kann daher jederzeit geändert werden. Im gezeigten Beispiel bekamen Projekte, die bereits auf dem Markt sind bzw. die Zulassung hinter sich haben, das Gewicht 3 („Launched“ und „Registered“). Projekten vor der Zulassung und in den klinischen Phasen II und III wurde das Gewicht 2 zugeordnet („Pre-registration“, „Phase III Clinical Trial“, „Phase II Clinical Trial“). Projekte in der Präklinik, in Phase I sowie ohne genaue Bezeichnung der klinischen Phase erhielten das Gewicht 1 („Preclinical“, „Phase I Clinical Trial“ und „Clinical Trial (unspecified phase)“). Projekte ohne Bericht über Entwicklung und beendete Projekte wurden durch Anwendung des Gewichts 0 von der weiteren Analyse ausgeschlossen („No Development Reported“, „Suspended“, „Unspecified“). Diese Verteilung des Gewichtungsfakors wird als konservative Gewichtung bezeichnet.

Die erste Abfrage zur Gewichtung („Weighting Base“, siehe Abb. 10.9) diente der Zuordnung der einzelnen Indikationen einer Substanz zu therapeutischen Gebieten, der Zuordnung der Substanzen zu Signalwegen sowie der Applikation

des Gewichtungsfaktors pro Forschungsprojekt. Zu diesem Zweck zählte die Abfrage die Anzahl der Indikationen einer Substanz pro therapeutischem Gebiet und Entwicklungsstatus (d.h. zum Beispiel für die Substanz X zwei Indikationen im kardiovaskulären Bereich, die bereits auf dem Markt sind). Durch Abgleich mit der Tabelle „Indication status types“ wurde den Indikationen entsprechend ihres Entwicklungsstatus ein Gewichtungsfaktor zugeordnet (z.B.: Indikationen auf dem Markt: Gewichtungsfaktor 3). Die Zuordnung der Substanzen zu Signalwegen erfolgte dabei durch das von ihnen adressierte Zielmolekül und die Verknüpfung der Rohdaten über die Inputtabelle „Targets and pathways“. Endprodukt dieser ersten Gewichtungsabfrage war eine Intermediärtabelle, die pro therapeutischem Gebiet und Signalweg alle Zielmoleküle und die von ihnen adressierten Substanzen listete, die Indikationen im entsprechenden therapeutischen Gebiet haben. In der Intermediärtabelle wurden darüber hinaus pro Substanz die Summe der Forschungsprojekte pro Entwicklungsstatus sowie der entsprechende Gewichtungsfaktor angezeigt. Die „Weighting Base“ Abfrage ist im Anhang abgebildet (Abb. 10.9).

Die zweite Abfrage zur Gewichtung („Weighting Base 1“, siehe Abb. 10.10) nutzte die Ergebnisse von „Weighting Base“ und berechnete pro Substanz und therapeutischem Gebiet das Gewicht für jeden Entwicklungsstatus (z.B. Gesamtgewicht sechs für Substanz X mit zwei Indikationen auf dem Markt, die je einen Gewichtungsfaktor von drei haben). Zu diesem Zweck wurde die Anzahl der Indikationen pro Entwicklungsstatus mit dem zugeordneten Gewichtungsfaktor multipliziert (siehe Abb. 10.10). Da die Substanzen einem Zielmolekül zugeordnet sind und das Zielmolekül Teil eines Signalwegs ist, wurde das Gewicht für jeden Signalweg spezifisch berechnet und aufgelistet. Ist ein Zielmolekül in mehr als einem Signalweg beteiligt, wurden die Gewichte für alle relevanten Signalwege aufgelistet.

Im dritten Gewichtungsschritt („Weighting Base 2“, Abb. 10.11) wurde pro Substanz und therapeutischem Gebiet ein Gesamtgewicht über alle Entwicklungsstufen ermittelt. Dieses setzte sich aus der Summe der Gewichte pro Entwicklungsstatus, die in der Abfrage „Weighting Base 1“ berechnet worden waren, zusammen. Um Variationen der Gewichtung, z.B. Durchschnittsbildung zu

erlauben, wurde in der gleichen Abfrage auch die Summe der Indikationen pro Substanz und therapeutischem Gebiet berechnet. Auch in dieser Abfrage wurden die Substanzen über ihre Zielmoleküle zu Signalwegen zugeordnet und die Berechnungsergebnisse für jeden relevanten Signalweg aufgelistet.

Im nächsten Gewichtungsschritt wurde ein Gesamtgewicht für jedes Zielmolekül pro therapeutischem Gebiet und Signalweg ermittelt. Zu diesem Zweck wurden die Gesamtgewichte der einzelnen Substanzen innerhalb eines therapeutischen Gebiets und Signalwegs für jedes Zielmolekül summiert („Weighting Base 2BC noav“, siehe Abb. 10.12). Dieses Gesamtgewicht entspricht der Summe der nach ihrem Entwicklungsstatus gewichteten Indikationen der Substanzen, die das Zielmolekül adressieren und in das therapeutische Gebiet gehören.

In den bisher beschriebenen Gewichtungsschritten wurde das Gewicht stets in Bezug auf ein therapeutisches Gebiet berechnet. Im folgenden Abfrageschritt wurde nun die Verbindung zweier therapeutischer Gebiete durch einen gemeinsamen Signalweg berücksichtigt (siehe „Weighting Base 3 noav“, Abb. 10.13).

Die Gewichtungsabfrage „Weighting Base 3 noav“ errechnete das Gewicht für einen verbindenden Signalweg als Summe aller gewichteten Zielmoleküle für beide therapeutischen Gebiete. Ein Zielmolekül zählte zur Gewichtung des verbindenden Signalwegs, wenn es durch Substanzen adressiert wird, die für Indikationen in einem oder in beiden der therapeutischen Gebiete erforscht werden (vergleiche „Expr3“ in Abb. 10.13). Dabei müssen die Zielmoleküle nicht zwangsweise direkt verbindend sein (d.h. für beide therapeutischen Gebiete Indikationen haben). Damit nur diejenigen Signalwege berücksichtigt werden, die in beiden therapeutischen Gebieten von Bedeutung sind, wurde die Inputtabelle der vorhergehenden Abfrage mit der Tabelle der Vernetzungsmatrix der therapeutischen Gebiete verknüpft. Durch diese Verbindung wurden bei der Abfrage nur diejenigen Signalwege berechnet, die bereits in der Vernetzungsanalyse als verbindend identifiziert worden waren.

Im letzten Gewichtungsschritt (siehe Abb. 10.14) wurden die Gewichte aller verbindenden Signalwege für die Verbindung von zwei therapeutischen Gebieten

zusammengerechnet. Somit entstand für die Verbindung von zwei therapeutischen Gebieten ein gewichteter Verbindungswert („Pathway Score“). Neben dem gewichteten Verbindungswert wurde in der Analyse ebenfalls zur Kontrolle die absolute Anzahl der verbindenden Signalwege bestimmt. Diese entspricht der Zahl, die bereits in Rahmen der Vernetzungsanalyse der therapeutischen Gebiete bestimmt wurde (siehe Abschnitt 6.4.1).

Zur vergleichenden Darstellung mit der ungewichteten Vernetzungsmatrix wurde eine Abfrage zur Erstellung einer Kreuztabelle der therapeutischen Gebiete erstellt. Der Aufbau der Abfrage für die Kreuztabelle sowie die Erstellung einer Outputtabelle mit Hilfe einer „Make-table query“ entsprach dem für die ungewichtete Vernetzungsmatrix (vergleiche 6.4.2) und wird daher nicht gezeigt.

6.4.5.2 Gewichtung mit einfacher Durchschnittsbildung

Als Variation der Gewichtung wurde die Durchschnittsbildung eingeführt. Im Fall der einfachen Durchschnittsbildung bedeutet dies, dass für jede Substanz innerhalb eines therapeutischen Gebiets ein durchschnittliches Gewicht ermittelt wurde. Dieses entstand durch Multiplikation des Gewichtungsfaktors mit dem durchschnittlichen Status der Indikationen der Substanz. Dies steht im Gegensatz zur Gewichtung ohne Durchschnittsbildung, bei der sich das Gewicht einer Substanz innerhalb eines therapeutischen Gebiets als Summe der mit dem Gewichtungsfaktor multiplizierten Einzelindikationen zusammensetzte. Bei der Gewichtung mit einfacher Durchschnittsbildung wurde also nicht mehr berücksichtigt, ob eine Substanz innerhalb eines therapeutischen Gebiets mehrere Indikationen hat oder nicht. Das Gewicht bestimmte sich lediglich durch den durchschnittlichen Entwicklungsstatus einer Substanz. Da die Durchschnittsbildung zunächst für jedes therapeutische Gebiet separat erfolgte, wurden direkt indikationsübergreifende Zielmoleküle, die bereits in beiden untersuchten therapeutischen Gebieten mit Indikationen aktiv sind, weiterhin zweifach berücksichtigt. Dies ist sinnvoll, um die besondere Bedeutung dieser direkten Verbindungen für den Signalweg zu berücksichtigen.

Hintergrund für die Einführung der Durchschnittsbildung ist das Phänomen, das Substanzen im Bereich Onkologie im Durchschnitt doppelt so viele Indikationen

haben, wie Substanzen, die in anderen therapeutischen Gebieten aktiv sind (vergleiche Abb. 7.16). Werden alle Indikationen einzeln berücksichtigt, wie das bei der Gewichtung ohne Durchschnittbildung der Fall ist, führt das zu einem erhöhten Gewicht aller Signalwege in der Onkologie.

Methodisch wurde die Gewichtung mit einfacher Durchschnittsbildung als Abweichung zur Gewichtung ohne Durchschnittsbildung in die Datenbank eingebaut. Zunächst wurden die ersten beiden Gewichtungsabfragen analog zur Gewichtung ohne Durchschnittsbildung berechnet („Weighting Base“ und „Weighting Base 1“). Darauf folgend wurde ein Zwischenschritt eingefügt, der das durchschnittliche Gewicht für jede Substanz pro Zielmolekül innerhalb eines therapeutischen Gebiets berechnete. Zu diesem Zweck wurde die Summe gewichteter Indikationen pro Substanz durch die absolute Anzahl der Indikationen pro Substanz dividiert (siehe Abb. 10.15).

Im nächsten Abfrageschritt wurden diese durchschnittlichen Gewichte der Substanzen dann – ähnlich wie bei der Gewichtung ohne Durchschnittsbildung die summierten Gewichte – für jedes Zielmolekül addiert („Weighting Base 2C“, siehe Abb. 10.16).

Das Endprodukt der Abfrage „Weighting Base 2C“ führte dann zu einer Intermediärtabelle, die strukturell der Intermediärtabelle des vierten Gewichtungsschritts der Gewichtung ohne Durchschnittsbildung entsprach (vergleiche „Weighting Base 2BC noav“, Abb. 10.12) und für jedes Zielmolekül pro Signalweg und therapeutischem Gebiet ein Gesamtgewicht enthielt.

Die anschließenden Gewichtungsschritte, die ein Gesamtgewicht zunächst für den Signalweg und dann für die Summe der Signalwege in der Verbindung zwischen den therapeutischen Gebieten bestimmten, wurden analog zur Gewichtung ohne Durchschnittsbildung definiert und sind daher nicht hier gezeigt. In Microsoft Access wurden sie als separate Abfragen abgelegt, damit jederzeit eine Auswahl zwischen den unterschiedlichen Gewichtungstypen zur Verfügung stand.

6.4.5.3 Gewichtung mit doppelter Durchschnittsbildung

Die Gewichtung mit doppelter Durchschnittsbildung wurde eingeführt, um einem weiteren Phänomen, der Varianz der Anzahl der pharmazeutisch aktiven Substanzen pro Zielmolekül, entgegenzuwirken. Pro therapeutisches Gebiet variiert die Anzahl der Substanzen, die es für ein und dasselbe Zielmolekül gibt. Dies ist unter anderem durch die Marktsituation (z.B. Umsatzvolumen, Wachstum) der therapeutischen Gebiete und durch die bisherige Forschungsintensität bedingt. Soll dieser Parameter bei der Gewichtung nicht mehr berücksichtigt werden, so muss sich das Gewicht pro Zielmolekül innerhalb eines therapeutischen Gebiets als Durchschnitt der Gewichte der Substanzen bilden und nicht wie bisher als deren Summe. Wie bei der Gewichtung mit einfacher Durchschnittsbildung wurden direkt indikationsübergreifende Zielmoleküle zweifach in der Gewichtung berücksichtigt.

Dies führte zu einer methodischen Veränderung der Gewichtungsabfragen. Zunächst wurden wie bei den anderen Gewichtungstypen die ersten beiden Abfragen („Weighting Base“ und „Weighting Base 1“) berechnet. Im nächsten Schritt wurde analog zur Gewichtung mit einfacher Durchschnittsbildung ein durchschnittliches Gewicht pro Substanz ermittelt (vergleiche Abb. 10.15, „Weighting Base 2B“). Diese durchschnittlichen Gewichte wurden im Folgeschritt aber nicht summiert, sondern wie in Abb. 10.17 gezeigt, zur Berechnung eines durchschnittlichen Gewichts pro Zielmolekül verwendet. Dies führte dazu, dass das Gewicht pro Zielmolekül lediglich im Bereich der durch die Gewichtungsfaktoren bestimmten Breite (z.B. zwischen 0 und 3) variierte.

Die anschließenden Gewichtungsschritte erfolgten analog zur Gewichtung ohne Durchschnittsbildung und wurden in Access als separate Abfragen abgelegt, um einen Wechsel der Gewichtung jederzeit mit minimalem Aufwand zu ermöglichen.

6.4.5.4 Gewichtung mit doppelter Durchschnittsbildung und Multiplikator

Wie bereits oben beschrieben, wurden im Rahmen der Gewichtung Zielmoleküle, die direkt indikationsübergreifend sind (d.h. bereits heute für Indikationen in beiden therapeutischen Gebiete erforscht werden), im Vergleich zu denen, die

bisher nur in einem der beiden Gebiete vorkommen und damit nur indirekt über den Signalweg verbindend wirken, zweifach berücksichtigt. Diese zweifache Berücksichtigung in der Gewichtung wurde durchgeführt, da diese direkten Molekülverbindungen eine Signalwegverbindung zwischen zwei Gebieten besonders stark stützen.

Die Gewichtung mit doppelter Durchschnittsbildung und Multiplikator ging noch einen Schritt weiter und führte einen Multiplikator ein, mit dem alle diejenigen Projekte multipliziert wurden, deren Zielmoleküle direkt indikationsübergreifend sind. Das bedeutet, dass sich das Gewicht eines Zielmoleküls für eine Signalwegverbindung nun nicht mehr direkt aus der Summe der durchschnittlichen Gewichte der zwei therapeutischen Gebiete zusammensetzte. Stattdessen wurde diese Summe noch zusätzlich mit einem Multiplikator verrechnet. Die Anwendung des Multiplikators führte zu dem Effekt, dass Verbindungen, die durch zahlreiche direkt indikationsübergreifende Zielmoleküle gestützt sind, stärker gewichtet werden als diejenigen, die zwar eine hohe Anzahl an Zielmolekülen haben, allerdings über wenig direkte Verbindungen verfügen. Eine solche Analyse ist dann von Bedeutung, wenn bei der Gewichtung die Erprobtheit einer Verbindung besonders betont werden soll.

Die Umsetzung dieser Gewichtung in der Datenbank baute auf der oben beschriebenen Gewichtung mit doppelter Durchschnittsbildung auf. Die in Abb. 10.17 dargestellte „Select“-Abfrage berechnete einen durchschnittlichen Status für ein Zielmolekül pro therapeutisches Gebiet. Aufbauend auf dem Ergebnis dieser Abfrage wurde mit einer weiteren Abfrage eine Liste aller Gewichte der Zielmoleküle für jedes Paar an therapeutischen Gebieten pro Signalweg erstellt (siehe Abb. 10.18). Das bedeutet, dass zum Beispiel für das Paar der therapeutischen Gebiete „CNS“ und „Oncology“ pro Signalweg alle diejenigen Zielmoleküle gelistet wurden, die entweder für „CNS“ oder „Oncology“ relevant sind (siehe logische Bedingung „Expr3“ in Abb. 10.18). Ein Zielmolekül, das sowohl für „CNS“ als auch für „Oncology“ wichtig ist, wurde daher insgesamt in zwei Reihen aufgelistet. Die Abfrage bezog dabei die Vernetzungsmatrix der therapeutischen Gebiete als Inputtabelle mit ein und stellte so sicher, dass nur verbindende Signalwege aufgelistet wurden.

Diese Liste wurde nun dazu gebraucht, diejenigen Zielmoleküle innerhalb eines Signalwegs zu identifizieren, die für ein Paar von therapeutischen Gebieten direkt indikationsübergreifend sind. Zu diesem Zweck wurde in der nächsten Abfrage die Anzahl der Zeilen gezählt, in denen ein Zielmolekül mit derselben Kombination an therapeutischen Gebieten und Signalweg aufgelistet worden war. Dies geschah durch eine einfache Zählfunktion der Zeilen aus dem Ergebnis der vorhergehenden Abfrage (siehe Abb. 10.19). Durch diese Abfrage erhielten alle Zielmoleküle, die für eine bestimmte Kombination an therapeutischen Gebieten und Signalweg direkt indikationsübergreifend sind, den Wert zwei und alle anderen den Wert eins.

Im nächsten Schritt wurden nun die Gewichte der direkt indikationsübergreifenden Zielmoleküle in zwei Schritten mit dem Multiplikator multipliziert. Um eine flexible Veränderung des Multiplikators zu ermöglichen, wurde der Wert in einer separaten Inputtabelle gespeichert (siehe Abb. 6.9).

Direct connection	Muliplicator
1	1
2	2

Abb. 6.9: Inputtabelle für den Wert des Multiplikators
Die Abbildung zeigt die Inputtabelle, in der der Wert des Multiplikators festgelegt wurde. Je nachdem, ob ein Zielmolekül bereits für beide therapeutischen Gebieten indikationsübergreifend ist („Direct connection“ = 2) oder nur in einem der beiden Gebiete von Bedeutung ist („Direct connection“ = 1), wurde der Multiplikator („Multiplicator“) zugeordnet. Im dargestellten Beispiel war der Multiplikator für direkte Verbindungen auf 2 eingestellt. Durch einfaches Ändern des Werts in der Multiplikator-Spalte konnte diese Einstellung verändert werden.

Der erste Abfrageschritt für die Anwendung des Multiplikators ordnete zunächst jedem Zielmolekül den entsprechenden Multiplikator zu (siehe Abb. 10.20).

Die eigentlich Multiplikation des durchschnittlichen Zielmolekülgewichts mit dem Multiplikator erfolgte in der nächsten Abfrage durch eine algebraische Operation (siehe Abb. 10.21). Ergebnis war ein durchschnittliches Gewicht pro Kombination an therapeutischen Gebieten und Signalweg für jedes Zielmolekül unter Berücksichtigung des Multiplikators. Das bedeutet, wurde ein Multiplikator von beispielsweise zwei verwendet, wurden alle Gewichte der Zielmoleküle, die direkt indikationsübergreifend sind, mit zwei multipliziert.

Im nächsten Schritt (vergleiche Abfrage aus Abb. 10.22) wurde nun die Gesamtstärke einer Signalwegverbindung zwischen zwei therapeutischen

Gebieten berechnet. Dies geschah durch Summieren der Gewichte der einzelnen Zielmoleküle und entspricht nun grundsätzlich dem, was auch in den anderen Gewichtungsarten berechnet wurde. Endprodukt der Abfrage war ein Wert für die Verbindungsstärke eines jeden Signalwegs, der für die Verbindung von zwei therapeutischen Gebieten relevant ist. Dies entsprach dem Endprodukt der Abfragen des vierten Gewichtungsschritts der anderen Gewichtungstypen (vergleiche „Weighting Base 2BC noav“, Abb. 10.12). Alle nachfolgenden Schritte wurden analog der Gewichtung ohne Durchschnittsbildung durchgeführt und werden daher hier nicht gezeigt. Die Abfragen wurden in Microsoft Access separat angelegt, um jederzeit einen Wechsel zwischen den Gewichtungsarten zu ermöglichen.

6.4.5.5 Normierte Vernetzungsmatrix

Je nach Verwendung unterschiedlicher Gewichtungsarten (z.B. mit und ohne Durchschnitt) variieren die absoluten Werte, die für eine Verbindung zwischen zwei therapeutischen Gebieten gewonnen wurden. Da die absoluten Werte aufgrund ihrer Größe zur Auswertung nicht gut geeignet und lediglich Unterschiede in der Vernetzung interpretierbar sind, wurden die Verbindungswerte zur Darstellung der Vernetzungsmatrix auf Werte von 0 bis 100 normiert. Um dies zu erreichen, wurde der maximale Wert der Matrix ermittelt und alle anderen Werte wurden durch diesen dividiert (vergleiche (6.3).

$$Normierter\ Verbindungswert = \frac{Absoluter\ Verbindungswert}{Maximaler\ Verindungswert\ pro\ Matrix} \qquad (6.3)$$

6.4.6 Gewichtung der Vernetzung von Indikationsgruppen

Genauso wie die Stärke der Verbindungen zwischen einzelnen therapeutischen Gebieten gewichtet wurde, wurde auch eine Gewichtung auf Ebene der Indikationsgruppen vorgenommen. Auch für diese Gewichtung galt, dass die Stärke der Gesamtverbindung der Summe der Stärke der einzelnen Signalwegverbindungen entsprach. Wie stark die Verbindung durch einen einzelnen Signalweg ist, hing wiederum davon ab, wie viele Substanzen es gibt, die Zielmoleküle innerhalb des Signalwegs haben, die für die Indikationsgruppe

relevant sind und welchen Entwicklungsstatus diese Substanzen haben (z.B. bereits auf dem Markt vs. in präklinischer Forschung).

Grundsätzlich ist eine Verbindung immer dann stark, wenn sie durch viele Substanzen mit weit fortgeschrittenen Entwicklungsstatus gestützt wird. Auch bei der Gewichtung auf Ebene der Indikationsgruppen galt, dass Zielmoleküle, die bereits direkt verbindend sind, da sie für beide Indikationsgruppen relevant sind, zweifach in der Gewichtung berücksichtigt wurden.

Ja nachdem welche Gewichtungsform gewählt wurde, variierte die Verrechnung der Gewichtungsparameter.

6.4.6.1 Gewichtung ohne Durchschnittsbildung

Für die Gewichtung ohne Durchschnittsbildung wurden die Indikationen der einzelnen Substanzen im ersten Schritt entsprechend ihres Entwicklungsstatus mit einem Gewichtsfaktor versehen. Zu diesem Zweck wurde derselbe Gewichtungsfaktor wir für die Gewichtung der therapeutischen Gebiete verwendet (siehe Tab. 6.1). Auch für die Gewichtung der Indikationsgruppen konnte das Gewicht durch Verändern der Tabelle variiert werden.

Grundsätzlich wurden die Gewichtungsabfragen analog zur Gewichtung der therapeutischen Gebiete angelegt. In den ersten Abfragen war der Hauptunterschied zur Gewichtung der therapeutischen Gebiete die Einführung der zusätzlichen Detailinformation der Indikationsgruppe. In der Datenbank wurde für den Begriff der Indikationsgruppe das kürzere Synonym „SubTA“ verwendet. In der ersten Gewichtungsabfrage für Indikationsgruppen („SubTA Weighting Base 1“, siehe Abb. 10.23) wurden die einzelnen Indikationen einer Substanz zu einer Indikationsgruppe und einem therapeutischen Gebiet zugeordnet. Pro Substanz wurde die Anzahl aller Indikationen innerhalb einer Indikationsgruppe ermittelt, die denselben Entwicklungsstatus haben. Den einzelnen Indikationen wurde entsprechend ihres Entwicklungsstatus durch Abgleich mit der Tabelle „Indication status types“ ein Gewichtungsfaktor zugeordnet.

Im nächsten Schritt („SubTA Weighting Base 1“, Abb. 10.24) wurden die Ergebnisse der ersten Abfrage genutzt und pro Substanz und Indikationsgruppe

ein Gesamtgewicht für jeden Entwicklungsstatus berechnet. Analog zur Gewichtung der Vernetzung der therapeutischen Gebiete wurde dazu die Anzahl der Indikationen einer Substanz im gleichen Entwicklungsstatus mit dem entsprechenden Gewichtungsfaktor multipliziert. Ist ein Zielmolekül einer Substanz Teil mehrerer Signalwege, so wurden die Gesamtgewichte pro Entwicklungsstatus für jeden relevanten Signalweg aufgelistet.

Im nächsten Gewichtungsschritt („SubTA Weighting Base 2“, Abb. 10.25) wurden alle gewichteten Indikationen einer Substanz zu einem Gesamtgewicht zusammengeführt. Zu diesem Zweck wurden alle Gewichte pro Entwicklungsstatus für eine Substanz summiert. Um weitere Formen der Gewichtung (z.B. mit Durchschnittsbildung) zu erlauben, wurde darüber hinaus die Summe aller Indikationen pro Substanz ermittelt. Dieser Parameter war jedoch für die Gewichtung ohne Durchschnittsbildung nicht relevant. Genauso wie im vorherigen Gewichtungsschritt erfolgte die Berechnung jeweils spezifisch für einen Signalweg, d.h. Zielmoleküle, die zu mehr als einem Signalweg gehören, wurden mit ihren Substanzen für jeden relevanten Signalweg einzeln aufgelistet.

Während alle bisherigen Gewichtungsschritte stets die Gewichtung auf Ebene der Substanzen berechneten, wurde im nächsten Schritt nun ein Gesamtgewicht pro Zielmolekül für jede Indikationsgruppe und jeden Signalweg ermittelt („SubTA Weighting Base 2BC noav“, Abb. 10.26). Zu diesem Zweck wurden die Gewichte der einzelnen Substanzen pro Zielmolekül innerhalb einer Indikationsgruppe summiert. Das Gesamtgewicht des Zielmoleküls entsprach daher der Summe der einzelnen gewichteten Substanzen. Wie bei den bisherigen Berechnungen erfolgte die Berechnung jeweils spezifisch für jeden Signalweg.

Der nachfolgende Gewichtungsschritt (Abb. 10.27, „SubTA Weighting Base 3 noav“) stellte nun – analog zur Gewichtung der Verbindungen zwischen therapeutischen Gebieten – jeweils zwei Indikationsgruppen in Relation zueinander und berechnete ein Gesamtgewicht für einen verbindenden Signalweg. Das Gewicht eines zwei Indikationsgruppen verbindenden Signalwegs berechnete sich als Summe aller gewichteten Forschungsprojekte, die Zielmoleküle in dem Signalweg adressieren und zu den beiden Indikationsgruppen gehören. Ein

Forschungsprojekt wird dabei als präklinische, klinische oder auch postklinische Aktivität einer Substanz für eine bestimmte Indikation verstanden. In der Abfrage „SubTA Weighting Base 3 noav“ wurde ein Forschungsprojekt dann für die Berechnung des Signalweggewichts verwendet, wenn die Indikation des Projekts einer der beiden durch den Signalweg verbundenen Indikationsgruppen zugeordnet werden konnte (siehe „Expr3“ in Abb. 10.27). Die Gewichtungsabfrage erstellte eine Beziehung zwischen der Tabelle „SubTA Matrix“ und der Intermediärtabelle der vorhergehenden Abfrage und stellte dadurch sicher, dass nur diejenigen Signalwege berücksichtigt wurden, die eine Verbindung zwischen den Indikationsgruppen darstellen. Um die Redundanzen in der von der Abfrage befüllten Intermediärtabelle zu verhindern, wurde mit Hilfe einer logischen Bedingung (siehe „Expr4“, Abb. 10.27) ausgeschlossen, dass Signalwege, die in der Tabelle „SubTA Matrix“ aufgrund ihrer Matrixsstruktur doppelt eingetragen sind, auch in Abb. 10.27 doppelt angezeigt werden.

Im anschließenden Gewichtungsschritt („SubTA Weighting Base 3B noav“, Abb. 10.28) wurden die einzelnen Gewichte der verbindenden Signalwege zwischen zwei Indikationsgruppen zu einem Gesamtwert der Verbindung addiert. Die Stärke der Verbindung zweier Indikationsgruppen miteinander entsprach somit der Summe der gewichteten einzelnen Signalwegverbindungen zwischen ihnen. Neben diesem Gesamtgewicht („Pathway Score“) wurde in der Analyse zur Kontrolle erneut die Anzahl der verbindenden Signalwege bestimmt („Pathway Count“). Diese musste mit der Anzahl der verbindenden Signalwege, die bereits in der Vernetzungsanalyse der Indikationsgruppen bestimmt worden war, übereinstimmen (vergleiche Abschnitt 6.4.3).

Um die gewichtete Vernetzungsmatrix mit der ungewichteten Vernetzungsmatrix vergleichen zu können, wurde ebenfalls eine Kreuztabelle erstellt. Basierend auf dieser Kreuztabelle wurde mit Hilfe einer „Make-Table“-Abfrage eine dauerhafte Outputtabelle. Die hierzu verwendeten Abfragen waren analog der Abfrage für die ungewichtete Matrix und werden hier nicht gezeigt.

6.4.6.2 Gewichtung mit einfacher Durchschnittsbildung

Auch für die Gewichtung der Verbindungen zwischen Indikationsgruppen wurde die Durchschnittsbildung eingeführt. Für den Fall der Indikationsgruppen bedeutete dies, dass für jede Substanz innerhalb einer Indikationsgruppe ein Gewicht ermittelt wurde, das im Gegensatz zur Gewichtung ohne Durchschnittsbildung nicht der Summe der gewichteten Indikationen einer Substanz entsprach, sondern dem gewichteten durchschnittlichen Entwicklungsstatus aller Indikationen einer Substanz. Mathematisch entstand das Gewicht durch Multiplikation des Gewichtungsfaktors mit dem durchschnittlichen Entwicklungsstatus aller Indikationen der Substanz. Im Gegensatz zur Gewichtung ohne Durchschnittsbildung war bei der Gewichtung mit einfacher Durchschnittbildung nicht mehr relevant, wie viele Indikationen eine Substanz pro Indikationsgruppe hat sondern lediglich wie weit fortgeschritten die Entwicklung der Substanz im Mittel über ihre Indikationen ist. Vorteil dieser Gewichtungsform ist, dass Substanzen unabhängig von der Zahl der Indikationen innerhalb einer Indikationsgruppe, gleichermaßen zur Verbindungsstärke eines Signalwegs beitragen können. Dies spielt besonders im Bereich Onkologie eine Rolle. Substanzen, die bei onkologischen Indikationen aktiv sind, haben im Durchschnitt doppelt so viele Indikationen pro Indikationsgruppen wie Substanzen aus anderen Bereichen. Gewichtung mit Durchschnittsbildung sorgte dafür, dass diese Besonderheit des therapeutischen Gebiets nicht zu einer „Übergewichtung“ aller Signalwegverbindungen mit der Onkologie führte.

In der Datenbank wurden die Gewichtungsvariationen durch separate Abfragen realisiert. Die ersten beiden Gewichtungsabfragen („SubTA Weighting Base“, „SubTA Weighting Base 1“ und „SubTA Weighting Base 2“) konnten dabei auf die Abfrage der Gewichtung ohne Durchschnittsbildung zurückgreifen. Im Anschluss wurde ein zusätzlicher Zwischenschritt zur Berechnung des durchschnittlichen Gewichts pro Substanz und Zielmolekül innerhalb einer Indikationsgruppe eingefügt („SubTA Weighting Base 2B“, siehe Abb. 10.29). Die Berechnung erfolgte durch Division der Summe der gewichteten Indikationen pro Substanz durch die Anzahl der Indikationen pro Substanz.

Im nächsten Analyseschritt wurde das Gewicht pro Zielmolekül berechnet (siehe Abb. 10.30). Zu diesem Zweck wurden die Durchschnitts-Gewichte der einzelnen Substanzen, die in der vorhergehenden Abfrage berechnet worden waren, pro Zielmolekül summiert. Dieser Rechenschritt führte zu einem Gesamtgewicht pro Zielmolekül, Indikationsgebiet und Signalweg analog der Gewichtungsabfrage „SubTA Weighting Base 2BC noav“. Basierend auf diesem Ergebnis konnten die letzten Schritte der Abfragen („Weighting Base 3“ und „Weighting Base 3B“) identisch zur Gewichtung ohne Durchschnittsbildung aufgebaut werden. Ebenso wie bei der Gewichtung der Verbindung therapeutischer Gebiete wurden die Abfragen in Microsoft Access separat angelegt, um jederzeit einen Wechsel zwischen den Gewichtungstypen zu erlauben.

6.4.6.3 Gewichtung mit doppelter Durchschnittsbildung

Auch für Verbindungen zwischen Indikationsgruppen wurde die Gewichtung mit doppelter Durchschnittbildung eingeführt, um der Varianz der Anzahl der Substanzen pro Zielmolekül entgegenzuwirken. Im Gegensatz zur Gewichtung ohne Durchschnittsbildung bzw. mit einfacher Durchschnittsbildung ergab sich bei der Gewichtung mit doppelter Durchschnittsbildung das Gewicht pro Zielmolekül innerhalb einer Indikationsgruppe nicht als Summe der Gewichte der einzelnen Substanzen, sondern als deren Durchschnitt. Das Gewicht pro Zielmolekül war damit also weder von der Anzahl der Indikationen pro Substanz noch von der Anzahl der Substanzen pro Zielmolekül innerhalb der Indikationsgruppe abhängig. Welches Gewicht ein Zielmolekül bekam, war daher nur durch das Vorkommen von Substanzen (ja/nein) und deren durchschnittlichen Entwicklungsstatus bestimmt (z.B. präklinisch vs. bereits auf dem Markt).

Für die ersten Gewichtungsschritte konnte auch bei doppelter Durchschnittsbildung auf dieselben Abfragen wie bei der Gewichtung ohne Durchschnitt („SubTA Weighting Base“, „SubTA Weighting Base 1“ und „SubTA Weighting Base 2“) zurückgegriffen werden. Im nächsten Schritt wurde analog zur Gewichtung mit einfacher Durchschnittsbildung ein Durchschnittsgewicht pro Substanz berechnet (siehe Abb. 10.30, „SubTA Weighting Base 2B“). Aus diesen Gewichten pro Substanz wurde im Folgeschritt

der zweite („doppelte“) Durchschnitt gebildet (siehe Abb. 10.31). Diese doppelte Durchschnittsbildung führte dazu, dass die Gewichte pro Zielmolekül und Indikationsgruppe lediglich um die Spannbreite der anfänglich festgesetzten Gewichtungsfaktoren (im Beispiel aus Tab. 6.1 zwischen 0 und 3) variierten. Die anschließenden Gewichtungsschritte wurden analog zur Gewichtung ohne Durchschnittsbildung in Microsoft Access als separate Abfragen angelegt, um einen Wechsel zwischen den Gewichtungsarten zu vereinfachen.

6.4.6.4 Gewichtung mit doppelter Durchschnittsbildung und Multiplikator

Analog zur Gewichtung der Verbindungen der therapeutischen Gebiete wurde auch bei der Gewichtung der Indikationsgruppenverbindungen die Möglichkeit geschaffen, diejenigen Signalwegverbindungen besonders stark zu gewichten, die über viele direkt indikationsübergreifende Zielmoleküle verfügen.

Einziger Unterschied zur Gewichtung bei therapeutischen Gebieten war, dass eine Beschränkung auf bestimmte Gebiete vor dem Berechnen der Analyse vorgenommen wurde. Das bedeutet, dass bei der Analyse – im Gegensatz zu den oben beschriebenen Gewichtungsarten – nicht alle therapeutischen Gebiete mit ihren Indikationsgruppen untersucht wurden, sondern nur eine bestimmte Auswahl. Diese Einschränkung war bei der Gewichtung der Verbindung von Indikationsgruppen mit Multiplikator notwendig, um die Rechenzeiten von Microsoft Access in einem benutzerfreundlichen Umfang zu halten.

Die Gewichtung mit doppelten Durchschnitt und Multiplikator baut auf den ersten Gewichtungsschritten der oben beschriebenen doppelten Durchschnittsbildung auf. Basierend auf der Abfrage aus Abb. 10.20 für therapeutische Gebiete wurden aber zusätzliche Zwischenschritte eingeführt. Der erste Zwischenschritt (siehe Abb. 10.32) erstellte eine Liste derjenigen Zielmoleküle, die für eine spezifische Kombination aus Indikationsgruppen („SubTAs“) und Signalweg („Pathway“) relevant sind. Für jedes Zielmolekül wurde der errechnete durchschnittliche, gewichtete Entwicklungsstatus (vergleiche doppelte Durchschnittsbildung, Abb. 10.31) aufgelistet. Durch Verknüpfung der Intermediärtabelle der vorausgehenden Analyse (T222L) mit der Vernetzungsmatrix der Indikationsgruppen wurde sichergestellt, dass nur diejenigen Signalwege aufgelistet wurden, die tatsächlich

für die Verbindung der Indikationsgruppen eine Rolle spielen. Zwei logische Bedingungen sorgten dafür, dass nur die Zielmoleküle, die eine Indikation in einer oder beiden Indikationsgruppen haben, zur Berechnung beitrugen (siehe „Expr3“ in Abb. 10.32) sowie gleiche Zeilen nicht mehrfach angezeigt wurden (siehe „Expr4“ in Abb. 10.32). Um die Begrenzung der Rechenzeit zu erreichen, wurden für die Felder „TA name1“ und „TA name2“ jeweils eine „Parameter“-Funktion verwendet. Diese „Parameter“-Funktion führte jeweils dazu, dass sich in Microsoft Access ein Fenster öffnet, welches die Eingabe des Namens des therapeutischen Gebiets verlangt. Auf diese Weise wurde die Analyse auf die eingegebenen therapeutischen Gebiete beschränkt.

Im nächsten Schritt (siehe Abb. 10.33) wurde überprüft, welche Zielmoleküle direkt indikationsübergreifend für die Verbindung zwischen zwei Indikationsgruppen sind und welche nur indirekt über den Signalweg verbinden. Zu diesem Zweck wurde eine Zählfunktion auf die Ergebnisse der vorherigen Abfrage angewendet. Diejenigen Moleküle, die direkt indikationsübergreifend sind, waren in der vorhergehenden Abfrage (Abb. 10.32) in zwei Zeilen gelistet worden und erhielten durch die Zählfunktion den Wert 2.

Im nächsten Schritt (siehe Abb. 10.34) wurde den Zielmolekülen entsprechend der des Ergebnisses der Zählfunktion der Multiplikator (vergleiche Abb. 6.9) zugeordnet. Mit diesem Wert wurde der durchschnittliche Status des Zielmoleküls im Folgenden multipliziert.

Abb. 10.35 zeigt den nächsten Abfrageschritt, der die eigentliche Multiplikation mit dem Multiplikator durchführte. Endprodukt der Abfrage war der multiplizierte durchschnittliche Status jedes Zielmoleküls für die Kombination von zwei Indikationsgruppen und einem Signalweg. Ist ein Zielmolekül für beide Indikationsgruppen von Bedeutung, so wurde es zweimal aufgelistet.

Im nächsten Schritt (siehe Abb. 10.36) wurden die Gewichte der einzelnen Zielmoleküle für einen Signalweg summiert. Im Ergebnis wurde so jeder Signalwegverbindung zwischen Indikationsgruppen ein Gesamtgewicht zugeordnet.

Die darauf folgenden Schritte entsprachen den Abfragen der Gewichtung ohne Durchschnittsbildung und wurden nur separat in der Datenbank angelegt, um jederzeit einen Wechsel zwischen den Gewichtungsarten zu erlauben.

6.4.6.5 Normierte Vernetzungsmatrix

Genauso wie bei der gewichteten Vernetzungsmatrix für therapeutische Gebiete wurden die Vernetzungsmatrizen auf Indikationsebene in normierter Weise dargestellt. Dazu wurde der Maximalwert pro Matrix ermittelt und diesem der Wert 100 zugeordnet. Alle anderen gewichteten Verbindungswerte wurden durch den Maximalwert geteilt, so dass die normierte Vernetzungsmatrix Werte zwischen 0 und 100 enthielt.

6.4.7 Gewichtung der Vernetzung ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen

Wie bereits in Abschnitt 6.4.4 beschrieben, war es sinnvoll, neben der Analyse der gesamten gewichteten und ungewichteten Vernetzungsmatrix von Indikationsgruppen ebenfalls die Analyse von ausgewählten therapeutischen Gebieten auf Ebene der Indikationsgruppen zu ermöglichen. Die Auswahl und damit Beschränkung auf einzelne therapeutische Gebiete ermöglichte eine übersichtlichere Analyse mit reduzierten Rechenzeiten. Neben der Berechnung der Vernetzungsmatrix ohne Gewichtung, wie in Abschnitt 6.4.4 beschrieben, wurde daher auch die Gewichtungsmethodik für ausgewählte therapeutische Gebiete eingerichtet.

Grundsätzlich gilt, dass die Gewichtung der Verbindungen zwischen ausgewählten Gebieten mit Hilfe der gleichen Abfragelogik durchgeführt werden konnte, die für die Gesamtmatrix beschrieben wurde. Methodisch bedeutete dies jedoch, dass alle Abfragen für die Gewichtung einzelner Gebiete unabhängig von denen der Gesamtmatrix in der Datenbank angelegt werden mussten, um die Ergebnisse der vollständigen Analyse zu erhalten. Die Abfragen zur Gewichtung ausgewählter Gebiete wurden in der Datenbank daher separat angelegt und mit dem Namenszusatz „Select“ versehen. Da bei der Gewichtung mit doppelter Durchschnittsbildung und Multiplikator aus Gründen der Rechenzeit ohnehin auf

eine vollständige, alle Gebiete einschließende Analyse, verzichtet werden musste, ist einzig für diesen Gewichtungstyp keine „Select“ Analyse in der Datenbank zu finden.

Zur Gewichtung ausgewählter Gebiete wurden zunächst die gleichen ersten Schritte genutzt, die auch für die Gewichtung der Gesamtmatrix verwendet worden waren. Im Gegensatz zur uneingeschränkten Gewichtung wurde jedoch für die Abfrage zum Gewichtungsschritt „SubTA Weighting Base 3“ nicht die Gesamtmatrix aller Indikationsgruppen verwendet, sondern diejenige Matrix, bei der bereits die Auswahl der therapeutischen Gebiete durchgeführt worden war (Tabelle „SubTA Matrix Select“, nicht gezeigt). Die eigentlichen Gewichtungsabfragen unterscheiden sich daher lediglich in der Wahl der Inputtabelle sowie in der Nomenklatur der Tabellen und Felder von den Gewichtungsabfragen ohne Beschränkung auf therapeutische Gebiete und werden daher hier nicht gezeigt.

Genauso wie für die Gesamtmatrix, wurden auch für die Gewichtung ausgewählter therapeutischer Gebiete sowohl die Gewichtung ohne also auch mit einfacher oder doppelter Durchschnittsbildung angelegt. Endprodukt aller Gewichtungsschritte war eine Kreuztabelle, die die Matrix der ausgewählten Gebiete zeigte und durch eine separate Abfrage in eine permanente Outputtabelle umgewandelt wurde.

Aufgrund der deutlich reduzierten Rechenzeit wird empfohlen, wann immer möglich, die Abfragen ausgewählter therapeutischer Gebiete den Gesamtabfragen vorzuziehen.

6.4.8 Signalwegranglisten

Während durch die Vernetzungsanalysen der therapeutischen Gebiete und Indikationsgruppen diejenigen Kontaktbereiche identifiziert werden konnten, die besonders eng miteinander verbunden sind, stellte sich im nächsten Schritt die Frage, welche Signalwege für diese Vernetzung verantwortlich sind. Wie in den Abschnitten 6.4.5 und 6.4.6 beschrieben wurde, setzt sich der „Pathway Score“, also der gewichtete Wert der Verbindungsstärke zwischen zwei therapeutischen

Gebieten oder zwei Indikationsgruppen, aus der Summe der einzelnen gewichteten Signalwegverbindungen zusammen. Das bedeutet, dass es für jede Verbindungsstelle zwischen therapeutischen Gebieten und Indikationsgruppen eine Liste an Signalwegen gibt, die unterschiedlich stark zum Gesamtwert in der Vernetzungsmatrix beitragen. Auch wenn die Information der Verbindungsstärke eines einzelnen Signalwegs zwischen zwei Gebieten also bereits implizit in der gewichteten Vernetzungsmatrix enthalten war, war es sinnvoll, diese Information so aufzubereiten, dass sie für eine weitere Analyse der Vernetzung geeignet war. Zu diesem Zweck wurden in der Datenbank Signalwegranglisten erstellt, die für eine ausgewählte Verbindungsstelle alle verbindenden Signalwege listen und ihre jeweilige Verbindungsstärke vergleichen. Zusätzlich wurde die Anzahl der Zielmoleküle berechnet, die innerhalb eines Signalwegs für die Verbindung zwischen den Gebieten direkt und indirekt verbindend sind.

6.4.8.1 Signalwegranglisten für Verbindungen zwischen therapeutischen Gebieten

Die Abfrage für therapeutische Gebiete griff auf die Berechnung des Gewichts der einzelnen verbindenden Signalwege zurück, welches im Rahmen der Gewichtungsabfragen berechnet wurde (vergleiche „Weighting Base 3“, Abschnitt 6.4.5). Entsprechend der gewählten Gewichtungsart (kein Durchschnitt, einfacher Durchschnitt, doppelter Durchschnitt, doppelter Durchschnitt mit Mulitplikator) variierte das Gewicht des Signalwegs. Um eine vollständige Rangliste mit möglichst umfangreicher Information zu erhalten, wurde die Abfrage zur Erstellung der Signalwegrangliste so konzipiert, dass sie alle vier möglichen Gewichte des Signalwegs auflistete, sowie zusätzlich pro Signalweg die Anzahl der direkt und indirekt verbindenden Zielmoleküle nannte.

Die Abfrage setzte mit Hilfe von Beziehungen alle unterschiedlichen Gewichtungsergebnisse miteinander in Verbindung, indem die gleichartigen Felder „TA name1“, „TA name2“ und „Pathway1“ durch Beziehungen verbunden wurden. Dies führte dazu, dass für eine bestimmte Kombination von therapeutischen Gebieten alle Gewichtungen unabhängig errechnet, dann jedoch die Ergebnisse der gleichen Kombination zugeordnet werden konnten. Da für die

Berechnung der Signalwegranglisten alle Arten der Gewichtung in Microsoft Access berechnet werden mussten, wurde eine „Parameter“-Funktion verwendet, um die Berechnung auf eine bestimmte Kombination von therapeutischen Gebieten einzuschränken. Dies reduzierte die Dauer der Berechnung deutlich.

Während der eigentlichen Berechnung griff die Abfrage auf alle vier Intermediärtabellen „Weighting Base 3“ der unterschiedlichen Gewichtungsarten zu und zog aus diesen Abfragen die jeweilige Verbindungsstärke eines Signalwegs. Zusätzlich zu den Gewichtungstabellen war für die Signalwegrangliste das Ergebnis der „Direct target check“-Abfrage (vergleiche Abb. 10.34) wichtig. Diese wurde wird im Rahmen der Gewichtung mit doppelter Durchschnittsbildung und Multiplikator berechnet und klassifizierte für eine Kombination aus therapeutischen Gebieten und Signalweg alle Zielmoleküle als direkt oder indirekt verbindend. Im Rahmen der Signalwegrangliste wurden die Gesamtzahl der Zielmoleküle sowie die Anzahl der direkt und indirekt verbindenden Zielmoleküle ermittelt. Die Unterscheidung zwischen direkte und indirekt verbindenden Zielmolekülen erfolgte dabei durch eine selektive Zählfunktion, die die Moleküle nur dann zählte, wenn sie durch eine 1 (Wert für indirekte Verbindungen in der „Direct target check“-Abfrage) oder durch eine 2 (Wert für direkte Verbindungen in der „Direct target check“-Abfrage) beschrieben waren (siehe Abb. 10.37).

Das Endprodukt der Abfrage wurde als permanente Outputtabelle in Microsoft Access gespeichert. In der Grundeinstellung wurden die verbindenden Signalwege nach Anzahl der relevanten Zielmoleküle und nach dem Gewicht mit einfacher Durchschnittsbildung sortiert. Die Verwendung der Sortierfunktion erlaubte allerdings jederzeit eine Sortierung nach einem anderen Gewicht.

6.4.8.2 Signalwegranglisten für Verbindungen zwischen Indikationsgruppen

Das Grundprinzip der Abfrage zur Erstellung von Signalwegranglisten von Verbindungen zwischen Indikationsgruppen (siehe Abb. 10.38) entsprach dem zur Verbindung zwischen therapeutischen Gebieten. Auch hier wurden diejenigen Einzelgewichte von Signalwegen genutzt, die im Rahmen der Gewichtung berechnet worden waren (vergleiche Abschnitt 6.4.6). Größter Unterschied zur

Rangliste bei therapeutischen Gebieten war, dass die Abfrage nicht die Intermediärtabellen der Gewichtung der therapeutischen Gebiete sondern die der Indikationsgruppen nutzte. Um die Rechenzeit zu begrenzen, bezog sich die Abfrage nicht auf die Gesamtmatrix aller Indikationsgruppen, sondern auf die „Select"-Intermediärtabellen, bei denen bereits eine Auswahl der therapeutischen Gebiete stattgefunden hatte. Für die Gewichtung mit doppelter Durchschnittsbildung und Multiplikator galt, dass die Berechnung ohnehin immer mit Begrenzung auf eine Kombination von therapeutischen Gebieten verlief.

Genauso wie oben für die therapeutischen Gebiete beschrieben, wurde eine „Parameter"-Funktion genutzt, um den Namen der gewünschten Indikationsgruppen abzufragen. Bei der Auswahl der Indikationsgruppen musste Bezug auf die therapeutischen Gebiete genommen werden, die bei Berechnung der „Select"-Intermediärtabellen ausgewählt worden waren. Die Abfrage (siehe Abb. 10.38) zog die gewichteten Verbindungsstärken der einzelnen Signalwege aus den Intermediärtabellen („Weighting Base 3") der vier Gewichtungsarten. Genauso wie auf Ebene der therapeutischen Gebiete wurden auch für die Rangliste der Indikationsgruppenverbindungen die Anzahl der verbindenden Zielmoleküle sowie deren Unterteilung in direkt und indirekt verbindende Moleküle vorgenommen. Auch diese Abfrage wurde in Access so angelegt, dass eine permanente Outputtabelle zur weiteren Analyse zur Verfügung stand.

6.4.8.3 Anteilige Verbindungsstärke

Die Signalwegranglisten für therapeutische Gebiete und Indikationsgruppen enthalten absolute Werte für die Verbindungsstärke eines Signalwegs. Im Rahmen der Auswertung wurden die absoluten gewichteten Verbindungsstärken der Signalwege als Anteil der Gesamtstärke einer Verbindung zwischen zwei therapeutischen Gebieten bzw. Indikationsgruppen dargestellt. Zu diesem Zweck wurde der totale Wert der Verbindungsstärke entweder der gewichteten Vernetzungsmatrix entnommen, oder als Summe der einzelnen Verbindungsstärken der Signalwege errechnet. Die anteilige Verbindungsstärke eines Signalwegs für eine Verbindung beschreibt den Anteil, den ein ausgewählter Signalweg an der Gesamtverbindung hat (vergleiche (6.4).

$$Anteil\ Signalweg\ A(\%) = \frac{Verbindungsstärke\ Signalweg\ A}{Gesamtverbindungsstärke\ aller\ Signalwege} \quad (6.4)$$

6.4.9 Identifikation von gemeinsamen Zielmolekülen

Die bisher beschriebenen Analysen ermöglichten neben der Untersuchung der Vernetzung von therapeutischen Gebieten und Indikationsgruppen untereinander die Identifizierung der für die Vernetzung verantwortlichen Signalwege. Nächster Analyseschritt war nun, für einen ausgewählten Signalweg diejenigen Zielmoleküle zu identifizieren, die für die Verbindung der Indikationsgruppen auf molekularer Ebene verantwortlich sind. Zu diesem Zweck wurden drei unterschiedliche Arten von Zielmolekülen unterschieden:

- **Bereits heute indikationsübergreifende Zielmoleküle („Common targets“):** Zielmoleküle, für die es bereits heute aktive Substanzen in beiden Indikationsgruppen gibt
- **Kandidaten für zukünftig indikationsübergreifende Zielmoleküle („Future common targets“)**: Zielmoleküle, die Teil eines verbindenden Signalwegs sind, aber erst innerhalb einer der Indikationsgruppen pharmazeutisch genutzt werden
- **Neue Zielmoleküle („New targets“):** Zielmoleküle, die Teil eines verbindenden Signalwegs sind, aber bisher für keine der beiden Indikationsgruppen pharmazeutisch genutzt werden

Ziel der Analysen war es, für eine ausgewählte Schnittstelle von Indikationsgruppen und einen verknüpfenden Signalweg alle drei Gruppen von Zielmolekülen zu identifizieren, aufzulisten und – sofern vorhanden – weitere Informationen über ihre pharmazeutische Nutzung aus der Datenbank zu extrahieren.

In der ersten Abfrage wurde eine Parameterfunktion verwendet, um durch den Nutzer die Namen der ausgewählten Indikationsgruppen sowie des zu untersuchenden Signalwegs einzugeben. Die Abfrage erstellte eine Liste aller Zielmoleküle innerhalb des ausgewählten Signalwegs, die für mindestens eine der beiden Indikationsgruppen bereits momentan pharmazeutisch genutzt werden.

Dabei wurden für jedes Zielmolekül die „Gene ID“, der Name des Moleküls sowie die Indikationsgruppe, für die das Zielmolekül relevant ist, aufgelistet. Ist ein Zielmolekül bereits jetzt für beide Indikationsgruppen relevant, wurde es zweifach aufgeführt. Wie Abb. 10.39 zeigt, nutzte die Abfrage neben den Inputtabellen „TAs and indications“, „Indications and drugs“, „Drugs and targets“ und „Targets and pathways“ auch die Vernetzungsmatrix der Indikationsgruppen in der die verbindenden Signalwege zwischen zwei Indikationsgruppen aufgelistet worden waren. Gleichartige Felder wurden durch Beziehungen untereinander verknüpft. Die Verbindung der Inputtabellen mit den Intermediärtabellen war sinnvoll, um weitere Informationen zu den Zielmolekülen (wie zum Beispiel Firmen und Indikationsstatus) zuordnen zu können. Die Auswahl derjenigen Zielmoleküle, die über Substanzen und deren Indikationen den Indikationsgruppen zugeordnet werden können, erfolgte mit Hilfe einer logischen Formel (siehe „Expr3“, Abb. 10.39).

In nächsten Abfrageschritt wurden die Ergebnisse aus der ersten Abfrage über eine Zählfunktion so zusammengefasst, dass direkt ablesbar wurde, ob ein Zielmolekül bereits jetzt ein „Common target“ ist, d.h. für beide ausgewählte Indikationsgruppen von Bedeutung ist, oder erst für eine der beiden Indikationsgruppen pharmazeutisch genutzt wird und damit in die Gruppe der „Future common targets“ fällt. Die Zählfunktion kennzeichnete ein Zielmolekül, das in der vorhergehenden Abfrage zwei Einträge hatte, d.h. für beide Indikationsgruppen vorkam, mit einer „2“. Zielmoleküle, die nur einen Eintrag in der Abfrage aus Abb. 10.39 hatten, wurden mit einer „1“ markiert (siehe Abb. 10.40).

Die nächste Abfrage erstellte nun eine Tabelle, die nur die bereits heute indikationsübergreifenden Zielmoleküle enthielt, d.h. der Moleküle, die in der vorhergehenden Abfrage eine Kennziffer „2“ erhalten hatten (siehe Abb. 10.41). Zur Dokumentation wurde eine „Make-table“-Abfrage angeschlossen, durch die eine permanente Outputtabelle erstellt wurde (nicht gezeigt).

Analog zu der in Abb. 10.41 gezeigte Abfrage wurde mit Hilfe einer weiteren Abfrage die Liste derjenigen Zielmoleküle erstellt, die momentan nur für eine der

beiden Indikationsgruppen relevant sind, die aber möglicherweise in Zukunft indikationsübergreifend sein könnten. Die dafür verwendete Abfrage wurde analog zu Abfrage aus Abb. 10.41 angelegt und wird daher nicht gezeigt. Einziger Unterschied war, dass in der zweiten Variation die Moleküle ausgewählt werden, die durch die Ziffer „1“ gekennzeichnet worden waren. Auch aus dieser Tabelle wurde eine permanente Outputtabelle erstellt (nicht gezeigt).

Sowohl für die bereits heute indikationsübergreifenden Zielmoleküle als auch für die Kandidaten für zukünftig indikationsübergreifende Zielmoleküle wurden zwei weitere Abfragen angeschlossen, mit denen detaillierte Informationen zu den einzelnen Molekülen zusammengestellt wurden. Diese waren für die Validierung der Ergebnisse und die Priorisierung der Kandidaten für zukünftig indikationsübergreifende Zielmoleküle hilfreich. Die erste Abfrage („Detailed info 1“) ist in Abb. 10.42 gezeigt und listete für jedes Zielmolekül Informationen über die pharmazeutisch aktiven Substanzen, die es adressieren (Name der Substanz, Indikationen und Entwicklungsstatus der Forschungsprojekte).

Die zweite Abfrage („Detailed info 2“, Abb. 10.44) listete die entwickelnde Firma („Originator“) sowie mögliche Lizenznehmer („Licensee“). Diese Information konnte nicht gemeinsam mit den Indikationen aus der vorhergehenden Abfrage aufgelistet werden, da die Firmenzugehörigkeit in den Pharmaprojects Rohdaten unabhängig von den Indikationen war und daher eine gemeinsame Auflistung einen nicht vorhandenen Zusammenhang vortäuschen würde (z.B. Entwicklung einer bestimmten Indikation nur vom Originator und nicht vom Lizenznehmer).

Die Abfragen zur Auflistung der detaillierten Informationen zu den Kandidaten für zukünftig indikationsübergreifende Zielmoleküle entsprachen denen, die für die indikationsübergreifenden Zielmoleküle in den Abb. 10.42 und Abb. 10.44 gezeigt wurden. Einziger Unterschied war, dass sich die Abfragen nicht auf die Tabelle mit der Liste der „Common Targets“ sondern auf die der „Potential future common targets“ bezogen. Sie werden sie hier nicht gezeigt.

Um die dritte Kategorie der neuen Zielmoleküle zu identifizieren, wurde eine Abfrage in drei Schritten definiert. Der erste Schritt („All Kegg Targets“, Abb. 10.43) erstellte eine Liste derjenigen Gene („Gene IDs“), die entsprechend der

KEGG Rohdatentabelle „Targets and pathways“ einem Signalweg zugeordnet waren. Dabei spielte es keine Rolle, ob die Zielmoleküle pharmazeutisch aktiv waren oder nicht.

Im nächsten Schritt wurde eine Liste derjenigen Zielmoleküle erstellt, die bereits für eines der beiden Gebiete pharmazeutisch genutzt werden und daher nicht mehr ganz neu sind (siehe Abb. 10.45).

Im letzten Analyseschritt (siehe Abb. 10.46) wurde nun eine Differenz gebildet, die die bereits bekannten Zielmoleküle (Ergebnis aus Abb. 10.45, „Known targets in industry“) von der Gesamtmenge der Zielmoleküle (Ergebnis aus Abb. 10.43, „All KEGG targets“) abzog. Die daraus erstellte Ergebnistabelle enthielt nur die Zielmoleküle, die bisher für keine der beiden Indikationsgruppen pharmazeutisch genutzt werden, jedoch ebenfalls Teil des verbindenden Signalwegs sind.

6.4.10 Untersuchungen zur Charakterisierung der Datenbasis

Die folgenden Analysen wurden zur Darstellung der Datenbasis genutzt. Sie basieren entweder direkt auf den Rohdatentabellen der Datenbank oder nutzen Intermediärtabellen der Vernetzungsanalysen. Untersuchungen zur Charakterisierung der Datenbasis waren notwendig, um die Ergebnisse von Vernetzung- und Signalweganalyse richtig interpretieren zu können. Die Abfragen zur allgemeinen Statistik wurden alle als „Select“-Abfragen formuliert und stehen daher nicht mit Outputtabellen im Zusammenhang.

6.4.10.1 Eigenschaft der therapeutischen Gebiete

Die im Folgenden beschriebenen Analysen dienten dazu, die Verteilung der Rohdaten auf die unterschiedlichen therapeutischen Gebiete zu erfassen und mögliche Unterschiede zwischen den einzelnen Gebieten darzustellen. In der ersten Analyse wurde für jedes therapeutische Gebiet die Anzahl der Substanzen ermittelt, die in der Datenbank für Analysen zur Verfügung stehen. Wie bereits im Abschnitt 6.2.1 (Rohdaten Pharmaprojects) beschrieben, wurde im Rahmen der Rohdatenaufbereitung jede Kombination von Substanz und Indikation als einzelnes Forschungsprojekt und damit eigener Eintrag in der Datenbank angelegt.

Das bedeutet, dass eine Substanz für jedes therapeutische Gebiet relevant ist, für das sie mindestens eine Indikation in Forschung und Entwicklung bzw. auf dem Markt hat. Da es sich bei therapeutischen Gebieten nicht um vordefinierte Einheiten handelt, unterscheiden sich einzelne Gebiete in ihrer Größe, die als Anzahl der Forschungsprojekte bestimmt wird. Um die Größe der einzelnen Gebiete miteinander vergleichen zu können, wurde für jedes Gebiet die Anzahl der Forschungsprojekte ermittelt. Zur Analyse wurden die beiden Rohdatentabellen „TAs and indications“ und „Indications and drugs“ über das gleichartige Feld der Indikation verknüpft (vergleiche Abb. 10.47). Eine Zählfunktion ermittelte für jedes therapeutische Gebiet („TA“) die Zahl der Forschungsprojekte.

Wie die Abbildung Abb. 6.4 im Abschnitt 6.2.1 bereits zeigen konnte, verteilen sich die Forschungsprojekte auf verschiedene Phasen von Forschung und Entwicklung. Da eine Variation dieser Verteilung je nach therapeutischem Gebiet vorliegt, wurde dies ebenfalls durch eine Rohdatenanalyse dargestellt. Für jeden Entwicklungsstatus (d.h. zum Beispiel Phase I der Klinik) wurde die Anzahl der Forschungsprojekte nach therapeutischem Gebiet gelistet. Je nach Kontext der Auswertung wurde die Reihenfolge der Spalten „Any Indication Status“ und „TA name“ in der Microsoft Access Abfrage getauscht, um ebenfalls eine Verteilung der Phasen innerhalb eines therapeutischen Gebiets übersichtlich sehen zu können.

6.4.10.2 Eigenschaften der Signalwege

Wie bereits in Abschnitt 6.2.2 beschrieben wurde, sind insgesamt 56 humane Signalwege Teil der Datenbank. Diese Signalwege unterscheiden sich in ihrer natürlichen Größe (Anzahl der Zielmoleküle innerhalb eines Signalwegs) und tragen daher auch unterschiedlich stark zur Vernetzung von therapeutischen Gebieten bei. Zunächst wurde in einer Analyse dargestellt, wie viele Zielmoleküle die Signalwege entsprechend KEGG enthalten. Zu diesem Zweck wurde eine Abfrage basierend auf der Inputtabelle „Targets and pathways“ formuliert, die die Anzahl der Gene („Target IDs“) pro Signalweg zählte (siehe Abb. 10.49).

Die eben beschriebene und in Abb. 10.49 dargestellte Abfrage zeigte nur die Anzahl der Zielmoleküle, die in den Rohdaten vorhanden sind. Sie stellte keinen Zusammenhang mit den Pharmaprojects Daten her und konnte daher keine Auskunft darüber geben, wie viele dieser Zielmoleküle auch pharmazeutisch genutzt werden und somit zur Berechnung der Vernetzung beitragen.

Um dies zu überprüfen und damit die Anzahl der vernetzungsrelevanten Zielmoleküle pro Signalweg zu bestimmen, wurden zwei weitere Abfragen angeschlossen. Im ersten Schritt der Analyse (siehe Abb. 10.50) wurden für jeden Signalweg diejenigen Zielmoleküle aufgelistet, die bereits pharmazeutisch adressiert sind. Dies wurde erreicht, indem das Feld „Any Gene ID“ nicht aus der Rohdatentabelle „Targets and pathways“ gezogen wurde, sondern stattdessen aus der Rohdatentabelle „Drugs and targets“. In der „Drugs and targets“-Tabelle sind nur die Zielmoleküle enthalten, für die es auch Substanzen in Forschung und Entwicklung gibt.

Im zweiten Schritt wurde eine einfache Zählfunktion angeschlossen, die basierend auf dem Ergebnis des ersten Schritts ermittelte, wie viele aktive Zielmoleküle pro Signalweg vorkommen (siehe Abb. 10.51).

Die zuvor beschriebenen Ansätze wurden genutzt, um die unterschiedliche Größe der Signalwege (biologisch insgesamt und pharmazeutisch adressiert) zu ermitteln.

Der Einfluss der unterschiedlichen Größe der Signalwege auf die Vernetzung der therapeutischen Gebiete und Indikationsgruppen hängt vor allem von der Anzahl der therapeutischen Gebiete bzw. der Indikationsgruppen ab, in denen ein Signalweg eine Rolle spielt. Ein Signalweg spielt immer dann für ein therapeutisches Gebiet oder eine Indikationsgruppe eine Rolle, wenn es ein Zielmolekül im Signalweg gibt, das von einer Substanz adressiert wird, für die es eine Indikation in diesem Gebiet bzw. in dieser Indikationsgruppe gibt. Die Abfrage zur Bestimmung der Anzahl der Gebiete, für die ein Signalweg wichtig ist (vergleiche Abb. 10.52), stellt die Umkehrung der Intermediärtabelle „TAs and Pathways“ dar, die im Rahmen der Vernetzungsanalyse erstellt worden war und Signalwege zu therapeutischen Gebieten zugeordnet hatte. Mit Hilfe einer

Zählfunktion wurde für jeden Signalweg die Zahl der relevanten therapeutischen Gebiete ermittelt.

Die Ermittlung der Anzahl Indikationsgruppen, in denen ein Signalweg eine Rolle spielt, erfolgte analog zur oben dargestellten Ermittlung der therapeutischen Gebiete. Einziger Unterschied war, dass die Abfrage statt der Tabelle „TAs and Pathways“ die Intermediärtabelle „SubTAs and Pathways“ nutzte, die die Zuordnung der Indikationsgruppen („SubTAs“) zu Signalwegen im Rahmen der Vernetzungsanalyse auf Indikationsgruppenebene gespeichert hatte. Aufgrund der Ähnlichkeit wird die Abfrage hier nicht gezeigt.

6.4.10.3 Eigenschaften der Zielmoleküle

Zur näheren Beschreibung der Zielmoleküle wurden zwei Eigenschaften untersucht: Die Zugehörigkeit von Zielmolekülen zu Signalwegen und die Anzahl von Substanzen pro Zielmolekül. Ersteres ist für die Auswertung relevant, da ein Zielmolekül grundsätzlich Teil von mehreren Signalwegen sein kann. Je größer diese Anzahl an Signalwegen ist, desto „universeller“ ist das Zielmolekül für den Organismus und desto größer ist gleichzeitig seine Möglichkeit, unterschiedliche Signalwege und therapeutische Gebiete miteinander zu vernetzen.

Da lediglich diejenigen Zielmoleküle vernetzen können, die auch durch Substanzen adressiert werden, wurden die Zielmoleküle von der Analyse ausgeschlossen, die zwar in den KEGG Rohdaten vorkommen, aber kein Äquivalent auf Seiten der Pharmaprojects Daten haben. In Microsoft Access wurde dies dadurch abgebildet, dass nur diejenigen Zielmoleküle („Target IDs“) in der Abfrage berücksichtigt wurden, für die es ein Gegenstück („Any target name“) aus der Pharmaprojects Tabelle „Drugs and targets“ gab. Jedem dieser „aktiven“ Zielmoleküle wurden im ersten Schritt der Analyse alle Signalwege zugeordnet, in denen das Zielmolekül entsprechend der KEGG Rohdaten vorkommt (siehe Abb. 10.53). Im zweiten Schritt wurde dann durch eine Zählfunktion die Anzahl der Signalwege pro Zielmolekül ermittelt (siehe Abb. 10.54).

Es wurde weiterhin untersucht, wie viele Substanzen für ein und dasselbe Zielmolekül existieren. Wie im Abschnitt 6.2.1 bereits beschrieben wurde, wird

eine Substanz immer dann in der Datenbank aufgeführt, wenn sie unabhängig von einer Firma entwickelt wurde (d.h. nicht unter Lizenzverträgen). Die Anzahl der Substanzen pro Zielmolekül ist also sowohl von den Eigenschaften des Zielmoleküls (z.B. Eignung für pharmazeutische Entwicklung) als auch vom Interesse der Firmen an Medikamenten, die am Zielmolekül angreifen, abhängig. Um mögliche Schwankungen zwischen therapeutischen Gebieten darzustellen, wurde die durchschnittliche Anzahl von Substanzen für ein Zielmolekül für jedes therapeutische Gebiets ermittelt. Ist ein Zielmolekül durch die damit verbundenen Forschungsprojekte für mehr als ein therapeutisches Gebiet von Bedeutung, so wurde es mit seinen entsprechenden Substanzen in jedem Gebiet berücksichtigt. Im ersten Schritt der Analyse wurde pro therapeutisches Gebiet die Anzahl der Substanzen pro Zielmolekül aufgelistet (siehe Abb. 10.55). Im zweiten Schritt wurde der Durchschnitt der Anzahl der Substanzen pro Zielmolekül über jedes therapeutische Gebiet mit Hilfe einer Durchschnittsfunktion („Avg“, siehe Abb. 10.56) ermittelt.

6.4.10.4 Eigenschaften der Substanzen

Wie in Abschnitt 6.4.5 beschrieben, haben einzelne Substanzen eine oder mehrere Indikationen, die jeweils einzeln bei der Gewichtung der Verbindungsstärke eines Signalwegs berücksichtigt wurden. Da sich einzelne therapeutische Gebiete durch die durchschnittliche Anzahl von Indikationen pro Substanz unterscheiden können, musste diese bei der Interpretation von Ergebnissen berücksichtigt werden und/oder bereits bei der Gewichtung als Korrekturfaktor eingehen (siehe Gewichtung mit Durchschnittsbildung). Zur Ermittlung von gebietsspezifischen Unterschieden in der Anzahl der Indikationen pro Substanz wurde eine Analyse in zwei Schritten durchgeführt. Im ersten Schritt (siehe Abb. 10.57) wurden basierend auf den beiden Inputtabellen „TAs and indications“ und „Indications and drugs“ für jedes therapeutische Gebiet alle Substanzen gelistet und pro Substanz die Anzahl der Indikationen gezählt. Ähnlich wie für die Analyse der durchschnittlichen Anzahl von Substanzen pro Zielmolekül wurde dann im zweiten Schritt ein Durchschnitt pro therapeutisches Gebiet gebildet. (siehe Abb. 10.58).

6.5 Statistische Methoden

Im Rahmen der Doktorarbeit wurde lineare und nicht-lineare Regressionsanalyse zur Überprüfung von Zusammenhängen zwischen Inputdaten verwendet. Die Anwendung beschränkte sich auf die Untersuchung einer Korrelation einer Variablen (z.B. Anzahl der Indikationsgruppen, in denen ein Signalweg vorkommt) und einer anderen Variablen (z.B. Größe des Signalwegs).

Bei der linearen Regression wurden die Rohdaten mit Hilfe von Microsoft® Office Excel und GraphPad Prism® in einem *X/Y*-Diagramm dargestellt. Durch die Analysefunktion der linearen Regression wurde eine Trendlinie hinzugefügt und das Bestimmtheitsmaß r^2 durch Verwendung der „minimum sum-of-squares" Formel berechnet. Dabei errechnet das verwendete Programm die Trendlinie so, dass die Summe der Fehlerquadrate der einzelnen Datenpunkte so klein wie möglich ist. r^2 wird dann als Quotient der Fehlerquadrate der Regressionsgerade und der Fehlerquadrate einer Geraden durch den Mittelwert der Datenpunte angegeben. Je näher r^2 dabei bei 1 liegt, desto besser ist die Trendlinie geeignet, um den Datenzusammenhang zu beschreiben. Ob ein Zusammenhang zwischen den Variablen *X* und *Y* bestand, wurde dann entsprechend von r^2 beurteilt. Die statistische Signifikanz von r^2 wurde in GraphPad Prism® bzw. Microsoft® Office Excel mit Hilfe eines F-Tests bestimmt. r^2-Werte mit einem P-Wert unter 0,05 wurden als signifikant bezeichnet und mit * gekennzeichnet. Bei $P < 0{,}01$ wurden ** und bei $P < 0{,}001$ wurden *** zur Kennzeichnung des r^2 Werts verwendet. Die nicht-lineare Regression wurde in GraphPad Prism® durchgeführt. Dabei wurde das Regressionsmodell so ausgewählt, dass das Bestimmtheitsmaß R^2 möglichst hoch ist und sich bei einem Runs-Test keine singifikante Abweichung der Daten vom Regressionsmodell feststellen lässt. GraphPad Prism® berechnet R^2 dabei erneut nach der „minimum sum-of-squares"-Formel.

Die oben beschriebene Regressionsanalyse wurde im Rahmen der vorliegenden Arbeit zur Untersuchung der folgenden Zusammenhänge verwendet:

- Abhängigkeit des relativen Vernetzungsgrads von der Anzahl der Projekte pro therapeutisches Gebiet (mit und ohne Gewichtung)
- Abhängigkeit der Anzahl der durch einen Signalweg verbundenen Gebiete von der Anzahl der Zielmoleküle pro Signalweg
- Abhängigkeit der aktiven Größe (Anzahl pharmazeutisch adressierter Zielmoleküle) eines Signalwegs von der absoluten Größe (Anzahl aller Zielmoleküle) eines Signalwegs

6.6 Bewertungsmatrix und Bewertungsmethoden

Zur Bewertung der im Rahmen der Verbindungsanalysen identifizierten Kandidaten für mögliche zukünftig indikationsübergreifende Zielmoleküle wurde ein Bewertungsschema mit den folgenden vier Bewertungsbereichen entworfen:

- Biologische Evidenz für indikationsübergreifende Funktion
- Pharmazeutische Machbarkeit
- Wettbewerbsintensität
- Strategische Eignung

Zielsetzung der Bewertung war die Priorisierung der Kandidaten für mögliche zukünftig indikationsübergreifende Zielmoleküle mit Hilfe von klar definierten qualitativen und quantitativen Kriterien. Im Rahmen der vorliegenden Arbeit wurden für alle vier Bereiche Bewertungskriterien erarbeitet. Darüber hinaus wurden exemplarisch fünf Zielmoleküle aus dem Signalweg Apoptose für die Verbindung der Indikationsgruppen rheumatoide Arthritis und Lungenkrebs bewertet.

Da eine Bewertung der strategischen Eignung nur im firmenspezifischen Zusammenhang sinnvoll durchführbar ist, wurden für diesen Bereich nur Bewertungskriterien definiert, jedoch keine exemplarische Bewertung durchgeführt.

Für alle Bereiche wurden Bewertungskriterien identifiziert und drei unterschiedliche Ausprägungsstufen für die Kriterien definiert. Für jedes Zielmolekül wurde die Ausprägung des Bewertungskriteriums untersucht, mit den definierten Ausprägungen verglichen und die Bewertung entsprechend der am besten zutreffenden Ausprägung vergeben. Da es sich bei den meisten Kriterien um qualitative Einschätzungen handelt, wurde von einer numerischen Berechnung der Gesamtbewertung abgesehen. Um trotzdem einen groben Vergleich der Bewertungen zu ermöglichen, wurde eine sogenannte „Mond“-Darstellung verwendet, die die Ausprägung eines Bewertungskriteriums symbolisiert (vergleiche Abb. 6.10, Abb. 6.11, Abb. 6.12 und Abb. 6.13).

Die beispielhafte Bewertung der Zielmoleküle verlief in mehreren Schritten. Im ersten Schritt wurden für alle zu bewertenden Zielmoleküle Kurzprofile erstellt, die allgemeine Informationen über das Zielmolekül wie Funktion, Struktur und Regulation enthalten. Im zweiten Schritt wurde die Literatur gesichtet, die Informationen zur mögliche Funktion des Moleküls in bisher noch nicht durch das Zielmolekül erschlossenen Indikationsgruppe enthielt. Relevante Veröffentlichungen wurden zusammengefasst. Im dritten Schritt wurden für jedes Molekül Proteindatenbanken (insbesondere Uniprot (The UniProt, 2008) durchsucht und die wichtigsten Informationen zusammengefasst. Im letzten Schritt wurden dann die Analysen durchgeführt, die die Datenbank verwenden. Basierend auf diesen Rohdaten wurde die Bewertung entlang der im Folgenden beschriebenen Bereichen durchgeführt sowie eine zusammenfassende Aussage formuliert.

6.6.1 Biologische Evidenz für indikationsübergreifende Zielmoleküle

Kandidatenmoleküle für indikationsübergreifende Funktion sind Zielmoleküle, die Teil eines verbindenden Signalwegs zwischen zwei Indikationsgruppen sind und bisher nur für eine der beiden Indikationsgruppen adressiert wurden. Die biologische Evidenz der indikationsübergreifenden Funktion beschreibt die Wahrscheinlichkeit, dass ein Kandidatenmolekül in Zukunft ebenfalls für die andere, „fehlende“ Indikationsgruppe genutzt werden kann. Dabei handelt es sich um eine qualitative Abschätzung der Möglichkeiten basierend auf den momentan verfügbaren Informationen. Aufgrund der qualitativen Natur der Einschätzungen ist eine abweichende Interpretation der Informationen denkbar.

Die biologische Evidenz wurde mit Hilfe der drei folgenden Kriterien bewertet:

- Position im Signalweg
- Anhaltspunkte für Funktion in der fehlenden Indikation
- Vernetzungseigenschaft des Zielmoleküls über Signalwege

Die Informationen für die Bewertung der Position im Signalweg wurden aus der KEGG Datenbank entnommen. Anhaltspunkte für die Funktion des Zielmoleküls

wurden durch Literaturrecherche gewonnen. Zu diesem Zweck wurde die PubMed Datenbank verwendet (U.S. National Library of Medicine, 2008b). Für jeden Suchbegriff (d.h. Name der Indikationsgruppe) wurden mit Hilfe des Mesh-Browsers (U.S. National Library of Medicine, 2008a) alle gängigen Bezeichnungen und Synonyme herausgesucht und iterativ für die Suche verwendet. Die Suche wurde auf Literatur ab dem Veröffentlichungsjahr 2000 begrenzt. Darüber hinaus wurden für die Fragestellung relevante Referenzen aus den einzelnen Veröffentlichungen verfolgt, auch wenn sie nicht durch Suche mit den Suchbegriffen gefunden worden waren. Im Rahmen der exemplarischen Bewertung wurden die folgenden Suchbegriffe in unterschiedlichen Kombinationen mit den Operatoren OR und AND verwendet:

- **Akt1**: „Akt1“, „protein kinase B“, „PKB“, „rheumatoid arthritis“
- **Interleukin 1 β**: “interleukin-1 beta”, “IL-1 beta”, “IL1 beta”, “lung cancer”, “lung carcinoma”
- **Interleukin 1 Rezeptor Typ I**: “interleukin-1 receptor type I”, “CD121a”, “interleukin-1 receptor”, “lung cancer”, “lung carcinoma
- **p53**: „p53“, „rheumatoid arthritis“
- **Death receptor 4**: “rheumatoid arthritis”, “TNFRSF10A”, “TRAIL” “DR4”, “death receptor 4”, “tumor necrosis factor receptor superfamily, member 10a”

Insgesamt wurden so je nach Zielmolekül zwischen 20 und 90 Treffer erzielt. Jede Veröffentlichung wurde als für die Fragestellung „sehr relevant“, „bedingt relevant“ oder „nicht relevant“ beurteilt. Für alle Veröffentlichungen, die als „sehr relevant“ oder „bedingt relevant“ eingestuft worden waren, wurde ein Kommentar zum Hauptergebnis in Bezug auf das gesuchte Zielmolekül und die Indikation sowie die verwendete Methode erstellt. Die gesamte Literaturrecherche wurde in einem Microsoft Excel Dokument detailliert dokumentiert. Zwischen 30 und 70 % der Treffer der Datenbanksuche waren für die Fragestellung sehr relevant.

Die Vernetzungseigenschaft des Zielmoleküls über Signalwege wurde mit Hilfe derselben zweistufigen Abfrage aus der Datenbank bestimmt, die bereits zur

Charakterisierung der Zielmoleküle verwendet wurde (vergleiche Abschnitt 6.4.10, Abb. 10.53 und Abb. 10.54). Abb. 6.10 zeigt die Bewertungskriterien des ersten Bereichs der detaillierten Beschreibung der Ausprägungen.

Ausprägung der Bewertungskriterien	○	◐	●
Position im Signalweg	• Zielmolekül ist in separatem Unterbereich des Signalwegs ohne funktionellen Zusammenhang mit indkationsübergreifenden Zielmolekülen • Keine funktionelle Verbindung zu Molekülen, die für die „fehlende“ Indikationsgruppe relevant sind	• Zielmolekül steht über mehrere Zwischen-schritte in funktionellem Zusammenhang mit Molekülen, die bereits heute indikations-übergreifend sind • Zielmolekül ist an zentralem Schnittpunkt des Signalwegs	• Zielmolekül steht in direktem funktionellen Zusammenhang zu Molekülen, die bereits heute indikations-übergreifend sind • Direkter funktioneller Zusammenhang mit anderen Kandidaten-Molekülen, die heute für die „fehlende“ Indikationsgruppe relevant sind
Anhaltspunkte für Funktion in der fehlenden Indikationsgruppe	• Keine Anhaltspunkte zur Funktion des Moleküls in der “fehlenden” Indikationsgruppe • Keine Evidenz der Funktion anderer Subtypen	• In vitro Untersuchungen legen eine mögliche Funktion des Zielmoleküls in der fehlenden Indikations-gruppe nahe • Belege für Funktion ähnlicher Zielmoleküle vorhanden (z.B. anderer Subtypen)	• Evidenz für in vivo Funktion des Zielmoleküls in der „fehlenden“ Indikationsgruppe (z.B. im Tiermodell)
Vernetzungs-eigenschaft des Zielmoleküls über Signalwege	• Zielmolekül ist spezifisch für Signalweg	• Zielmolekül ist in bis zu fünf Signalwegen relevant	• Stark vernetzendes Zielmolekül, das in mehr als fünf Singalwegen vorkommt

Abb. 6.10: Die Bewertung der biologischen Evidenz der indikationsübergreifenden Funktion wurde mit drei Kriterien bewertet
Die Abbildung zeigt die drei Kriterien, die zur Bewertung der biologischen Evidenz der indikationsübergreifenden Funktion verwendet wurden: Position im Signalweg, Anhaltspunkte für die Funktion in der fehlenden Indikationsgruppe und Vernetzungseigenschaft des Zielmoleküle über Signalwege. Für jedes Kriterium wurden drei Ausprägungen definiert, die mit einem Mondsymbol gekennzeichnet sind. Der leere Mond (○) symbolisiert eine Ausprägung, die für die indikationsübergreifende Funktion eher schlecht ist während der volle Mond (●) eine Ausprägung beschreibt, die vermuten lässt, dass indikationsübergreifende Funktion möglich ist. Durch Vergleich der Situation bei dem zu bewertenden Zielmolekül mit den definierten Ausprägungen, konnte dem Zielmolekül eine Bewertung zugeordnet werden. Die Übergänge zwischen den einzelnen Ausprägungen sind dabei fließend, d.h. es sind auch Einordnungen zwischen den einzelnen Ausprägungen denkbar.

6.6.2 Pharmazeutische Machbarkeit

Zur Bewertung der pharmazeutischen Machbarkeit wurden zwei Kriterien genutzt:

- Biologische Eignung des Zielmoleküls für pharmazeutische Intervention
- Evidenz der Entwicklungsfähigkeit von Medikamenten gegen das Zielmolekül

Ausprägung der Bewertungskriterien	○	◑	●
Biologische Eignung des Zielmoleküls für pharmazeutische Intervention	• Zielmolekülfamilie ist z.B. durch ihre Lokalisierung in der Zelle schwer erreichbar (z.B. innere Membran-proteine, GPCRs) • Nur innovative Techniken können zum Erreichen des Zielmoleküls verwendet werden (z.B. Gentransfer, RNAi)	• Proteinstruktur und Bindestellen mäßig gut charakterisiert • Nur Ansätze mit „small molecules“ oder kleinen monoklonalen Antikörper (-fragmenten) möglich • Historisch Schwierigkeiten mit Spezifität und Medikamenten-entwicklung (z.B. HDACs, Proteasen)	• Bekannte Proteinstruktur mit gut charakterisierten Bindestellen • Verschiedene Ansätze möglich (“small molecules”, therapeutische Antikörper, therapeutische Viren, etc.) • Viel Erfahrung in Industrie mit Molekülklasse in Medikamentenentwicklung (z.B. Membranrezeptoren, Kinasen)
Evidenz der Entwicklungs-fähigkeit von Medikamenten gegen das Zielmolekül	• Zielmolekül im gesamten Spektrum therapeutischer Gebiete momentan nur für eine Indikationsgruppe adressiert • Projekt(e) ausschließlich in präklinischer Entwicklung	• Zielmolekül im gesamten Spektrum therapeutischer Gebiete nur für einige wenige Indikations-gruppen in max. zwei Gebieten adressiert • Mehrere Projekte in präklinischer und klinischer Entwicklung	• Zielmolekül im gesamten Spektrum therapeutischer Gebiete bereits für viele verschiedene Indikations-gruppen in mehreren therapeutischen Gebieten adressiert • Viele Projekte in mehreren Indikationen pro Zielmolekül • Projekte verteilt über alle Entwicklungsstadien • Mindestens ein Projekt bereits zugelassen

Abb. 6.11: Die pharmazeutischen Machbarkeit wurde mit zwei Kriterien bewertet
Die Abbildung zeigt die zwei Kriterien, die zur Bewertung der pharmazeutischen Machbarkeit dienten: Biologische Eignung des Zielmoleküls für pharmazeutische Intervention und Evidenz der Entwicklungsfähigkeit von Medikamenten gegen das Zielmolekül. Für jedes Kriterium wurden drei Ausprägungen definiert, die mit einem Mondsymbol gekennzeichnet sind. Der leere Mond (○) symbolisiert eine Ausprägung, die für die indikationsübergreifende Funktion eher schlecht ist, während der volle Mond (●) eine Ausprägung beschreibt, die vermuten lässt, dass indikationsübergreifende Funktion möglich ist. Durch Vergleich der Situation bei dem zu bewertenden Zielmolekül mit den definierten Ausprägungen, konnte dem Zielmolekül eine Bewertung zugeordnet werden.. Die Übergänge zwischen den einzelnen Ausprägungen sind dabei fließend, d.h. es sind auch Einordnungen zwischen den einzelnen Ausprägungen denkbar.

Ähnlich wie bei den Kriterien aus dem vorhergehenden Abschnitt wurden zur Bewertung der biologischen Eignung externe qualitative Datenquellen (v.a. UniProt (The UniProt, 2008) und Pubmed (U.S. National Library of Medicine, 2008b) verwendet.

Zur Bewertung der Entwicklungsfähigkeit von Medikamenten wurden die Abfragen zu zusätzlichen Informationen über die Kandidaten für mögliche zukünftig indikationsübergreifenden Zielmolekülen verwendet (vergleiche Abschnitt 6.4.9) und die Forschungsprojekte in den einzelnen Indikationsgruppen und Phasen gezählt. Dabei wurden diejenigen Projekte von der Analyse ausgeschlossen, bei denen keine definierte Indikation vorhanden war („unspecified“). Abb. 6.11 zeigt die detaillierte Beschreibung der Ausprägungen der einzelnen Bewertungskriterien.

6.6.3 Wettbewerbsintensität

Zur Bewertung der Wettbewerbsintensität konnte die Datenbank genutzt werden. Die Wettbewerbsintensität wurde mit Hilfe von zwei Kriterien bewertet:

- Anzahl und Status der bereits existierenden Substanzen
- Wettbewerberstruktur

Datengrundlage bieten erneut die Analysen mit den detaillierten Informationen über die potentiellen zukünftigen indikationsübergreifenden Zielmolekülen (vergleiche Abschnitt 6.4.9). Auch bei der Analyse zur Wettbewerbsintensität wurden alle Substanzen ohne definierte Indikation ausgeschlossen. Abb. 6.12 zeigt die detaillierte Definition der Ausprägungen der einzelnen Bewertungskriterien. Die Wettbewerbsintensität kann ohne firmenspezifische Informationen von außen („outside-in“) bewertet werden. Eine firmenspezifische Interpretation ist aber für die finale Priorisierungsentscheidung notwendig.

Ausprägung der Bewertungskriterien	○	◑	●
Anzahl und Status der bereits existierenden Substanzen für das Zielmolekül	• Hohe Anzahl verschiedener Substanzen für das Zielmolekül (>10) • Mehrere Substanzen bereits zugelassen • Noch nicht zugelassene Substanzen überwiegend in später klinischer Forschung (Phase III)	• Mittlere Substanzanzahl pro Zielmolekül (5-10) • Substanzen überwiegend in präklinischer oder früher klinischer Forschung	• Geringe Substanzanzahl pro Zielmolekül (< 5) • Noch keine zugelassenen Substanzen für das Zielmolekül
Wettbewerber-struktur	• Zahlreiche unterschiedliche Entwickler mit mehreren Substanzen für das Zielmolekül • Substanzen bereits mehrfach lizensiert • Substanzen im Programm einzelner Firmen verteilt über alle Entwicklung-phasen • Mehrere große Pharma-konzerne involviert	• Mehrere Firmen haben Substanzen unabhängig entwickelt • Mehrheit der Firmen hat eine Substanz in Entwicklung • Einzelne kleine Firmen haben größere Pro-gramme • Keine/kaum Lizenzaktivitäten	• Keine Firma hat mehr als eine Substanz in Entwicklung • Keine Lizenzaktivitäten • Keine großen Pharmakonzerne involviert

Abb. 6.12: Die Wettbewerbsintensität wurde mit zwei Kriterien bewertet
Die Abbildung zeigt die zwei Kriterien, die zur Bewertung der Wettbewerbsintensität verwendet wurden: Anzahl und Status der bereits existierenden Substanzen für das Zielmolekül und Wettbewerberstruktur. Für jedes Kriterium wurden drei Ausprägungen definiert, die mit einem Mondsymbol gekennzeichnet sind. Der leere Mond (○) symbolisiert eine Ausprägung, die für die indikationsübergreifende Funktion eher schlecht ist während der volle Mond (●) eine Ausprägung beschreibt, die vermuten lässt, dass indikationsübergreifende Funktion möglich ist. Durch Vergleich der Situation bei dem zu bewertenden Zielmolekül mit den definierten Ausprägungen, konnte dem Zielmolekül eine Bewertung zugeordnet werden. Die Übergänge zwischen den einzelnen Ausprägungen sind dabei fließend, d.h. es sind auch Einordnungen zwischen den einzelnen Ausprägungen denkbar.

6.6.4 Strategische Eignung

Die Bewertung der strategischen Eignung eines Zielmoleküls für Forschung und Entwicklung ist stets vom strategischen Kontext der jeweiligen Firma abhängig und kann daher nicht von außen erfolgen. Da es sich bei dieser Dimension trotzdem um einen sehr wichtigen Entscheidungsparameter handelt, wurden auch für diese Dimension Bewertungskriterien definiert.

- Erfahrung und Methodenkompetenz für Signalweg
- Erfahrung und Methodenkompetenz für Molekülklasse
- Existierende Pipeline
- Übereinstimmung mit F&E-Gesamtstrategie

Abb. 6.13 zeigt die detaillierte Beschreibung der Ausprägung der einzelnen Kriterien.

Ausprägung der Bewertungskriterien	○	◐	●
Erfahrung und Methoden-kompetenz für Signalweg	• Keine Erfahrung mit Signalweg • Notwendige Methoden (z.B. Tiermodelle, „Read-out"-Systeme) nur unzureichend vorhanden	• Wenig Erfahrung mit Signalweg, externe Partner oder interne Experten verfügbar • Benötigte Methoden begrenzt vorhanden, Investitionen zum Aufbau überschaubar	• Ausreichend eigene Erfahrung mit Signalweg, • Notwendige Methoden und Modelle vorhanden oder leicht und günstig verfügbar
Erfahrung und Methoden-kompetenz für Molekülklasse	• Keine Erfahrung mit Molekülklasse • Keinen externen Partner oder internen Wissensträger verfügbar	• Begrenzte Erfahrungen mit Molekülklasse • Externe Partner oder interne Wissensträger verfügbar	• Langjährige Erfahrung mit Molekülklasse • „Stand-alone"-Entwicklung möglich
Existierende Pipeline	• Keine aktiven Substanzen in der Pipeline, die für Molekül getestet werden könnten	• Keine aktiven Projekte, die für Molekül getestet werden könnten, jedoch Ideen und Konzepte vorhanden	• Substanzen in Pipeline vorhanden, die für Zielmolekül getestet werden könnten
Übereinstimmung mit F&E-Gesamtstrategie	• Aktivitäten lassen sich nicht in existierende Struktur eingliedern • Aktivitäten passen inhaltlich nicht zur F&E-Gesamtstrategie	• Integration in Struktur mit mäßigen Schwierigkeiten durchführbar • Aktivität von mäßigem Interesse für F&E-Gesamtstrategie	• Integration in bisherige Struktur ohne Schwierigkeiten möglich • Aktivität von zentraler Bedeutung für F&E-Gesamtstrategie

Abb. 6.13: Für die Bewertung der strategische Eignung wurden vier Kriterien entwickelt

Die Abbildung zeigt die zwei Kriterien, die zur Bewertung der strategischen Eignung dienen könnten: Erfahrung und Methodenkompetenz für Signalweg, Erfahrung und Methodenkompetenz für Molekülklasse, existierende Pipeline, Übereinstimmung mit F&E-Gesamtstrategie. Für jedes Kriterium wurden drei Ausprägungen definiert, die mit einem Mondsymbol gekennzeichnet sind. Der leere Mond (○) symbolisiert eine Ausprägung, die für die indikationsübergreifende Funktion eher schlecht ist während der volle Mond (●) eine Ausprägung beschreibt, die vermuten lässt, dass indikationsübergreifende Funktion möglich ist. Durch Vergleich der Situation bei dem zu bewertenden Zielmolekül mit den definierten Ausprägungen, konnte dem Zielmolekül eine Bewertung zugeordnet werden. Die Übergänge zwischen den einzelnen Ausprägungen sind dabei fließend, d.h. es sind auch Einordnungen zwischen den einzelnen Ausprägungen denkbar.

7 Ergebnisse

7.1 Entwicklung einer Methodik zur Identifizierung von indikationsübergreifenden Signalwegen

Ziel der Arbeit war es, eine Methodik zu entwickeln, mit deren Hilfe es möglich ist,

- Signalwege zu identifizieren, die bestimmte therapeutische Gebiete und Indikationsgruppen verbinden
- Moleküle zu identifizieren, die eine erhöhte Chance haben, für zwei ausgewählte Indikationsgruppen von Bedeutung zu sein.

Grundlegender Ansatz um diese Zielsetzungen zu erfüllen war die Konstruktion einer Datenbank, die durch Kombination von biologischen und pharmazeutischen Daten eine systematische Analyse der Vernetzung aller therapeutischen Gebiete erlaubt. Ausgangpunkt der Methodik ist die Tatsache, dass Moleküle, die durch Medikamente adressiert werden („Zielmoleküle“), keine isolierte Position im Organismus haben, sondern mit anderen Molekülen in Form von Signalwegen im funktionellen Zusammenhang stehen. Zur Erfüllung biologischer Funktionen ist in der Regel nicht nur ein Molekül wichtig, sondern die Wirkung vieler Moleküle in einem Signalweg. Das legt die Vermutung nahe, dass die Wirkung von Medikamenten häufig weniger auf der Inhibition eines einzelnen Moleküls als vielmehr auf dem Eingriff in einen Signalweg beruht. In Konsequenz bedeutet das, dass wenn ein Molekül innerhalb eines Signalwegs eine bestimmte therapeutische Funktion hat, auch andere Moleküle desselben Signalwegs eine erhöhte Chance haben, im diesem therapeutischen Zusammenhang eine Funktion haben. Neben den bekannten Verbindungen von zwei therapeutischen Gebieten über ein gemeinsames Zielmolekül (sogenannte direkte Verbindungen), sollen daher in der vorliegenden Arbeit auch Verbindungen von zwei therapeutischen Gebieten durch einen gemeinsamen Signalweg berücksichtigt werden. Diese

Signalwegverbindungen müssen nicht unbedingt ein gemeinsames Zielmolekül aufweisen und werden daher als indirekte Verbindungen bezeichnet. Für eine Verbindung reicht es aus, wenn ein Molekül aus dem Signalweg für das erste therapeutisches Gebiet wirksam ist und ein anderes aus demselben Signalweg für das zweite therapeutische Gebiet.

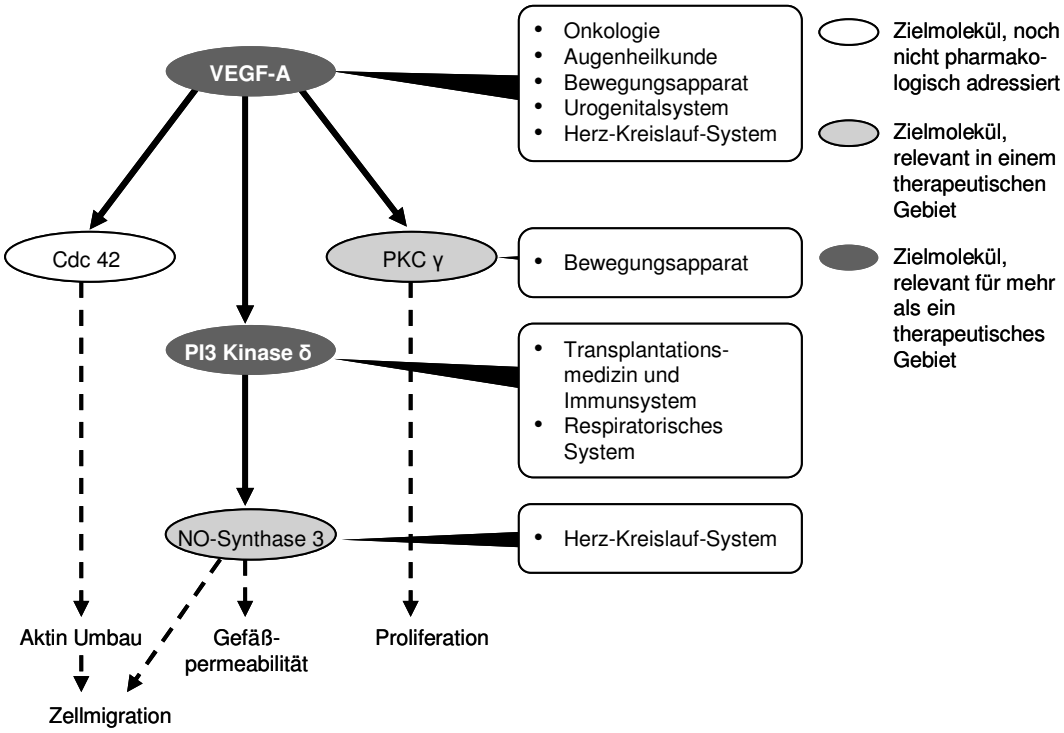


Abb. 7.1: Ausschnitt des VEGF-Signalwegs illustriert direkte und indirekte Verbindungen zwischen therapeutischen Gebieten und unterschiedliche Klassen von Zielmolekülen

Die Abbildung zeigt fünf Moleküle, die Teil des VEGF-Signalwegs sind sowie die therapeutischen Gebiete, in denen es Medikamente bzw. Substanzen in Forschung und Entwicklung gibt, die dieses Molekül adressieren.

Verbindungen: Therapeutische Gebiete, die sich ein Zielmolekül teilen, sind direkt miteinander verbunden. Zum Beispiel verbindet die PI3 Kinase δ direkt Transplantationsmedizin und Immunsystem mit dem respiratorischen System. Indirekte Verbindungen bestehen zwischen Gebieten, die sich kein Zielmolekül teilen, aber jeweils Zielmoleküle im selben Signalweg haben. So ist zum Beispiel das respiratorische System indirekt über den VEGF-Signalweg mit der Onkologie verbunden.

Zielmoleküle: Für die Moleküle VEGF-A und PI3 Kinase δ gibt es bereits Medikamente/Substanzen mit Indikationen in mehreren therapeutischen Gebieten, d.h. sie sind direkt indikationsübergreifend. Die PKCγ ist nur für ein therapeutisches Gebiet (Bewegungsapparat) bisher von Bedeutung. Da sie aber Teil des VEGF-Signalwegs ist, hat sie gute Chancen, auch für die anderen angrenzenden Gebiete in Zukunft von Bedeutung zu sein. Die PKCγ ist damit ein Kandidat für zukünftig indikationsübergreifende Zielmoleküle. Auch Cdc 42 könnte in Zukunft indikationsübergreifend sein, allerdings ist hier die Wahrscheinlichkeit geringer, da das Molekül noch gar nicht pharmakologisch adressiert wird.

Abb. 7.1 illustriert dies am Beispiel eines Ausschnitts aus dem Signalwegs „VEGF signaling pathway“. Die gezeigten Zielmoleküle dieses Signalwegs

können in unterschiedliche Klassen gefasst werden. VEGF-A und PI3 Kinase δ sind bereits heute für mehrere therapeutische Gebiete relevant und verbinden die entsprechenden Gebiete daher direkt als indikationsübergreifende Zielmoleküle. Andere Moleküle, wie zum Beispiel die PKC γ, werden pharmazeutisch erst für ein therapeutisches Gebiet genutzt. Die PKC γ ist daher noch kein indikationsübergreifendes Zielmolekül, verbindet aber verschiedene therapeutische Gebiete (zum Beispiel Onkologie und Bewegungsapparat) indirekt mit Hilfe des gemeinsamen VEGF-Signalwegs. Die dritte Klasse an Zielmolekülen (zum Beispiel Cdc 42) ist überhaupt noch nicht pharmazeutisch genutzt. Als funktioneller Bestandteil des VEGF-Signalwegs liegt aber die Vermutung nahe, dass auch Cdc 42 für ein therapeutisches Gebiet wirksam sein könnte, in dem der Signalweg eine Rolle spielt.

Um sich der Identifikation von möglichen Kandidaten für indikationsübergreifende Zielmoleküle zu nähern, wurden die Analysen der vorliegenden Arbeit in fünf Hauptschritte aufgeteilt, die von der Ebene des therapeutischen Gebiets, über die Ebene der Indikationsgruppen und der Signalwege bis auf die Detailebene der Zielmoleküle gehen.

1. Schritt: Analyse der Verbindung von therapeutischen Gebieten durch gemeinsame Signalwege. Im diesem Schritt wurde untersucht, welche therapeutischen Gebiete besonders eng miteinander über gemeinsame Signalwege verbunden sind und damit eine erhöhte Wahrscheinlichkeit haben, sich indikationsübergreifende Zielmoleküle zu teilen (siehe Abschnitt 7.2).

2. Schritt: Analyse der Verbindung von Indikationsgruppen durch gemeinsame Signalwege. Hier wurde ein starker Verbindungspunkt („Hot-Spot“) von therapeutischen Gebieten herausgegriffen und detailliert untersucht. Die Analyse beantwortet die Frage, welche Indikationsgruppen innerhalb ausgewählter therapeutischer Gebiete besonders stark verbunden sind (siehe Abschnitt 7.3).

3. Schritt. Identifikation und Priorisierung von Signalwegen in der Verbindung von therapeutischen Gebieten und Indikationsgruppen. Im Rahmen des dritten Schritts wurden die Signalwege, die zwei therapeutische Gebiete bzw. zwei

Indikationsgruppen verbinden, identifiziert und nach ihrer Bedeutung für die Verbindung in einer Rangliste sortiert (siehe Abschnitt 7.4).

4. Schritt. Identifikation von möglichen Zielmolekülkandidaten innerhalb einer Signalweg-Verbindung. Im diesem Schritt wurde ein Signalweg, der für die Verbindung zwischen zwei Indikationsgruppen wichtig ist, herausgegriffen und auf molekularer Ebene analysiert. Dabei wurden Zielmoleküle, die bereits heute indikationsübergreifend sind, Kandidaten für mögliche zukünftig indikationsübergreifende Zielmoleküle sowie ganz neue Zielmoleküle identifiziert (siehe Abschnitt 7.5).

5. Schritt. Priorisierung der Zielmolekülkandidaten. Die im vierten Schritt identifizierten Kandidaten für zukünftig indikationsübergreifende Zielmoleküle wurden entsprechend ihren Chancen für die indikationsübergreifende Funktion priorisiert (siehe Abschnitt 7.6.).

Alle Abbildungen des Ergebnisteils beruhen auf der in dieser Arbeit erstellten Datenbank und verwenden daher auch die Rohdaten von Pharmaprojects (Informa UK Ltd, 2008) und KEGG (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa Laboratories, 2008; Kanehisa u. a., 2008).

7.1.1 Grundstruktur der relationalen Datenbank

Die Grundstruktur der relationalen Datenbank deckt mit fünf Tabellen drei Bereiche ab:

- Pharmazeutische Forschungs- und Entwicklungspipeline: Pharmaprojects Informationen zu Projekten in Forschung und Entwicklung sowie zu Medikamenten, die bereits auf dem Markt sind (Informa UK Ltd, 2008)
- Biologische Signalwege: Zuordnung von Zielmolekülen zu funktionellen humanen Signalwegen basierend auf Daten der Kyoto Encyclopedia of Genes and Genomes (Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa Laboratories, 2008; Kanehisa u. a., 2008)

- Therapeutische Gebiete: Zuordnung der Indikationen von Medikamenten und Substanzen in Forschung und Entwicklung zu Indikationsgruppen und therapeutischen Gebieten

Grundprinzip der Datenbank sowie aller Analysen ist die Vernetzung der Daten zu einem integrierten Datensatz. Zu diesem Zweck wurden die Grundtabellen der Datenbank durch Beziehungen gleichartiger Felder zu einem integrierten Datensatz verbunden. Details hierzu sind im Methodenteil beschrieben. Abb. 7.2 zeigt schematisch die Vernetzung der Rohdaten zur relationalen Datenbank.

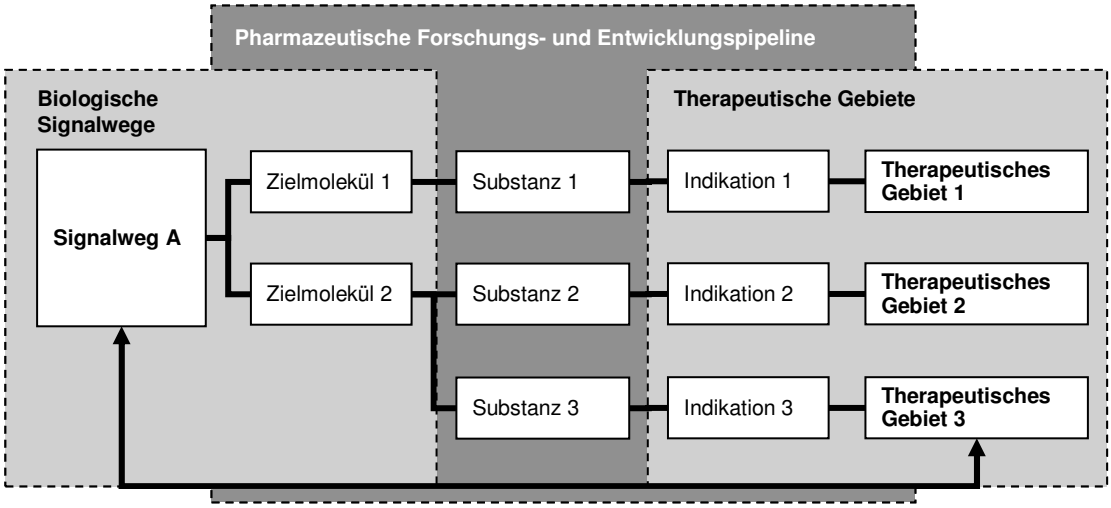


Abb. 7.2: Schematische Darstellung der Rohdatenvernetzung
Die Abbildung zeigt schematisch, wie die drei Rohdatenbereiche biologische Signalwege, pharmazeutische Forschungs- und Entwicklungspipeline und therapeutische Gebiete untereinander verbunden sind.
Durch Verbindung der Zielmoleküle zu Substanzen, die sie adressieren und den Indikationen der Substanzen, die Teil eines therapeutischen Gebiets sind, können indirekt Signalwege mit therapeutischen Gebieten verbunden werden. Dabei ist in der schematischen Darstellung das Zielmolekül 2 direkt indikationsübergreifend zwischen den Gebieten 2 und 3. Das Zielmolekül 1 hingegen verbindet die therapeutischen Gebiete 1 und 2 bzw. 1 und 3 nur indirekt über den Signalweg A.

Diese Vernetzung erlaubt einen Signalweg über die in ihm enthaltenen Zielmoleküle, über Substanzen, die diese Zielmoleküle adressieren und über die Indikationen, die diese Substanzen haben, mit therapeutischen Gebieten zu verbinden. In der schematischen Darstellung von Abb. 7.2 kann der Signalweg A als indirekte Verbindung zwischen den therapeutischen Gebieten 1 und 2 sowie 1 und 3 identifiziert werden. Im Gegensatz dazu sind die therapeutischen Gebiete 2 und 3 bereits auf Ebene der Zielmoleküle direkt über das indikationsübergreifende Zielmolekül 2 verbunden.

7.1.2 Allgemeine Statistik

Zur besseren Interpretation der Ergebnisse aus den fünf Schritten der Vernetzungsanalysen wurde zunächst die Grundstruktur der Daten deskriptiv analysiert.

7.1.2.1 Verteilung der Forschungsprojekte über therapeutische Gebiete

In der für die Datenbank verwendeten Datenbasis sind die Forschungsprojekte enthalten, von den Daten in dem verwendeten Auszug der Pharmaprojects Datenbank verfügbar waren. Dies entspricht einer guten Annährung an die Gesamtmenge aller öffentlich bekannten Forschungsprojekte (Informa UK Ltd, 2008). Aufgrund dessen spiegelt die Rohdatenbasis recht genau die tatsächlichen Größenverhältnisse der therapeutischen Gebiete wider.

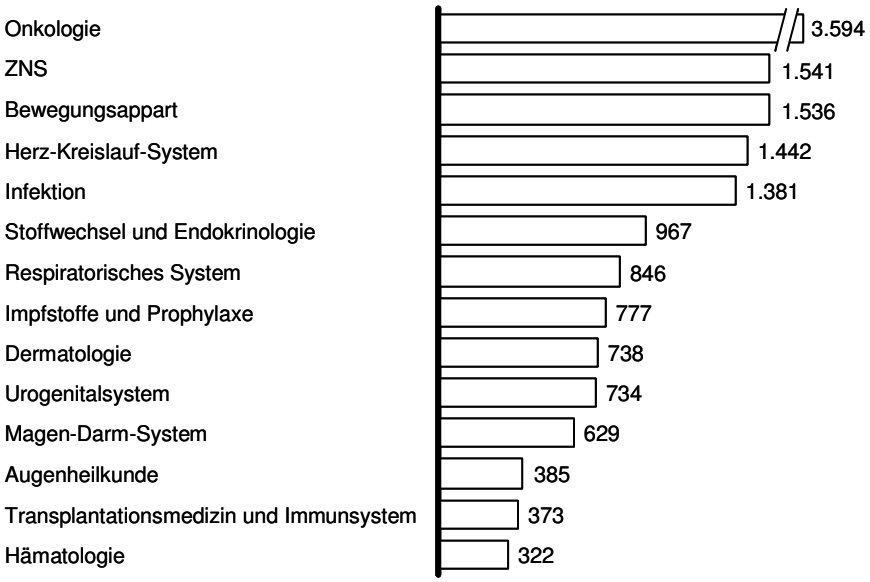


Abb. 7.3: Anzahl der Forschungsprojekte in der Datenbank nach therapeutischem Gebiet
Die Abbildung zeigt die Anzahl der Forschungsprojekte (d.h. unabhängige Kombinationen von Substanz und Indikation) in der Datenbank nach therapeutischem Gebiet. Die Gebiete „Other“ und „Ear and mouth“ wurden von der Analyse ausgeschlossen.
Da es sich bei den therapeutischen Gebieten nicht um künstlich definierte Einheiten sondern natürlich gewachsene Gebiete handelt, unterscheiden sie sich in ihre Größe deutlich. Während die Onkologie mit über 3500 Forschungsprojekten etwa 24 % aller Projekte beinhaltet, entfallen auf die kleinsten Gebiete mit 300 bis 400 Forschungsprojekten nur ca. 2 % der gesamten Projektzahl.

Die Größe der therapeutischen Gebiete variiert und ist neben der Anzahl an Indikationen und der Krankheitshäufigkeit (Prävalenz) und Anzahl der Neuerkrankungen (Inzidenz) der einzelnen Indikationen in diesem Gebiet auch

durch die wirtschaftliche Attraktivität des Gebiets für die pharmazeutischen Unternehmen bestimmt (Umsatzpotential, Erstattung von Medikamenten etc.).

Abb. 7.3 zeigt die Dominanz des therapeutischen Gebiets Onkologie auf dem pharmazeutischen Markt. 24 % aller Projekte gehören zu diesem therapeutischen Gebiet während die kleinsten Gebiete Transplantationsmedizin und Immunsystem sowie Hämatologie nur je 2 % der Projekte stellen.

7.1.2.2 Verteilung der Forschungsprojekte über Entwicklungsphasen

Die Forschungsprojekte der Datenbank verteilen sich auf unterschiedliche Entwicklungsphasen (siehe Abb. 7.4). Alle Substanzen, die keine aktiven Indikationen haben, wurden bereits vor der Konstruktion der Datenbank ausgeschlossen. Substanzen, die mehrere Indikationen in unterschiedlichen Gebieten haben, wurden mit jeder Indikation unabhängig als Forschungsprojekt erfasst und gehen daher pro Indikation einmal in Abb. 7.4 ein.

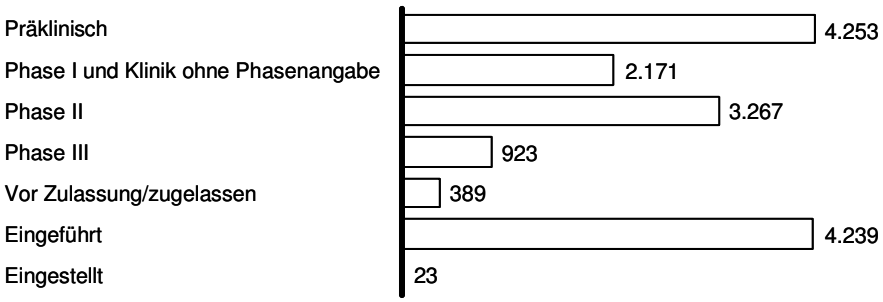


Abb. 7.4: Anzahl der Forschungsprojekte in der Datenbank nach Entwicklungsstatus
Die Abbildung zeigt die Anzahl der Forschungsprojekte (d.h. unabhängige Kombinationen von Substanz und Indikation) in der verwendeten Datenbank nach Entwicklungsstatus. Die Gebiete „Other“ und „Ear and mouth“ wurden von der Analyse ausgeschlossen. Substanzen, die keine einzige aktive Indikation haben, wurden ebenfalls nicht in den Datensatz aufgenommen. Projekte in der Klinik aber ohne Angabe der klinischen Phase wurden gemeinsam mit Projekten in Phase I erfasst.
Die Projekte verteilen sich auf alle Entwicklungsphasen. 23 Projekte sind eingestellt worden. Bei diesen Projekten handelt es sich um Substanzen, die in einer Indikation eingestellt wurden, aber in anderen Indikationen weiterhin aktiv sind und daher nicht vom Datensatz ausgeschlossen worden waren.

Abb. 7.4 macht deutlich, dass Projekte in allen Entwicklungsphasen im Datensatz enthalten sind. Etwa 4200 präklinischen Projekten steht eine etwa gleich große Zahl von Projekten gegenüber, die bereits auf dem Markt eingeführt wurden. Circa 6000 Projekte sind in den klinischen Entwicklungsphasen I-III. Eine sehr kleine Zahl von Forschungsprojekten (23) wurden eingestellt. Diese sind in der

Datenbank trotz vorheriger Bereinigung erhalten, weil es sich um vereinzelt beendete Indikationen von ansonsten noch aktiven Substanzen mit weiteren aktiven Indikationen handelt. Projekte, die sich in der klinischen Entwicklung befinden, jedoch nicht durch eine bestimmte Entwicklungsphase gekennzeichnet sind, wurde mit Projekten in Phase I zusammengefasst.

7.1.2.3 Forschungsaktivität nach therapeutischen Gebieten

Auch wenn die Projekte insgesamt über alle Entwicklungsphasen verteilt sind, gibt es deutliche Unterschiede in der Verteilung nach therapeutischen Gebieten (siehe Abb. 7.5). Während in den Gebieten Stoffwechsel und Endokrinologie sowie Infektion beispielsweise mehr als 50 % aller Projekte in der Präklinik und der ersten klinischen Phase sind, sind dies im Urogenitalsystem nur 24%. Ursachen für diese differenzierte Verteilung sind vermutlich nicht nur Unterschiede im aktuellen Forschungsgeschehen durch unterschiedliche Geschwindigkeit von Wissenszuwachs, sondern auch wirtschaftliche Gesichtspunkte, zum Beispiel das Umsatzpotential verschiedener therapeutischer Gebiete.

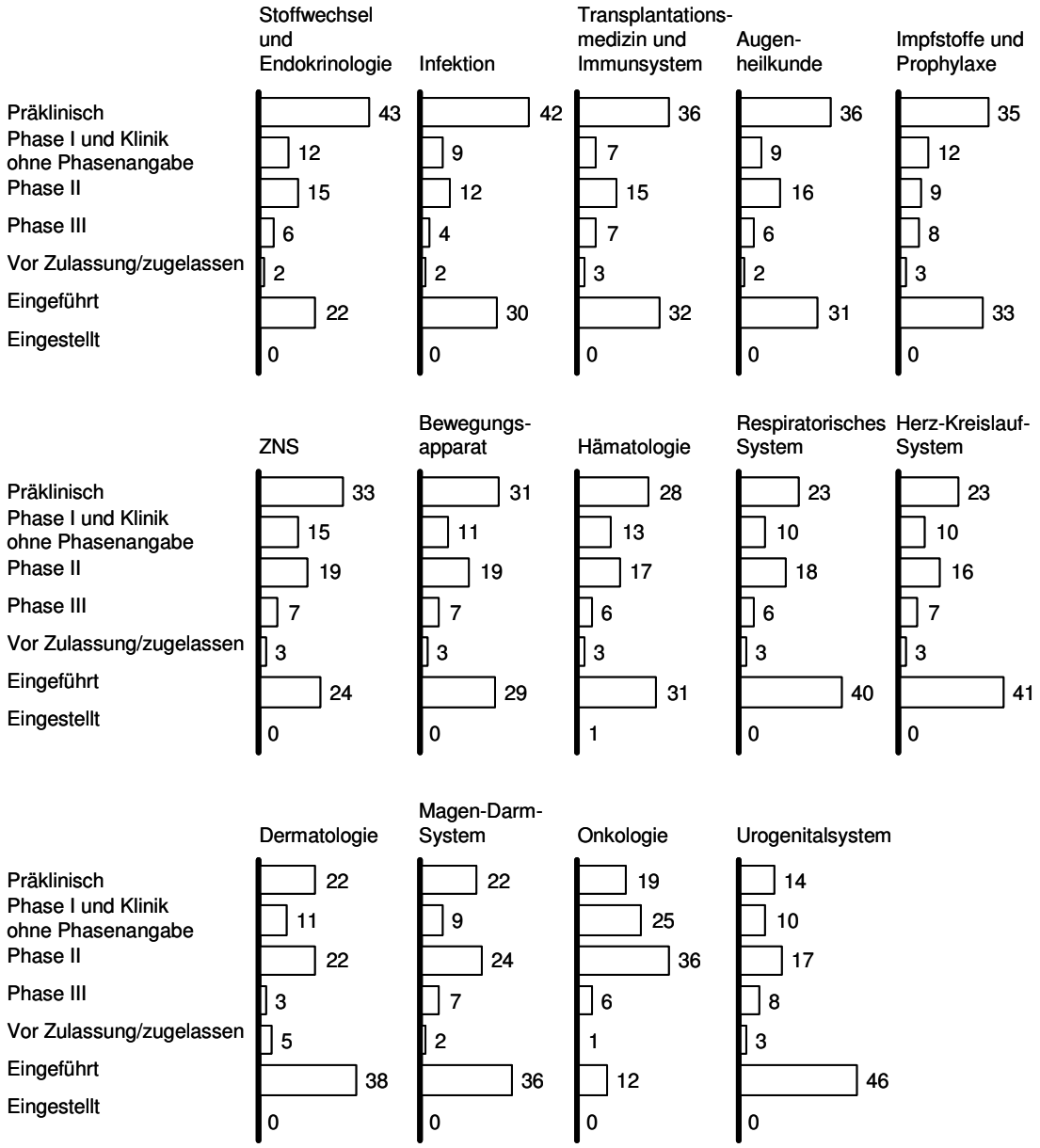


Abb. 7.5: Verteilung der Forschungsprojekte in der Datenbank pro Entwicklungsstatus nach therapeutischem Gebiet

Die Abbildung zeigt den Prozentsatz der Forschungsprojekte (d.h. unabhängige Kombinationen von Substanz und Indikation) in der Datenbank pro Entwicklungsstatus für alle therapeutischen Gebiete. Die Gebiete „Other“ und „Ear and mouth“ wurden von der Analyse ausgeschlossen. Substanzen, die keine einzige aktive Indikation haben, wurden nicht in den Datensatz aufgenommen. Projekte in der Klinik aber ohne Angabe der klinischen Phase wurden gemeinsam mit den Projekten in Phase I erfasst.

Die therapeutischen Gebiete unterscheiden sich deutlich in der Verteilung ihrer Forschungsprojekte auf unterschiedliche Entwicklungsphasen. Während beispielsweise das Gebiet Stoffwechsel und Endokrinologie 43 % seiner Projekte in der Präklinik hat, sind beim Urogenitalsystem nur 14 % der Projekte präklinisch, dafür aber 46 % aller Projekte bereits auf dem Markt eingeführt.

7.2 Analyse der Verbindungen von therapeutischen Gebieten durch gemeinsame Signalwege

7.2.1 Vernetzungsmatrix von therapeutischen Gebieten durch gemeinsame Signalwege

Mit Hilfe der Datenbank wurden die Verbindungen therapeutischer Gebiete untereinander über gemeinsame Signalwege untersucht. Dabei wurde die Anzahl der gemeinsamen Signalwege für jede Verbindung von zwei therapeutischen Gebieten bestimmt.

Abb. 7.6 zeigt das Ergebnis der Analyse in Form einer Vernetzungsmatrix. In den Feldern zwischen zwei therapeutischen Gebieten ist jeweils die Anzahl der verbindenden Signalwege aufgelistet. Insgesamt wurden 56 Signalwege in die Datenbank integriert, wobei nur 48 davon zur Vernetzung von therapeutischen Gebieten beitragen. Die restlichen Signalwege sind entweder überhaupt nicht pharmazeutisch adressiert oder nur in einem therapeutischen Gebiet von Bedeutung, so dass sie nicht als Verbindung gezählt werden können. Die Felder der Matrix sind entsprechend der Anzahl der verbindenden Signalwege eingefärbt, wobei die Intensität der Rottönung des Feldes die Stärke der Verbindung kennzeichnet. Je intensiver die Rottönung, desto mehr gemeinsame Signalwege teilen sich die beiden angrenzenden therapeutischen Gebiete. Insgesamt lässt sich eine hohe Anzahl von verbindenden Signalwegen zwischen fast allen therapeutischen Gebieten feststellen. 67 Verbindungsstellen sind durch mindestens die Hälfte aller analysierten Signalwege verbunden. Trotz der insgesamt hohen Anzahl von Verbindungen können auch Unterschiede in der Verbindungsdichte zwischen den einzelnen therapeutischen Gebieten festgestellt werden: Insbesondere die Verbindung zwischen dem therapeutischen Gebiet Bewegungsapparat und dem Gebiet Onkologie sticht hervor, da sich diese beiden Gebiete 44 Signalwege teilen. Dies ist in der Verbindungsmatrix der therapeutischen Gebiete die Kombination von Gebieten, die am engsten miteinander verbunden sind.

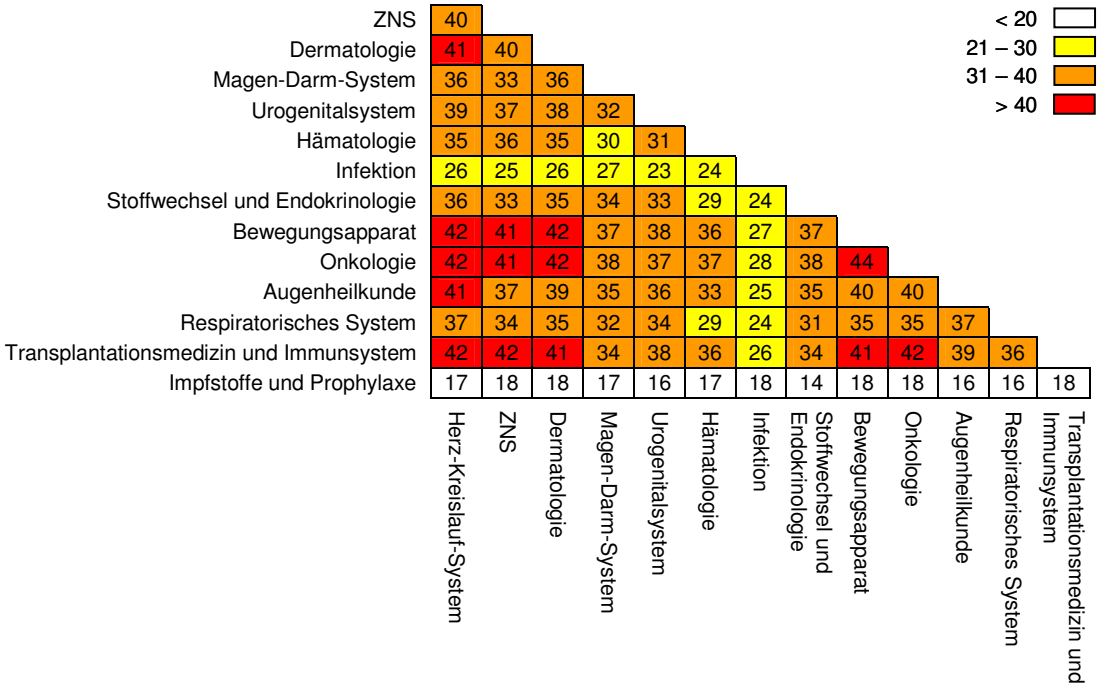


Abb. 7.6: Vernetzungsmatrix therapeutischer Gebiete: Anzahl gemeinsamer Signalwege
Die Abbildung stellt die Verbindung therapeutischer Gebiete untereinander durch gemeinsame Signalwege dar. In jedem Feld ist die Anzahl der Signalwege dargestellt, die sich die angrenzenden therapeutischen Gebiete teilen. Je intensiver die Rottönung des Feldes, desto mehr gemeinsame Signalwege liegen vor (vergleiche Legende oben rechts).
Von den 56 Signalwegen, die in der Datenbank enthalten sind, tragen 48 zur Vernetzung bei. Die engste Verbindung kann zwischen den Gebieten Onkologie und Bewegungsapparat beobachtet werden. Beide Gebiete teilen sich 44 Signalwege. Verbindungen mit dem Gebiet Impfstoffe und Prophylaxe zeichnen sich durch eine besonders geringe Anzahl an gemeinsamen Signalwegen aus.

Um eine bessere Aussage über die Verbindungsdichte der einzelnen therapeutischen Gebiete im Vergleich zueinander machen zu können, wurde der Parameter des Vernetzungsgrads eingeführt. Der Vernetzungsgrad beschreibt wie eng ein therapeutisches Gebiet mit allen anderen Gebieten vernetzt ist. Der absolute Vernetzungsgrad errechnet sich als Summe der Verbindungen eines therapeutischen Gebiets mit allen anderen. Der relative Vernetzungsgrad entspricht dem absoluten Vernetzungsgrad dividiert durch den höchsten Vernetzungsgrad im Vergleich der therapeutischen Gebiete und ist damit eine normierte Größe zwischen 0 und 1 (vergleiche Abschnitt 6.4.2). Dabei spielt es keine Rolle, ob der Vernetzungsgrad durch einzelne besonders eng verbundene Felder („Hot-spots“) oder durch eine gleichbleibend dichte Vernetzung aller Verbindungspunkte des Gebiets bestimmt wird. Unter Verlust dieser Detailinformation ermöglicht der Vernetzungsgrad, eine Aussage über die allgemeine Tendenz eines therapeutischen Gebiets, mit anderen Gebieten in

Zusammenhang zu stehen, zu machen. Abb. 7.7 zeigt den relativen Vernetzungsgrad aller untersuchten therapeutischen Gebiete. Dabei lässt sich bestätigen, was bereits Abb. 7.6 zeigte: Onkologie ist das am stärksten vernetzte Gebiet und teilt sich insgesamt am meisten Signalwege mit den anderen therapeutischen Gebieten. Neben der Onkologie sind auch der Bewegungsapparat und das Herz-Kreislauf-System sehr stark mit anderen Gebieten vernetzt. Auch diese Gebiete sind sehr groß und bieten damit vielfältige Möglichkeiten für Vernetzungen mit anderen Gebieten.

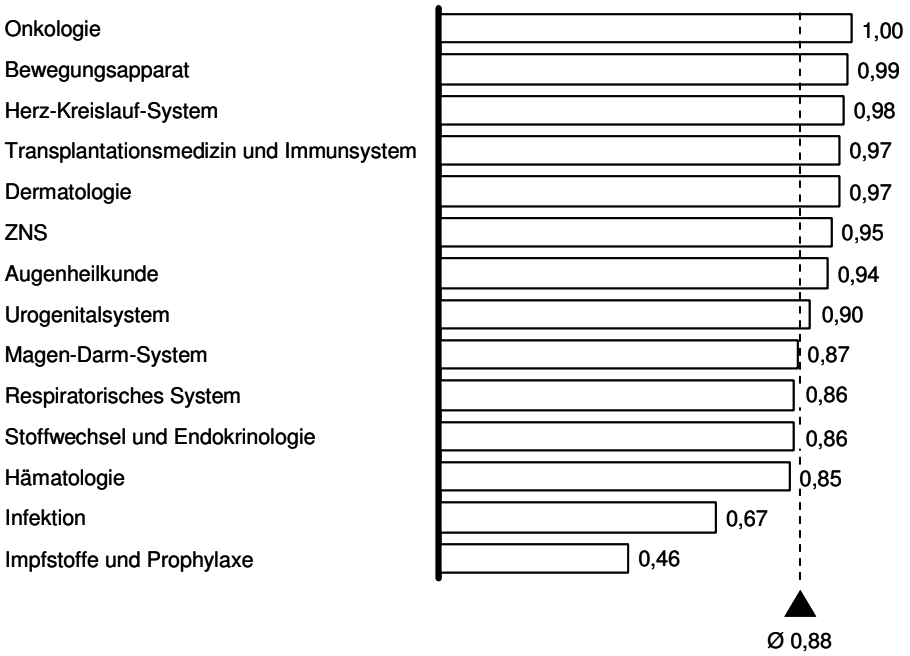


Abb. 7.7: Relativer Vernetzungsgrad der therapeutischen Gebiete
Die Abbildung zeigt den relativen Vernetzungsgrad pro therapeutischem Gebiet. Der relative Vernetzungsgrad ist ein Maß für die Tendenz eines Gebiets, sich Signalwege mit anderen Gebieten zu teilen und entspricht dem auf Werte zwischen 0 und 1 normierten absoluten Vernetzungsgrad (vergleiche Abschnitt 6.4.2).
Es fällt auf, dass beinahe alle Gebiete durch einen hohen Vernetzungsgrad nahe 0,9 charakterisiert sind. Dies bedeutet, dass alle Gebiete eine starke Tendenz haben, sich Signalwege mit anderen Gebieten zu teilen. Lediglich die Gebiete Infektion und Impfstoffe und Prophylaxe haben einen deutlich unter dem Mittelwert (0,88) liegenden Vernetzungsgrad.

Trotz dieser ersten Beobachtungen scheint die Größe jedoch nicht der ausschlaggebende Faktor für die Vernetzung zu sein, da zum Beispiel auch das sehr kleine Gebiet Transplantationsmedizin und Immunsystem durch eine intensive Vernetzung überrascht und beim Vernetzungsgrad auf dem vierten Rang liegt. Abb. 7.8 bestätigt diese Beobachtung und zeigt, dass es keine eindeutige

Korrelation zwischen der Größe eines therapeutischen Gebiets und seinem Vernetzungsgrad gibt.

Insgesamt ist auffällig, dass die meisten therapeutischen Gebiete eine starke Vernetzung zeigen. Dies spiegelt sich auch im Vernetzungsgrad (siehe Abb. 7.7) wider. Lediglich die Gebiete Infektion und Impfstoffe und Prophylaxe fallen durch deutlich geringere Vernetzung und einen Vernetzungsgrad deutlich unter dem Mittelwert auf. Der Vernetzungsgrad der übrigen Gebiete liegt im Bereich zwischen 0,85 und 1 und ist somit gleichbleibend hoch.

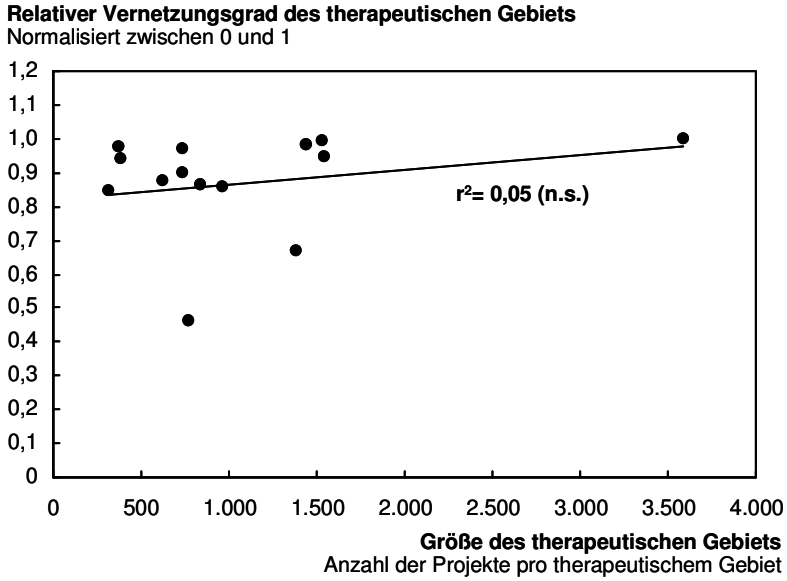


Abb. 7.8: Abhängigkeit des relativen Vernetzungsgrads von der Größe des therapeutischen Gebiets
Die Abbildung zeigt den Zusammenhang des relativen Vernetzungsgrads (y-Achse) von der Größe eines therapeutischen Gebiets (x-Achse, gemessen als Anzahl Projekte pro Gebiet).
Mit Hilfe von linearer Regressionsanalyse und dem Bestimmtheitskoeffizienten der linearen Regression (r^2=0,05, nicht signifikant mit P = 0,42 nach F-Test) kann gezeigt werden, dass kein eindeutiger Zusammenhang zwischen der Größe des Gebiets und seinem relativen Vernetzungsgrad besteht.

Mit Hilfe der Datenbank können für jede beliebige Verbindungsstelle von zwei therapeutischen Gebieten alle gemeinsamen Signalwege aufgelistet werden. Dies ist vor allem dann von Interesse, wenn eine Forschungsgruppe bereits über strategische Schwerpunkte in bestimmten therapeutischen Bereichen verfügt und für diese die Forschungsaktivitäten auf gemeinsame Signalwege ausrichten möchte. Als Beispiel für diese Auflistung zeigt Abb. 7.9 die verbindenden Signalwege zwischen den Gebieten Hämatologie und dem Magen-Darm-System. In diesem Beispiel sind 30 Signalwege für die Verbindung zwischen den beiden

therapeutischen Gebieten verantwortlich. Ohne eine detaillierte Gewichtung der einzelnen Signalwege kann jedoch keine Aussage darüber gemacht werden, welche Signalwege mehr und welche weniger wichtig für die Verbindung sind. Abb. 7.9 listet die Signalwege daher alphabetisch mit ihren englischen Originalbezeichnungen.

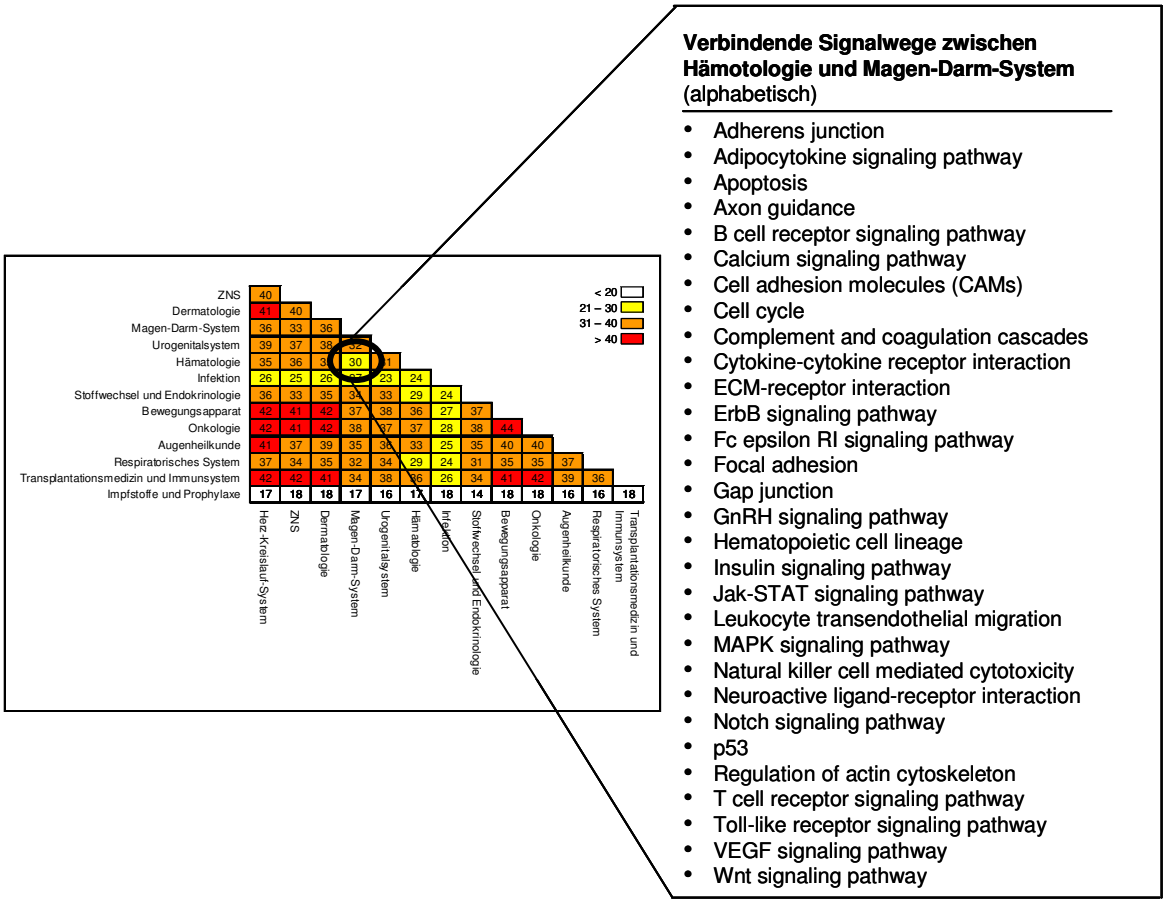


Abb. 7.9: Verbindende Signalwege für die therapeutischen Gebiete Hämatologie und Magen-Darm-System
Die Abbildung zeigt eine alphabetische Auflistung der Signalwege, die die therapeutischen Gebiete Hämatologie und Magen-Darm-System verbinden. Die Signalwege werden mit ihren englischen Originalbezeichnungen aufgelistet.

7.2.2 Vernetzungseigenschaften von Signalwegen und Zielmolekülen

Im Abschnitt 7.2.1 war vor allem die insgesamt hohe Vernetzung der meisten therapeutischen Gebiete durch gemeinsame Signalwege auffällig. Im Folgenden wurde daher untersucht, was die Ursachen hierfür sind und ob es Signalwege oder Zielmoleküle gibt, denen eine besondere Rolle in diesem Zusammenhang zukommt.

7.2.2.1 Vernetzungsfähigkeit eines Signalwegs in Abhängigkeit seiner Größe

Abb. 7.10 zeigt die 48 Signalwege, die zur Vernetzung der therapeutischen Gebiete beitragen. Für jeden Signalweg wurde die Anzahl der therapeutischen Gebiete und Indikationsgruppen angegeben, in denen der Signalweg eine Rolle spielt. Wie in Abb. 7.10 dargestellt ist, können die Signalwege entsprechend der Anzahl der durch sie verbundenen therapeutischen Gebiete sowie der Anzahl der verbundenen Indikationsgruppen in drei Gruppen eingeteilt werden. Die erste Gruppe der Signalwege (sogenannte „Generalisten“) verbindet alle untersuchten therapeutischen Gebiete. Auf Ebene der Indikationsgruppen lässt sich jedoch auch in der Gruppe der „Generalisten“ eine Differenzierung erkennen. Vor allem der Signalweg „Neuroactive-ligand-interaction“ dominiert die Vernetzung, in dem er in 73 der 77 untersuchten Indikationsgruppen eine Rolle spielt. Die große Bedeutung dieses Signalwegs in beinahe allen Indikationsgruppen ist nachvollziehbar, da es sich um eine besonders große funktionelle Gruppe von Zielmolekülen handelt. Dieser Signalweg enthält nicht nur sämtliche Neurotransmitter sondern auch deren Rezeptoren. Damit musste eine ubiquitäre Funktion des Signalwegs erwartet werden und kann als Kontrollkriterium für die Vernetzung dienen. Während zwölf Signalwege in die Gruppe der „Generalisten“ fallen, ist die Gruppe der „Hauptwege“, d.h. der Signalwege, die in mehr als zwei Drittel aller Gebiete eine Rolle spielen, deutlich größer. Insgesamt 25 Signalwege sind in zehn bis 13 Gebieten gleichzeitig von Bedeutung und tragen daher neben den „Generalisten“ maßgeblich zur hohen Verbindungsdichte bei.

	Signalweg	Anzahl der therapeutischen Gebiete, für die der Signalweg relevant ist	Anzahl der Indikations-gruppen, für die der Signalweg relevant ist
"Generalisten"– Signalwege, die alle therapeu-tischen Gebiete verbinden	• Neuroactive ligand-receptor interaction	14	73
	• Calcium signaling pathway	14	65
	• Cytokine-cytokine receptor interaction	14	64
	• MAPK signaling pathway	14	59
	• Jak-STAT signaling pathway	14	56
	• Toll-like receptor signaling pathway	14	52
	• Hematopoietic cell lineage	14	49
	• Adipocytokine signaling pathway	14	48
	• Apoptosis	14	47
	• Fc epsilon RI signaling pathway	14	45
	• T cell receptor signaling pathway	14	44
	• Cell cycle	14	33
“Hauptwege" – Signalwege, die mehr als 2/3 aller therapeu-tischen Gebiete verbinden	• Regulation of actin cytoskeleton	13	51
	• Natural killer cell mediated cytotoxicity	13	49
	• GnRH signaling pathway	13	47
	• VEGF signaling pathway	13	46
	• Leukocyte transendothelial migration	13	42
	• TGF-beta signaling pathway	13	39
	• Axon guidance	13	37
	• Complement and coagulation cascades	13	37
	• B cell receptor signaling pathway	13	34
	• Cell adhesion molecules (CAMs)	13	33
	• Gap junction	12	52
	• Focal adhesion	12	44
	• Insulin signaling pathway	12	35
	• ErbB signaling pathway	12	35
	• PPAR signaling pathway	12	32
	• Long-term depression	11	46
	• Long-term potentiation	11	44
	• Melanogenesis	11	32
	• Wnt signaling pathway	11	31
	• Notch signaling pathway	11	19
	• ECM-receptor interaction	10	31
	• Tight junction	10	26
	• mTOR signaling pathway	10	26
	• Phosphatidylinositol signaling system	10	23
	• p53	10	20
“Selektive Wege" – Signalwege, die weniger als 2/3 aller therapeu-tischen Gebiete verbinden	• Adherens junction	9	29
	• Renin-angiotensin system	9	16
	• Antigen processing and presentation	8	17
	• ABC transporters	8	14
	• Taste transduction	8	12
	• Hedgehog signaling pathway	8	11
	• Cell Communication	7	7
	• Dorso-ventral axis formation	6	13
	• Olfactory transduction	6	6
	• DNA polymerase	5	12
	• Regulation of autophagy	5	6

Abb. 7.10: Einteilung der Signalwege nach ihren Vernetzungseigenschaften
Die 48 Signalwege, die für die Vernetzung der therapeutischen Gebiete von Bedeutung sind, wurden nach Anzahl der durch sie verbundenen therapeutischen Gebiete in Gruppen eingeteilt. Die „Generalisten“ sind in allen 14 untersuchten Gebieten von Bedeutung. Die „Hauptwege“ verbinden zehn bis 13 therapeutische Gebiete und die „selektiven Wege“ sind in maximal neun Gebieten von Bedeutung. Neben der Anzahl der therapeutischen Gebiete ist auch die Zahl der Indikationsgruppen aufgelistet, in denen ein Signalweg eine Rolle spielt. Generell ist diese höher, wenn mehr therapeutische Gebiete adressiert werden.

Abb. 7.10 zeigt, dass es keine Signalwege gibt, die lediglich in einem einzelnen Gebiet eine Rolle spielen. Elf Signalwege sind allerdings in maximal neun Gebieten von Bedeutung. Damit sind diese Signalwege zwar nicht spezifisch für ein Gebiet, jedoch selektiver als die anderen und werden daher als „selektive Wege“ bezeichnet.

Die unterschiedliche Fähigkeit eines Signalwegs therapeutische Gebiete zu verbinden ist durch zwei wichtige Eigenschaften bedingt. Zum einen spielt die Größe der funktionellen Gruppe von Zielmolekülen eine starke Rolle, zum anderen ist auch die Eigenschaft der einzelnen Zielmoleküle eines Signalwegs von Bedeutung. Abb. 7.11 zeigt die Abhängigkeit der Vernetzungsfähigkeit eines Signalwegs von seiner Größe.

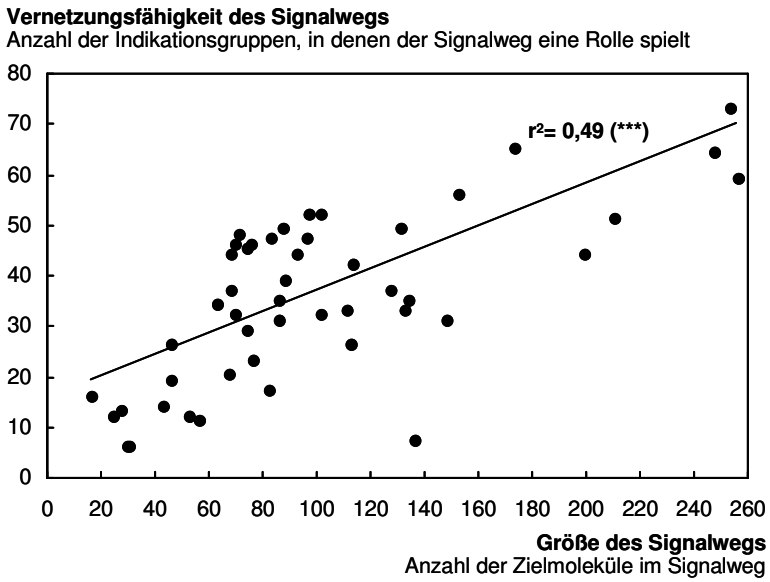


Abb. 7.11: Abhängigkeit der Vernetzungsfähigkeit eines Signalweg von seiner Größe
Die Abbildung zeigt den Zusammenhang des Vernetzungsfähigkeit eines Signalwegs (y-Achse, gemessen an der Anzahl der Indikationsgruppen, in denen der Signalweg eine Rolle spielt) von der Größe des Signalwegs (x-Achse, gemessen an der Anzahl der Zielmoleküle des Signalwegs).
Mit Hilfe von Regressionsanalyse wurde untersucht, in wie weit die Größe eines Signalwegs mit seiner Vernetzungsfähigkeit korreliert. Der Bestimmtheitskoeffizient der linearen Regression r^2 von 0,49 (hochsignifikant, $P < 0,0001$ nach F-Test) zeigt, dass nur ein mittlerer Zusammenhang zwischen der Größe des Signalwegs und seiner Vernetzung besteht.

Die Vernetzungsfähigkeit ist dabei durch die Anzahl der Indikationsgruppen, in denen der Signalweg eine Rolle spielt, definiert. Die Größe des Signalwegs wird durch die Anzahl seiner Zielmoleküle entsprechend der KEGG Rohdateninformation gemessen. Wie die Trendlinie und der

Bestimmtheitskoeffizient r^2 in Abb. 7.11 zeigen, gibt es nur einen mittleren Zusammenhang zwischen der Größe eines Signalwegs und seiner Vernetzungsfähigkeit.

Andere Parameter, in denen sich die untersuchten Signalwege unterscheiden, spielen offenbar für die Vernetzungsfähigkeit eines Signalwegs ebenfalls eine wichtige Rolle. Ein solcher Parameter ist zum Beispiel der Forschungs- und Nutzungsstatus eines Signalwegs. Damit ist gemeint, dass in manchen Signalwegen mehr Moleküle pharmazeutisch genutzt werden als das in anderen. Dieser aktive Anteil an Molekülen, also der Prozentsatz der Zielmoleküle innerhalb eines Signalwegs, der pharmazeutisch adressiert wird, hängt wiederum von ganz verschiedenen Faktoren ab.

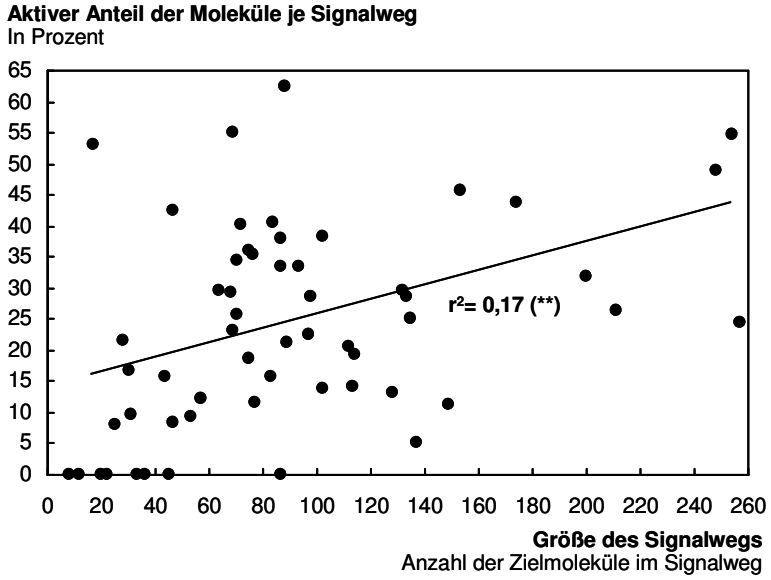


Abb. 7.12: Abhängigkeit des aktiven Anteils eines Signalwegs von seiner Größe
Die Abbildung zeigt den Zusammenhang zwischen dem aktiven Anteil eines Signalwegs, d.h. dem Prozentsatz an Molekülen im Signalweg, die pharmazeutisch genutzt werden (y-Achse), und der Gesamtgröße des Signalwegs (d.h. Anzahl der Moleküle entsprechend KEGG Rohdaten, x-Achse).
Die Trendlinie und der Bestimmtheitskoeffizient der linearen Regression r^2 von 0,17 (hochsignifikant, P = 0,0014 nach F-Test) zeigen, dass es keinen eindeutigen Zusammenhang zwischen der Größe des Signalwegs und seinem aktiven Anteil gibt.

Abb. 7.12 zeigt, dass ein Signalweg, der entsprechend KEGG viele Moleküle enthält, nicht automatisch auch durch einen hohen aktiven Anteil gekennzeichnet ist. Die Größe eines Signalwegs beeinflusst seinen aktiven Anteil also nicht unbedingt. Es wird vermutet, dass sowohl die Historie der biologischen und pharmazeutischen Forschung entlang eines Signalwegs als auch typische

Charakteristika der Zielmoleküle (zum Beispiel Rezeptoren versus intrazelluläre Moleküle) hierbei eine Rolle spielen.

Betrachtet man allerdings die aktive Größe eines Signalwegs, d.h. die absolute Anzahl von pharmazeutisch adressierten Zielmolekülen, so lässt sich eine deutliche Abhängigkeit der Vernetzungsfähigkeit eines Signalwegs von seiner aktiven Größe erkennen (siehe Abb. 7.13).

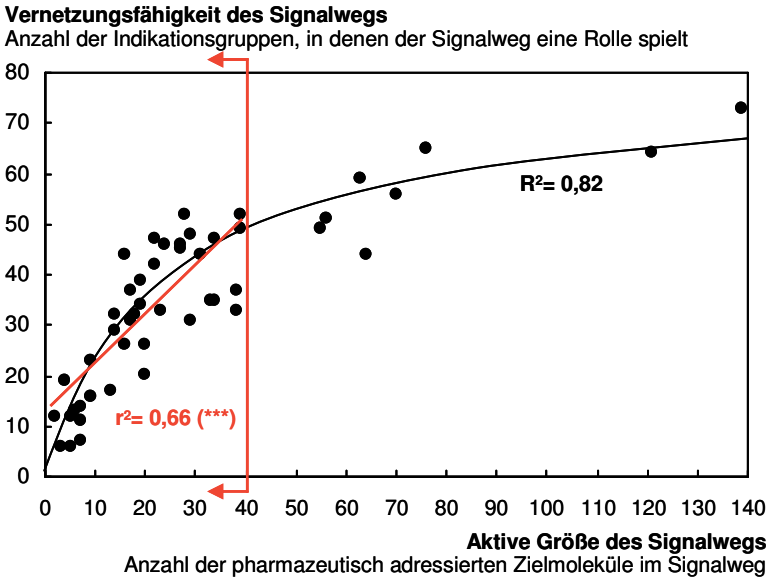


Abb. 7.13: Abhängigkeit der Vernetzungsfähigkeit eines Signalwegs von seiner aktiven Größe

Die Abbildung zeigt den Zusammenhang zwischen der Vernetzungseigenschaft eines Signalwegs (y-Achse, gemessen an der Anzahl der Indikationsgruppen, in denen er eine Rolle spielt) und seiner aktiven Größe (x-Achse, gemessen an der Anzahl pharmazeutisch adressierter Zielmoleküle im Signalweg.)
Die Trendlinie der nichtlinearen Regression (schwarze Hyperbel) und deren Bestimmtheitskoeffizient R^2 von 0,82 zeigen, dass es einen Zusammenhang zwischen der aktiven Größe und den Vernetzungseigenschaften gibt (Hyperbelfunktion, Abweichung der Daten vom nicht-linearen Modell nach Runs-Test nicht signifikant).
Ab einer aktiven Größe von ca. 40 Zielmolekülen, steigt die Vernetzungsfähigkeit nicht mehr so stark an wie bei einer geringeren aktiven Größe. Eine lineare Regression (rote Gerade) unter einem Grenzwert von 40 aktiven Molekülen (r^2 = 0,66, hochsignifikant, $p < 0{,}0001$ nach F-Test) zeigt, dass aber auch in diesem Bereich die Vernetzungsfähigkeit des Signalwegs nicht ausschließlich durch seine aktive Größe bestimmt wird.

Dies entspricht den Erwartungen, denn je mehr pharmazeutisch adressierte Zielmoleküle ein Signalweg hat, desto größer ist auch seine Chance viele Indikationsgruppen zu vernetzen. In Abb. 7.13 kann darüber hinaus erkannt werden, dass eine Sättigung eintritt, wenn eine kritische Anzahl von etwa 40 aktiven Zielmolekülen überschritten ist. In anderen Worten: Hat ein Signalweg mehr als 40 Zielmoleküle, die pharmazeutisch genutzt werden, ist die

Wahrscheinlichkeit sehr hoch, dass er in nahezu allen Indikationsgruppen eine Rolle spielt. Dies ist vor allem für die Signalwege „Neuroactive-ligand interaction“ und „Cytokine-cytokine signaling“ der Fall, deren aktive Größe mit 139 bzw. 121 pharmazeutisch adressierten Zielmolekülen mit Abstand am größten ist. Die Tatsache, dass diese beiden Signalwege in Abb. 7.9 unter den „Generalisten“ aufgeführt sind, ist also neben der Tatsache, dass sie verbreitete Funktionen des Organismus beschreiben auch davon bestimmt, dass sie sehr viele pharmazeutisch aktive Zielmoleküle haben. Eine lineare Regression bei Beschränkung auf kleinere Signalwege mit weniger als 40 aktiven Molekülen und einem Bestimmtheitsfaktor r^2 von 0,66 (siehe Abb. 7.13, rote Gerade) zeigt jedoch erneut, dass es neben der aktiven Größe noch einige andere Faktoren geben muss, die die Vernetzungseigenschaften eines Signalwegs bestimmen.

7.2.2.2 Vernetzungseigenschaften von Zielmolekülen

Neben der aktiven Größe des Signalwegs spielen auch die Eigenschaften der enthaltenen Zielmoleküle eine wichtige Rolle für die Vernetzungsfähigkeit eines Signalwegs. Während die meisten Zielmoleküle spezifisch nur in einzelnen Signalwegen vorkommen, gibt es eine kleine Gruppe von Molekülen, die in mehr als 10 Signalwegen vorkommt (siehe Abb. 7.14).

Diese Netzknoten-Zielmoleküle sind durch ihr häufiges Vorkommen in verschiedenen Signalwegen eine wichtige Ursache der hohen Vernetzungsdichte. Offenbar reicht bereits die geringe Anzahl von 13 Zielmolekülen in dieser Klasse aus, die starke Vernetzung der therapeutischen Gebiete anzutreiben. Bei den Netzknoten-Zielmolekülen handelt es sich vor allem um Moleküle, die an ubiquitär im Organismus vorkommenden Prozessen beteiligt sind (v.a. Proteinkinasen). Neben den 13 Netzknoten-Zielmolekülen gibt es eine Gruppe von knapp 40 Molekülen, die durch Teilhabe an 5 – 10 Signalwegen ebenfalls stark zur allgemeinen Vernetzung beitragen. Die große Masse der Zielmoleküle (> 90 %) gehören jedoch nur zu einem oder sehr wenigen Signalwegen.

	Anzahl der Signalwege, in denen das Zielmolekül vorkommt	Anzahl der Zielmoleküle in dieser Klasse	**Beispiele** (Anzahl der Signalwege in denen das Zielmolekül vorkommt)
Netzknoten Zielmoleküle	> 10	13	• PKC beta1/alpha/gamma (17/16/14) • MAPKK1 (15) • PI3-kinase alpha (15)
Stark vernetzende Zielmoleküle	5 - 10	38	• EGFR (10) • Jun oncogene (8) • Interleukin 1 beta (5)
Leicht vernetzende Zielmoleküle	2 - 4	278	• Fibronectin 1 (4) • Leptin receptor (4) • Adenosine A2a receptor (2)
Signalweg-spezifische Zielmoleküle	1	264	• Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1 • Coagulation factor X • Neuropeptide Y

Abb. 7.14: Klassifizierung der Zielmoleküle nach der Anzahl der Signalwege, in denen sie vorkommen
Eine kleine Gruppe von 13 Zielmolekülen kommt in mehr als zehn Signalwegen vor und ist damit ein starker Treiber für die Vernetzung. Zu diesen sogenannten Netzknoten-Zielmolekülen gehören vor allem ubiquitäre Proteinkinasen. 38 weitere Moleküle sind durch Teilhabe an fünf bis zehn Signalwegen ebenfalls sehr stark vernetzend wirksam. Die große Mehrheit der Zielmoleküle ist jedoch nur einem oder einigen wenigen Signalwegen von Bedeutung.

7.2.3 Gewichtung

7.2.3.1 Logik der Gewichtung

Die im Kapitel 7.2.1 beschriebene Vernetzung basiert auf der absoluten Anzahl der verbindenden Signalwege zwischen den therapeutischen Gebieten. Dabei wird nicht berücksichtigt, ob ein Signalweg zwei Gebiete mehr oder weniger stark miteinander verbindet. Um eine genauere Information über die Stärke der Vernetzung zu bekommen, wurde die Gewichtung eingeführt. Dabei werden nicht mehr nur alle Signalwegverbindungen zwischen zwei therapeutischen Gebieten addiert, sondern jedem Signalweg entsprechend seiner Verbindungsstärke für die Kombination der therapeutischen Gebiete ein spezifisches Gewicht gegeben. Wie in Abb. 7.15 gezeigt ist, setzt sich in Folge dessen die Verbindungsstärke zwischen zwei therapeutischen Gebieten aus der Summe der gewichteten Verbindungsstärken der einzelnen verbindenden Signalwege zusammen.

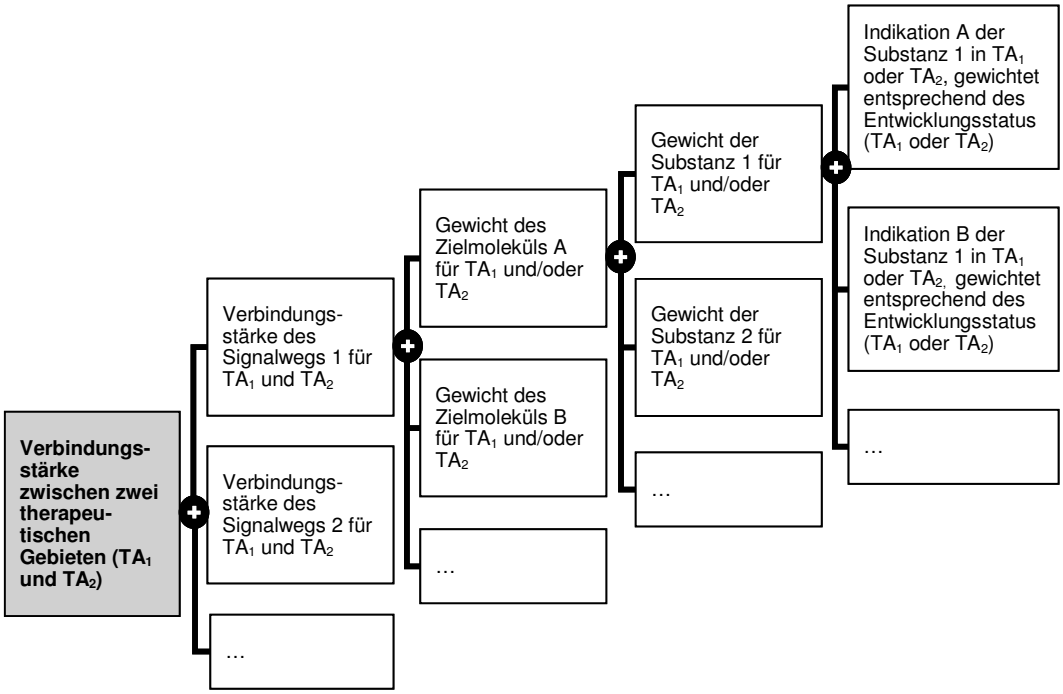


Abb. 7.15: Rechenlogik der Gewichtung ohne Durchschnittsbildung
Die Abbildung zeigt, wie sich bei Gewichtung ohne Durchschnittsbildung die Verbindungsstärke zwischen zwei therapeutischen Gebieten berechnet.
Die Verbindungsstärke zwischen zwei therapeutischen Gebieten („therapeutic area" TA_1 und TA_2) entspricht der Summe der Verbindungsstärken aller verbindenden Signalwege (1. Ebene von links). Die Verbindungsstärke eines Signalwegs wird als Summe der Gewichte derjenigen Zielmoleküle bestimmt, die zum Signalweg gehören und über Substanzen mit Indikationen in den Gebieten TA_1 und/oder TA_2 verfügen (2. Ebene von links). Das Gewicht eines Zielmoleküls bestimmt sich als Summe der gewichteten Substanzen mit Indikationen in den Gebieten TA_1 und/oder TA_2 (3. Ebene von links). Das Gewicht einer Substanz bestimmt sich als Summe der entsprechend ihres Entwicklungsstatus mit einem Gewichtungsfaktor gewichteten Indikationen. Dabei werden die Indikationen für beide Gebiete (TA_1 und TA_2) berücksichtigt (4. Ebene von links).

Bei der Berechnung der Verbindungsstärke eines einzelnen Signalwegs werden folgende Parameter berücksichtigt:

- Anzahl der Zielmoleküle innerhalb des Signalwegs, die für die Verbindung relevant sind
- Gewicht der einzelnen Zielmoleküle für die Verbindung

Generell gilt, dass ein Zielmolekül, das bereits heute indikationsübergreifend für beide therapeutischen Gebiete von pharmazeutischer Bedeutung ist, bei der Gewichtung stärker berücksichtigt wird als ein Zielmolekül, das bisher nur für ein Gebiet relevant ist. (vergleiche Abschnitt 6.4.5). Bei der Berechnung des Gewichts eines Zielmoleküls für ein therapeutisches Gebiet werden folgende Parameter berücksichtigt:

- Anzahl der Substanzen für dieses Zielmolekül mit Indikationen innerhalb dieses Gebiets bzw. dieser Gebiete
- Anzahl und Entwicklungsstatus der einzelnen Indikationen für die Substanzen

Im Rahmen der hier beschriebenen einfachen Gewichtung (auch „Gewichtung ohne Durchschnittsbildung“) wird jede Indikation innerhalb des therapeutischen Gebiets entsprechend ihres Entwicklungsstatus mit einem Gewichtungsfaktor versehen. Dieser Gewichtungsfaktor richtet sich nach dem Entwicklungsstatus der Substanz in der entsprechenden Indikation und ist bei der Gewichtung zunächst umso höher, je weiter die Substanz für diese Indikation die klinische Entwicklung bereits durchlaufen hat. Das Gewicht der Substanz für ein therapeutisches Gebiet errechnet sich durch die Summe der mit dem Gewichtungsfaktor versehenen Indikationen innerhalb eines therapeutischen Gebiets.

Im Ergebnis führt die Gewichtung dazu, dass jeder Signalweg für eine Verbindung zwischen zwei therapeutischen Gebieten eine unterschiedliche Verbindungsstärke hat. Genauso hat jedes Zielmolekül je nach Kombination von therapeutischen Gebieten ein unterschiedliches Gewicht für die Verbindung. Je mehr unterschiedliche Substanzen mit Indikationen in einem oder beiden therapeutischen Gebieten für das Zielmolekül entwickelt werden und je weiter die einzelnen Indikationen in der Entwicklung fortgeschritten sind, desto höher ist das Gewicht des Zielmoleküls. In anderen Worten: Ein Signalweg hat eine hohe Verbindungsstärke für zwei Gebiete, wenn die Verbindung durch zahlreiche Zielmoleküle gestützt wird und jedes dieser Zielmoleküle durch viele Substanzen mit vielen, weit fortgeschrittenen Indikationen in den entsprechenden Gebieten adressiert wird.

7.2.3.2 Variationen der Gewichtungslogik

Wie im vorhergehenden Abschnitt beschrieben wurde, werden bei der Gewichtung diejenigen Zielmoleküle stärker gewichtet, die durch eine große Anzahl von Substanzen mit möglichst zahlreichen weit fortgeschrittener Indikationen gestützt werden. Spezifische Eigenschaften von therapeutischen

Gebieten (vergleiche Abschnitt 7.1.2) könnten dabei aber gegebenenfalls eine Verzerrung der Gewichtung hervorrufen.

Um dies auszuschließen, wurden Untersuchungen zu den folgenden beiden Fragen durchgeführt:

- Gibt es gebietsspezifische Variationen der Anzahl der Indikationen pro Substanz?
- Gibt es gebietsspezifische Variationen der Anzahl der Substanzen pro Zielmolekül?

Abb. 7.16 und Abb. 7.17 zeigen die jeweiligen Ergebnisse.

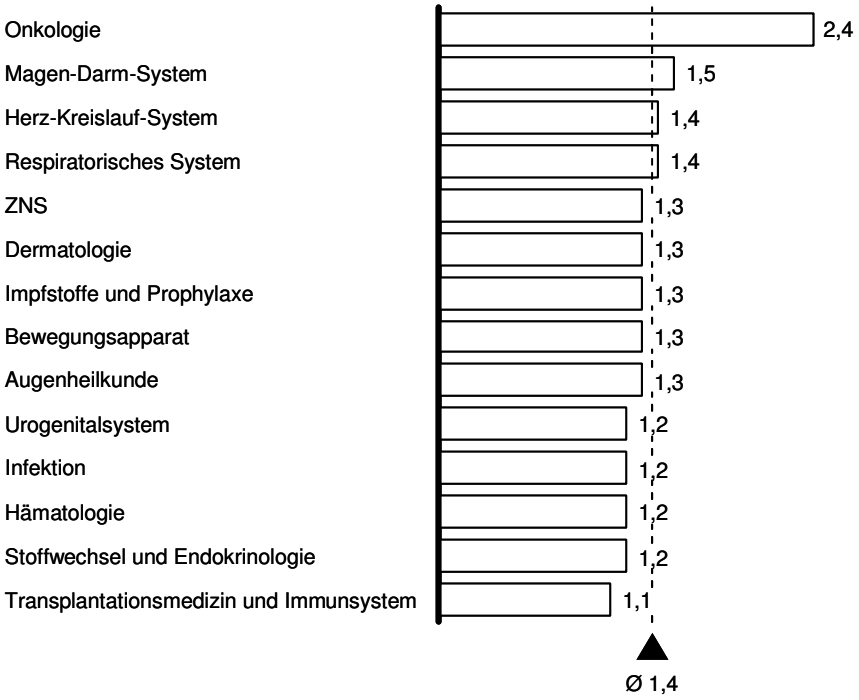


Abb. 7.16: Durchschnittliche Anzahl von Indikationen pro Substanz nach therapeutischen Gebieten
Die Abbildung zeigt die durchschnittliche Anzahl von Indikationen für ein therapeutisches Gebiet (zum Beispiel Infektion) pro Substanz.
Es fällt auf, dass fast alle Gebiete ca. 1,4 Indikationen pro Substanz haben. Substanzen im Bereich Onkologie verfügen jedoch mit 2,4 über eine deutlich höhere durchschnittliche Anzahl von Indikationen pro Substanz.

Abb. 7.16 zeigt, dass in den meisten therapeutischen Gebieten ca. 1,4 Indikationen pro Substanz vorliegen. Einzige Ausnahme ist die Onkologie, in der im Durchschnitt beinahe doppelt so viele Indikationen pro Substanz zu finden sind. Eine mögliche Erklärung für dieses Phänomen kann in den charakteristischen Eigenschaften des Gebiets gefunden werden: Während die meisten anderen

therapeutischen Gebiete auf Organe bzw. Organsysteme beschränkt sind, erstreckt sich die Onkologie auf organsystem-übergreifende Krankheitsmechanismen (zum Beispiel Zellzyklus-Deregulation), die für viele verschiedene Krebsformen ähnlich sind. Dies erleichtert die Erweiterung des Indikationsspektrums auf mehrere Krebsformen. Dazu kommt, dass die Nebenwirkungstoleranz durch den Leidensdruck einer Krebserkrankung verhältnismäßig hoch ist, was eine Anwendung in verschiedenen Indikationen erlauben kann, auch wenn dies mit erheblichen Nebenwirkungen verbunden ist.

Im Rahmen der Gewichtung ohne Durchschnittsbildung wurde jede Indikation einzeln zur Berechnung des Gewichts eines Zielmoleküls berücksichtigt. Im Anbetracht von Abb. 7.16 führt dies jedoch zu einer überproportionalen Gewichtung aller Verbindungen mit der Onkologie und sollte daher vermieden werden. Zu diesem Zweck wurde eine zweite Gewichtungsform eingeführt, die als „Gewichtung mit einfacher Durchschnittsbildung“ bezeichnet wird. Wie in Abb. 7.18 deutlich wird, wird bei dieser Gewichtungsform nicht die Summe aller nach ihrem Entwicklungsstatus gewichteten Indikationen gebildet, sondern ein Durchschnitt aller gewichteten Indikationen pro Substanz und therapeutischem Gebiet. Dies ermöglicht eine generelle Aussage darüber zu machen, wie weit fortgeschritten die Entwicklung der Indikationen innerhalb des therapeutischen Gebiets für diese Substanz ist.

Abb. 7.17 zeigt die durchschnittliche Anzahl von Substanzen pro Zielmolekül. Dabei gilt, dass nur chemisch verschiedene Substanzen gezählt werden, jedoch Substanzen mit gleichem Wirkmechanismus, die chemisch verschieden sind, individuell berücksichtigt werden. Die zu beobachtende Schwankung kann viele Ursachen haben. Neben einem unterschiedlichen Wissensstand von Forschung und Entwicklung, ist vermutlich auch der Einfluss der kommerziellen Attraktivität des Gebiets inklusive gesetzlicher Verordnungen (zum Beispiel Zulassungsbeschränkungen und Wiedererstattungsregelungen) von Bedeutung.

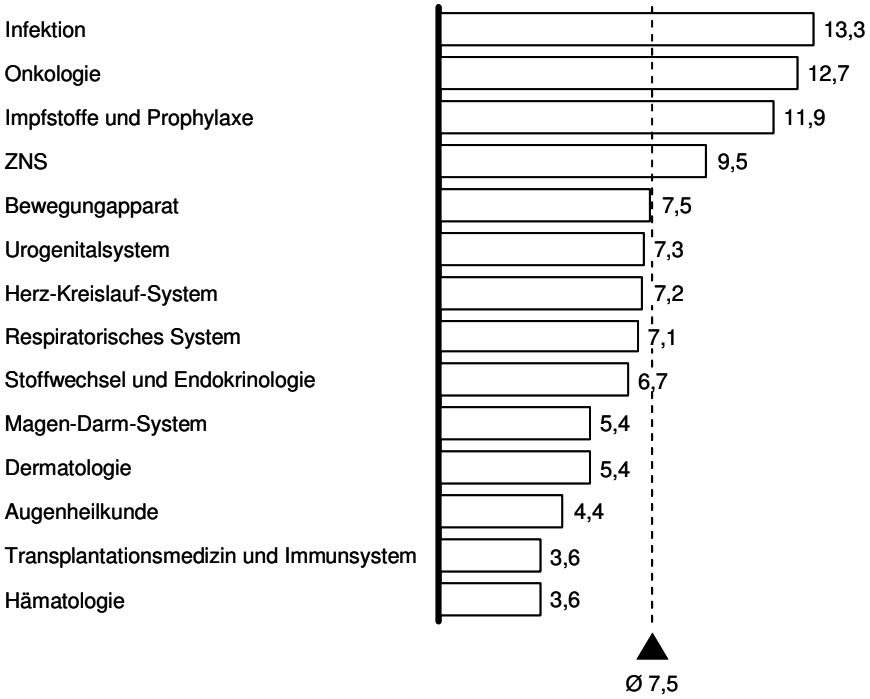


Abb. 7.17: Durchschnittliche Anzahl von Substanzen pro Zielmolekül nach therapeutischen Gebieten
Die Abbildung zeigt die durchschnittliche Anzahl von Substanzen pro Zielmoleküle in jedem therapeutischen Gebiet. Bei einem Durchschnitt von 7,5 Substanzen pro Zielmolekül lassen sich starke Schwankungen der Anzahl von 3,6 bis 13,3 feststellen.

Ob eine „Korrektur“ um diese Variation vergleichbar zu der oben beschriebenen Durchschnittsbildung durchgeführt werden sollte, hängt von der jeweiligen Fragestellung der Analyse ab. Grundsätzlich belegt eine hohe Anzahl von Substanzen pro Zielmolekül das Vertrauen der Pharmaforschung in die Bedeutung des Zielmoleküls und kann daher als Validitätskriterium verwendet werden. Eine Anpassung der Gewichtung sollte allerdings immer dann vorgenommen werden, wenn es verstärkt um den Aspekt der biologischen Verbindung geht und kommerzielle Effekte bereinigt werden sollen. Die Anpassung der Gewichtung erfolgt im diesem Fall durch eine zweite Durchschnittsbildung („Gewichtung mit doppelter Durchschnittbildung“), bei der ein durchschnittliches Gewicht pro Zielmolekül aus den einzelnen gewichteten Substanzen berechnet wird. Abb. 7.18 zeigt vergleichend die drei oben beschriebenen Gewichtungsvariationen.

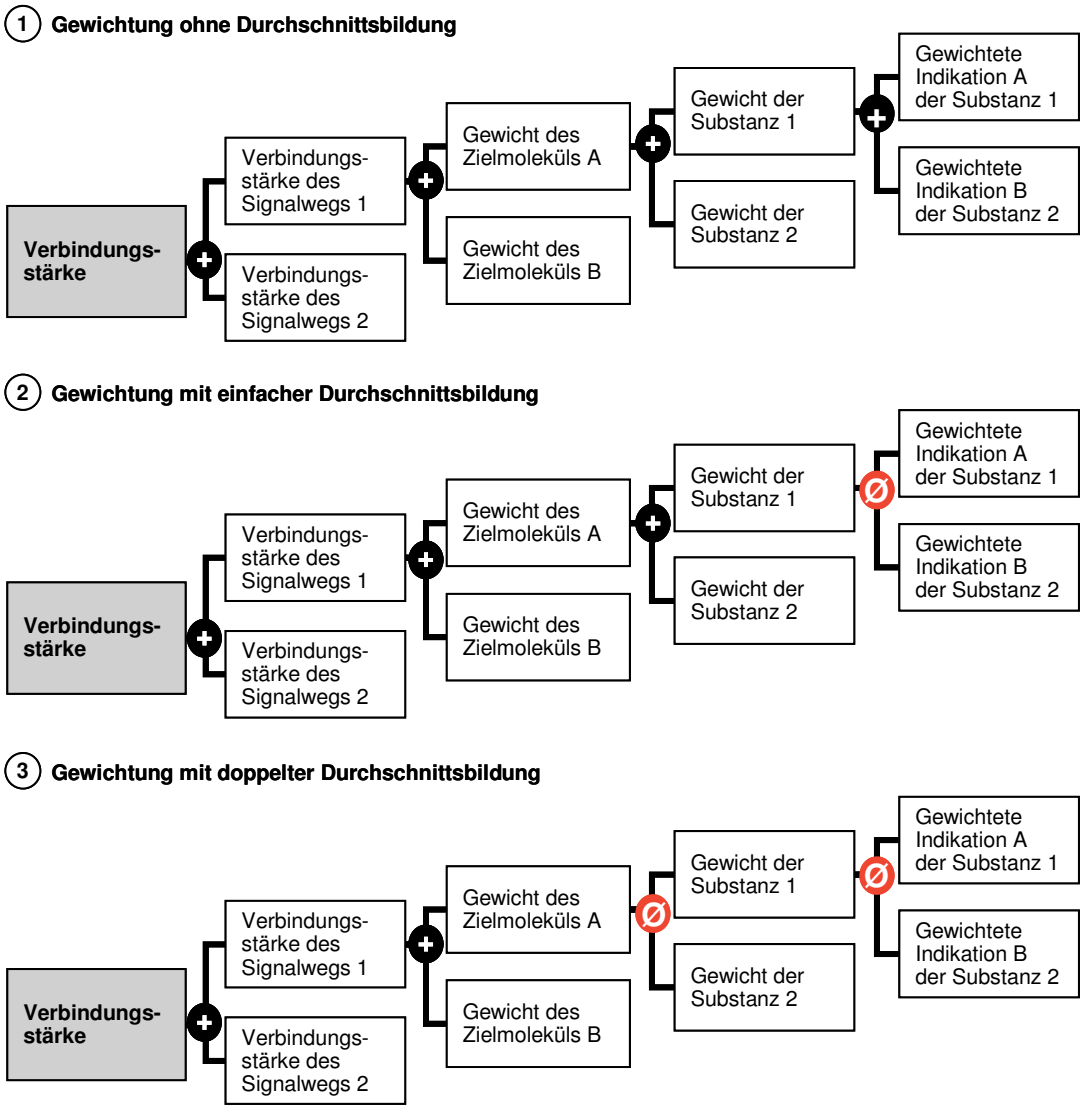


Abb. 7.18: Variationen der Gewichtungslogik
Die Abbildung zeigt die drei möglichen Variationen der Gewichtung ohne Durchschnittsbildung, mit einfacher und mit doppelter Durchschnittsbildung.
Die obere Abbildung (1) zeigt die Situation ohne Durchschnitt. Die Berechnung des Gewichts des verbindenden Signalwegs erfolgt durch vierfache Summenbildung. Dies entspricht der in Abb. 7.15 erläuterten Gewichtungsberechnung.
Die Gewichtung mit einfacher Durchschnittsbildung (2), errechnet das Gewicht pro Substanz nicht als Summe sondern als Durchschnitt der gewichteten Indikationen der Substanz. Alle weiteren Rechenschritte bleiben Summenbildungen.
Die Gewichtung mit doppelter Durchschnittsbildung (3) verwendet das durchschnittliche Gewicht der Substanz und berechnet daraus ein durchschnittliches Gewicht pro Zielmolekül. Die Gewichte der Zielmoleküle werden wie bisher zur Verbindungsstärke des Signalwegs addiert; die Summe der Verbindungsstärken aller verbindenden Signalwege entspricht der gesamten Verbindungsstärke.

Zielmolekülen, die bereits heute für beide untersuchten Gebiete relevant sind, werden im Gegensatz zu denen, die nur in einem Gebiet wichtig sind, bei der Gewichtung stärker berücksichtigt. Abb. 7.19 zeigt eine schematische Berechnung

mit doppelter Durchschnittsbildung. Während das Zielmolekül 1 für die therapeutischen Gebiete 1 und 2 („therapeutic area“, TA_1 und TA_2) von Bedeutung ist, gibt es für das Zielmolekül 2 nur Indikationen innerhalb des therapeutischen Gebiets 2 (TA_2). Im Rahmen der ersten Durchschnittsbildung wird der durchschnittliche Entwicklungsstatus für jedes Gebiet separat bestimmt. Im zweiten Schritt wird dann – ebenfalls für jedes Gebiet separat – der Durchschnitt der gewichteten Substanzen pro Zielmolekül gebildet. Ist ein Molekül für beide Gebiete vertreten, werden beide Durchschnitte addiert (vergleiche Zielmolekül 1 im Beispiel). Dies ist sinnvoll, um zu berücksichtigen, dass dieses Zielmolekül einen besonders großen Anteil an der Verbindung der beiden Gebiete hat.

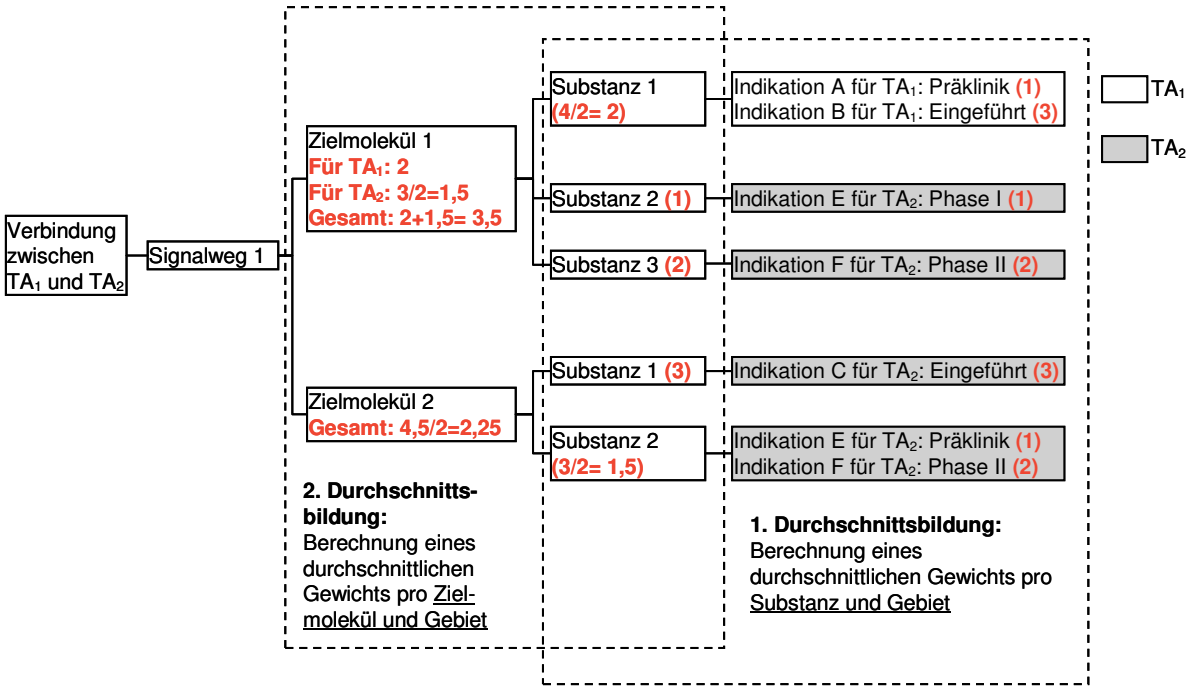


Abb. 7.19: Beispiel der Berechnung des Gewichts mit doppelter Durchschnittsbildung und Berücksichtigung direkt indikationsübergreifender Zielmoleküle
Die Abbildung zeigt eine beispielhafte Berechnung des Gewichts von zwei Zielmolekülen (Zielmolekül 1 und 2) aus dem Signalweg 1 für die Verbindung der therapeutischen Gebiete 1 und 2 („therapeutic area“ TA_1 und TA_2). Die Berechnungswerte sind in roter Schrift vermerkt.
Mit der ersten Durchschnittsbildung wird ein durchschnittliches Gewicht pro Substanz berechnet. Dies erfolgt für jedes therapeutische Gebiet getrennt. Bei der zweiten Durchschnittsbildung wird das durchschnittliche Gewicht pro Zielmolekül berechnet. Auch hier bleibt die Trennung pro therapeutisches Gebiet erhalten, so dass das Zielmolekül 1, welches im Beispiel für beide Gebiete relevant ist, zwei Durchschnittswerte erhält. Beide Durchschnittswerte werden dann addiert, so dass das Zielmolekül 1 ein höheres Gewicht als das Zielmolekül 2, welches nur für ein Gebiet genutzt wird, erhält.

7.2.3.3 Gewichtete Vernetzungsmatrizen

Die folgenden Abbildungen zeigen die Veränderung der Vernetzungsmatrix der therapeutischen Gebiete durch Gewichtung. Abb. 7.20 zeigt die Matrix mit Gewichtung ohne Durchschnittsbildung. Wie nach der Analyse aus Abb. 7.16 zu erwarten war, sind alle Verbindungen mit Onkologie überproportional stark betont. Diese Gewichtung eignet sich daher nicht für weitere Analysen.

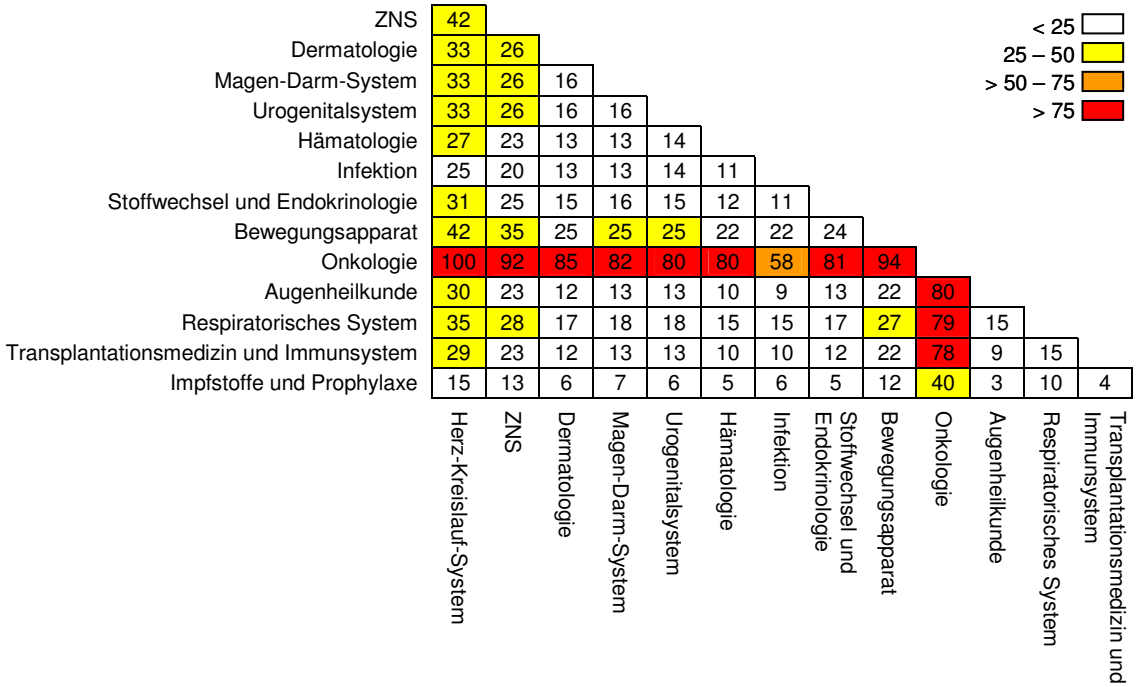


Abb. 7.20: Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung ohne Durchschnittsbildung
Die Abbildung zeigt die Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung ohne Durchschnittsbildung. Die Verbindungsstärke der Verbindung zwischen zwei therapeutischen Gebieten ist in den Feldern der Matrix als normierter Verbindungswert zwischen 0 und 100 eingetragen. Das bedeutet, dass die stärkste Verbindung den Wert 100 hat und die anderen in Verhältnis dazu eingestuft werden. Die Rottönung der Matrix ist an die Normierung angepasst (vergleiche Legende oben rechts in der Abbildung).
Es wird deutlich, dass vor allem Verbindungen mit dem therapeutischen Gebiet Onkologie stark hervorgehoben sind. Dies liegt vor allem an der hohen Anzahl an Indikationen pro Substanz (vergleiche Abb. 7.16). Aufgrund dieses Effekts ist die Gewichtung ohne Durchschnittsbildung für weitere Analysen nicht geeignet.

Sowohl einfache (Abb. 7.21) als auch doppelte Durchschnittsbildung (Abb. 7.22) erlauben eine Differenzierung zwischen der Vernetzung der therapeutischen Gebiete. Im Vergleich lässt sich folgendes beobachten:

- Unabhängig von der Gewichtungsart ist das therapeutische Gebiet Onkologie am stärksten mit anderen therapeutischen Bereichen verbunden. Dabei fallen die Verbindungen der Onkologie zum Herz-Kreislauf-System

sowie zum Bewegungsapparat als stärkste Verbindungspunkte („Hot-Spots“) der Matrix auf.

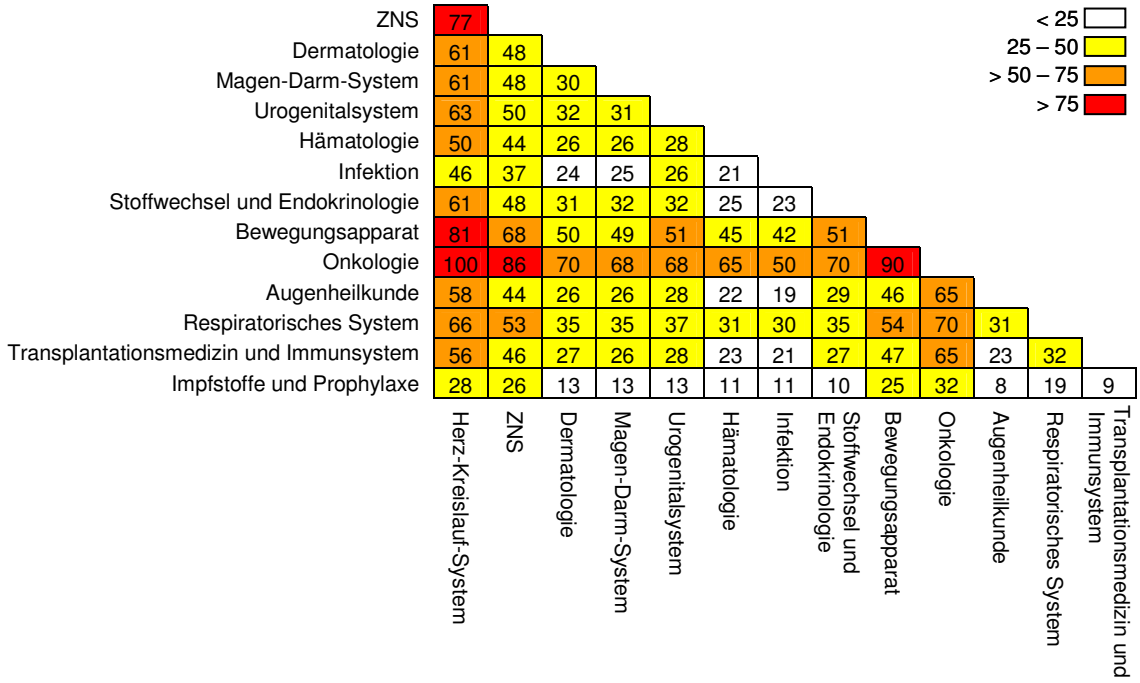


Abb. 7.21: Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit einfacher Durchschnittsbildung
Die Abbildung zeigt die Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit einfacher Durchschnittsbildung. Die Verbindungsstärke der Verbindung zwischen zwei therapeutischen Gebieten ist in den Feldern der Matrix als normierter Verbindungswert zwischen 0 und 100 eingetragen. Das bedeutet, dass die stärkste Verbindung den Wert 100 hat und die anderen in Verhältnis dazu eingestuft werden. Die Rottönung der Matrix ist an die Normierung angepasst (vergleiche Legende oben rechts in der Abbildung).
Die Kombination von therapeutischen Gebieten, die den höchsten Verbindungswert hat („Hot Spot“ der Matrix), ist die Verbindung des Herz-Kreislauf-Systems mit dem Gebiet Onkologie. Dies sind auch diejenigen therapeutischen Gebiete, deren Verbindungen insgesamt verhältnismäßig stark sind. Wie auch vor der Gewichtung, ist das Gebiet Impfstoffe und Prophylaxe durch schwache Verbindungen zu anderen therapeutischen Gebieten gekennzeichnet.

- Das therapeutische Gebiet Transplantationsmedizin und Immunsystem hat nach der Gewichtung im Vergleich zur Analyse ohne Gewicht (Abb. 7.6) weniger Verbindungen zu den anderen Gebieten. Offenbar verbinden zwar zahlreiche Signalwege das therapeutische Gebiet mit anderen, jedoch werden die Verbindungen bisher durch weniger Zielmoleküle und weniger Substanzen gestützt, als das in anderen Gebieten der Fall ist.

- Neben der Onkologie sind vor allem die therapeutischen Gebiete Bewegungsapparat und Herz-Kreislauf-System durch zahlreiche Verbindungen zu anderen Gebieten gekennzeichnet. Durch Vergleich der

Gewichtung mit einfacher und mit doppelter Durchschnittsbildung lässt sich feststellen, dass Verbindungen mit dem Herz-Kreislauf-System offenbar durch mehr Substanzen gestützt sind als im Bereich Bewegungsapparat, d.h. vor allem bei einfacher Durchschnittbildung einen hohen Verbindungswert erhalten.

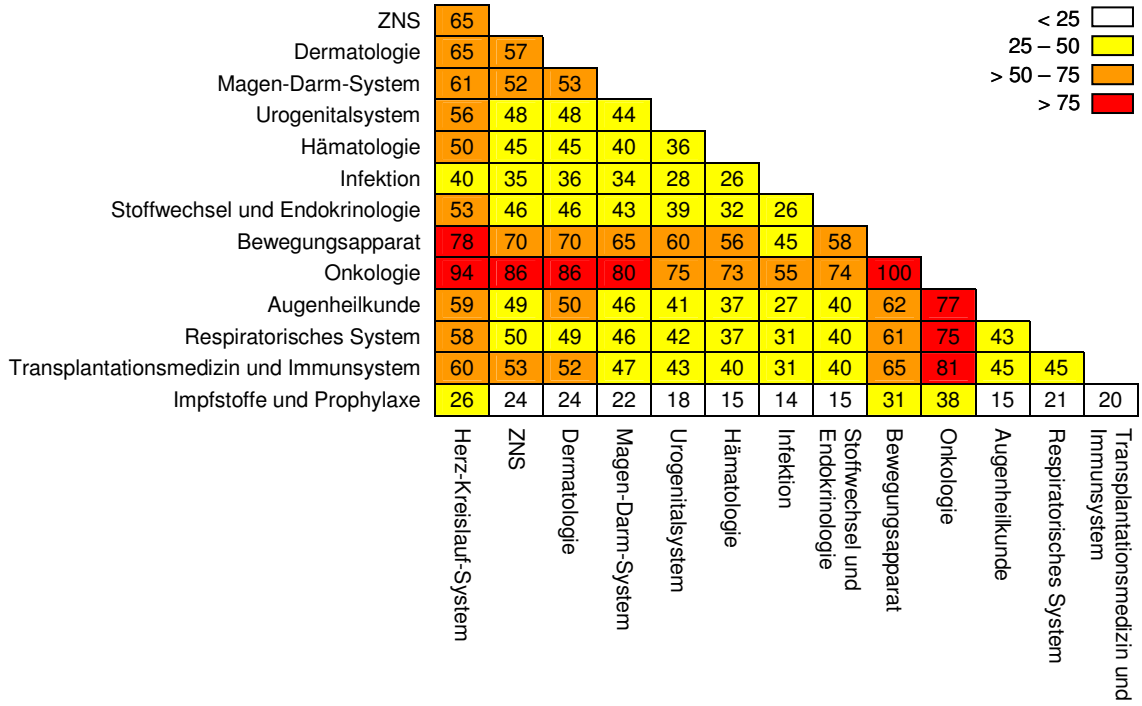


Abb. 7.22: Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit doppelter Durchschnittsbildung

Die Abbildung zeigt die Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit doppelter Durchschnittsbildung. Die Verbindungsstärke der Verbindung zwischen zwei therapeutischen Gebieten ist in den Feldern der Matrix als normierter Verbindungswert zwischen 0 und 100 eingetragen. Das bedeutet, dass die stärkste Verbindung den Wert 100 hat und die anderen in Verhältnis dazu eingestuft werden. Die Rottönung der Matrix ist an die Normierung angepasst (vergleiche Legende oben rechts in der Abbildung).
Der Verbindungspunkt mit dem höchsten Verbindungswert („Hot Spot“) liegt zwischen den therapeutischen Gebieten Onkologie und Bewegungsapparat. Im Vergleich zur Matrix nach einfacher Durchschnittsbildung (vergleiche Abb. 7.21) wird unter anderem deutlich, dass Verbindungen mit dem Herz-Kreislauf-System weniger hohe Verbindungswerte erhalten. Im Gegensatz dazu sind die Verbindungen mit dem therapeutischen Gebiet Bewegungsapparat überwiegend durch hohe Verbindungswerte gekennzeichnet. Die Onkologie bleibt das insgesamt am stärksten vernetzte therapeutische Gebiet.

- Die Vernetzung des therapeutischen Gebiets Impfstoffe und Prophylaxe verändert sich durch Gewichtung kaum. Dies entspricht den zuvor diskutieren besonderen Eigenschaften des Gebiets und bestätigt, dass durch Gewichtung keine artifiziellen Veränderungen der Analyse entstehen.

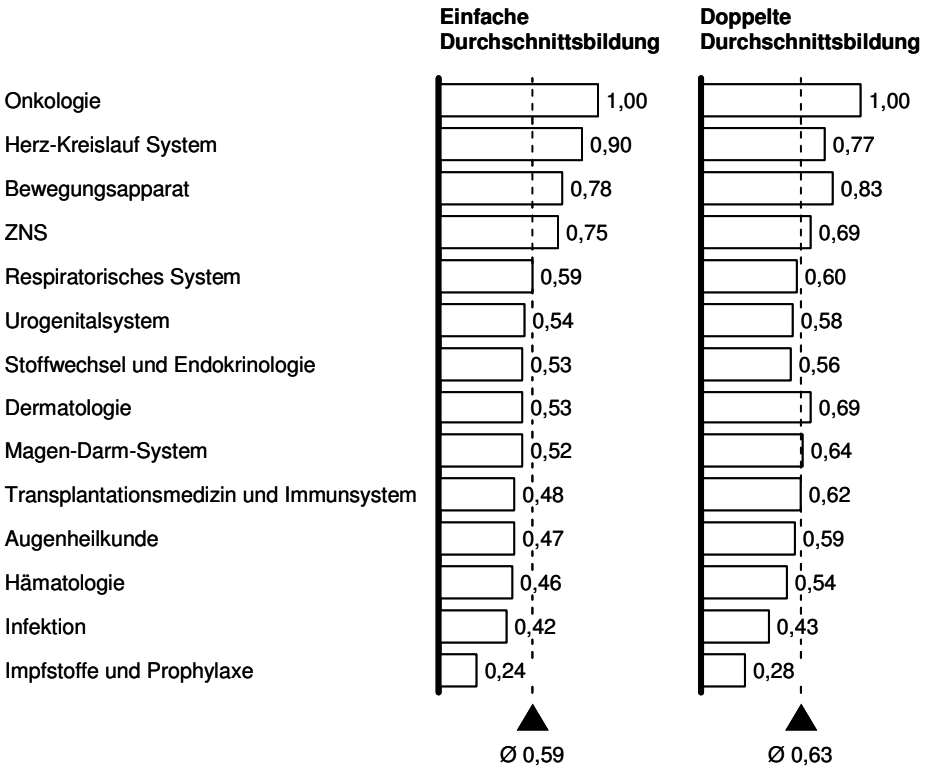


Abb. 7.23: Relativer Vernetzungsgrad der therapeutischen Gebiete nach Gewichtung
Die Abbildung zeigt den relativen Vernetzungsgrad (vergleiche Abschnitt 6.4.2) der therapeutischen Gebiete nach Gewichtung mit einfacher und doppelter Durchschnittsbildung.
Die Gewichtung erlaubt eine deutliche Streuung der Vernetzungsgradwerte und damit eine verbesserte Differenzierung zwischen den einzelnen therapeutischen Gebieten als das ohne Gewichtung der Fall war (vergleiche Abb. 7.7). Einfache und doppelte Durchschnittbildung führen zu Unterschieden im Vernetzungsgrad, die vor allem durch die stärkere Betonung der Anzahl der Substanzen pro Zielmolekül bei der einfachen Durchschnittsbildung hervorgerufen werden.

Der Vergleich des relativen Vernetzungsgrads nach Gewichtung (Abb. 7.23) zeigt im Vergleich zum relativen Vernetzungsgrad vor Gewichtung (Abb. 7.7) eine stärkere Streuung und damit auch eine bessere Aufteilung des Vernetzungsstärken durch Gewichtung. Deutliche Unterschiede der Vernetzung lassen sich neben den oben bereits beschriebenen Effekten von Herz-Kreislauf-System, Bewegungsapparat und Transplantationsmedizin und Immunsystem auch bei den Gebieten Zentralnervensystem (ZNS), Dermatologie, Magen-Darm-System, Augenheilkunde und Hämatologie erkennen. Während das ZNS bei einfacher Gewichtung höhere Verbindungswerte bekommt und daher offenbar durch zahlreiche Substanzen gestützt ist, beobachtet man den umgekehrten Effekt in den anderen Gebieten. Die Vernetzung ist hier offenbar durch gemeinsame Signalwege und Zielmoleküle gegeben und wird allerdings durch weniger Substanzen gestützt.

7.2.3.4 Variation des Gewichtungsfokus

Bei der zuvor beschriebenen Gewichtungslogik wurde der Gewichtungsfaktor für jede Indikation entsprechend ihres Entwicklungsstatus vergeben. Dabei wurde der Gewichtungsfaktor umso höher gewählt, je weiter der Entwicklungsstatus der Indikation fortgeschritten ist. Das bedeutet, dass Substanzen mit Indikationen, die bereits auf dem Markt sind, ein höheres Gewicht bekommen als Substanzen, die bisher nur präklinisch entwickelt werden. Dieses Verfahren legt den Fokus auf bereits validierte Signalwegverbindungen und wurde daher auch als konservative Gewichtung bezeichnet. Je nach Zielsetzung der Analyse ist es aber aufgrund der flexiblen Einstellung des Gewichtungsfaktors in der Datenbank möglich, den Schwerpunkt der Gewichtung zu verändern. Das bedeutet, dass zum Beispiel Substanzen, die bisher nur in frühen Entwicklungsphasen sind (zum Beispiel Präklinik, Phase I), mit einem höheren Gewichtungsfaktor versehen werden können als Substanzen, die bereits weiter fortgeschrittene Indikationen (zum Beispiel Phase III, nach Zulassung) haben. Abb. 7.24 zeigt ein Beispiel für eine solche Veränderung des Gewichtungsfaktors, die zum Beispiel sinnvoll ist, wenn der Schwerpunkt der Analyse auf Verbindungen, die primär durch neue Projekte gestützt sind, ausgerichtet ist.

Entwicklungsstatus des Forschungsprojekts	Gewichtungsfaktor: Schwerpunkt auf Signalwege, die durch validierte Projekten gestützt werden (konservative Gewichtung)	Gewichtungsfaktor: Fokus auf Signalwege, die durch präklinische/frühe klinische Projekte gestützt werden
Präklinisch	1	3
Klinik (ohne Phasenangabe)	1	2
Klinik Phase I	1	3
Klinik Phase II	2	2
Klinik Phase III	2	2
Vor Zulassung	2	2
Zugelassen	3	1
Eingeführt	3	1
Keine Entwicklung berichtet	0	0
Beendet	0	0
Nicht spezifiziert	0	0

Abb. 7.24: Veränderung des Gewichtungsfaktors bei unterschiedlichem Schwerpunkt der Analyse
Der Gewichtungsfaktor, mit dem die einzelnen Forschungsprojekte gewichtet werden, kann variabel an den Schwerpunkt der Untersuchung angepasst werden.
Liegt der Fokus auf Verbindungen, die durch validierte Projekte (zum Beispiel späte klinische Phasen oder bereits zugelassene Medikamente) gestützt werden, sollte der Gewichtungsfaktor so gewählt werden, dass weit fortgeschrittene Entwicklungsphasen einen hohen Wert bekommen (vergleiche linke Spalte, konservative Gewichtung). Sollen vor allem Verbindung, die durch neue Forschungsprojekte in frühen Entwicklungsphase gestützt sind, hervorgehoben werden, dann sollten diese einen hohen Gewichtungsfaktor bekommen (vergleiche rechte Spalte).
Der absolute Wert des Gewichtungsfaktors kann beliebig gewählt werden. Je größer die Spannbreite der Gewichtungsfaktoren, desto stärker ist die Betonung einzelnen Forschungsprojekte in einem bestimmten Entwicklungsstatus.

Abb. 7.25 und Abb. 7.26 zeigen das Ergebnis der Vernetzungsanalyse nach Veränderung des Gewichtungsfokus. In beiden Abbildungen liegt der Schwerpunkt nun auf Signalwegverbindungen, die durch Projekte in frühen Entwicklungsphasen gestützt werden.

Überraschenderweise lässt sich durch die Veränderung des Gewichtungsfaktors nur eine geringe Veränderung des Vernetzungsbilds erkennen. Hauptunterschied der Matrizen aus den Abb. 7.25 und Abb. 7.26 im Vergleich zu den zuvor beschriebenen mit konservativer Gewichtung (vergleiche Abb. 7.21 und Abb. 7.22) ist, dass Verbindungen mit dem Herz-Kreislauf-System beim Schwerpunkt auf neuen Projekten weniger stark gewichtet werden. Offenbar wird die Vernetzung des Herz-Kreislauf-Gebiets mit anderen therapeutischen Gebieten vor allem durch bereits eingeführte oder klinisch weit fortgeschrittene Substanzen gestützt. Die Onkologie bleibt auch nach Veränderung des Gewichtungsschwerpunkts das am stärksten vernetzte therapeutische Gebiet. Verbindungen mit der Onkologie sind also gleichermaßen durch neue und

validierte Forschungsprojekte gestützt. Im Vergleich der beiden Vernetzungsmatrizen mit einfacher Durchschnittsbildung (Abb. 7.21 und Abb. 7.25) fällt auf, dass der Verbindungspunkt mit der maximalen Verbindungsdichte („Hot-Spot“) von der Verbindung Herz-Kreislauf-System mit Onkologie zur Verbindung Bewegungsapparat mit Onkologie wandert. Dieser neue „Hot-Spot“ stimmt mit dem „Hot-Spot“, der bei doppelter Durchschnittsbildung unabhängig vom Gewichtungsfokus auftritt (Abb. 7.22 und Abb. 7.26), überein. Dieser Wechsel des „Hot-Spots“ ist ein weiterer Beleg dafür, dass Verbindungen mit dem Herz-Kreislauf-System vor allem durch weit fortgeschrittene Forschungsprojekte gestützt werden.

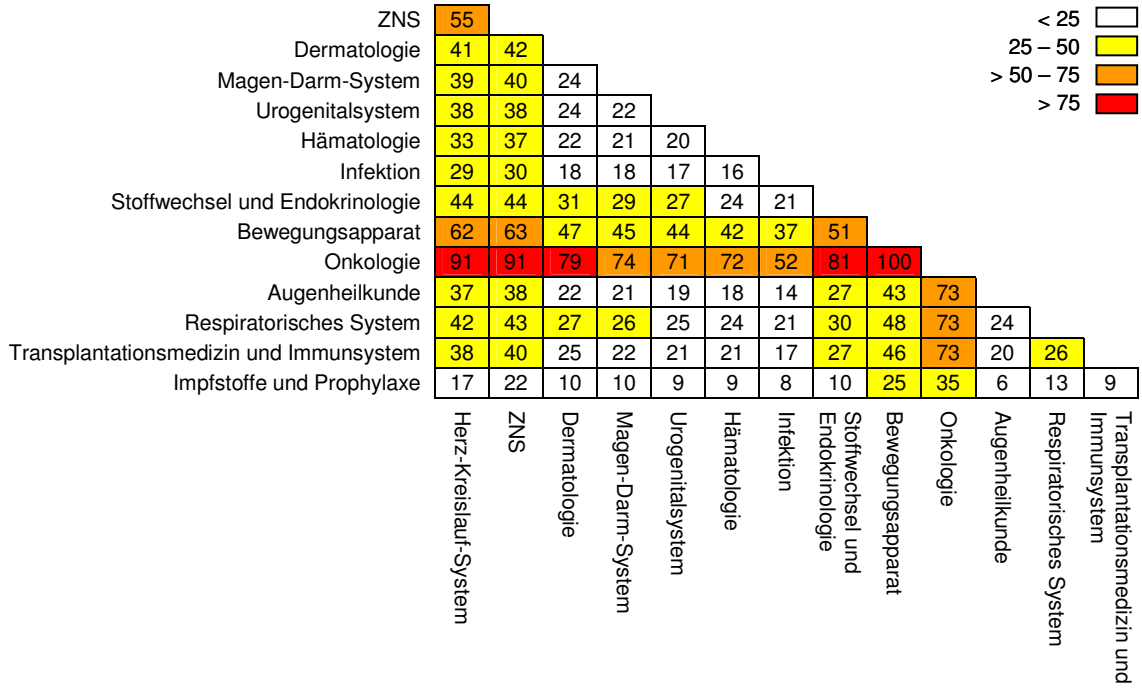


Abb. 7.25: Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit einfacher Durchschnittsbildung und Schwerpunkt auf Signalwegverbindungen, die durch präklinische/frühe klinische Projekte gestützt werden

Die Abbildung zeigt die Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit veränderten Gewichtungsfaktor und einfacher Durchschnittsbildung. Die Verbindungsstärke der Verbindung zwischen zwei therapeutischen Gebieten ist in den Feldern der Matrix als normierter Verbindungswert zwischen 0 und 100 eingetragen. Das bedeutet, dass die stärkste Verbindung den Wert 100 hat und die anderen in Verhältnis dazu eingestuft werden. Die Rottönung der Matrix ist an die Normierung angepasst (vergleiche Legende oben rechts in der Abbildung).
Der „Hot-Spot“ der Matrix, d.h. der Verbindungspunkt mit dem maximalen Verbindungswert, ist die Verbindung der Onkologie mit dem Bewegungsapparat. Trotz Veränderung des Gewichtungsfaktors und des Schwerpunkts auf Signalwegen, ähnelt die Matrix der aus Abb. 7.21. Lediglich Verbindungen mit dem Herz-Kreislauf-System haben insgesamt niedrigere Verbindungswerte.

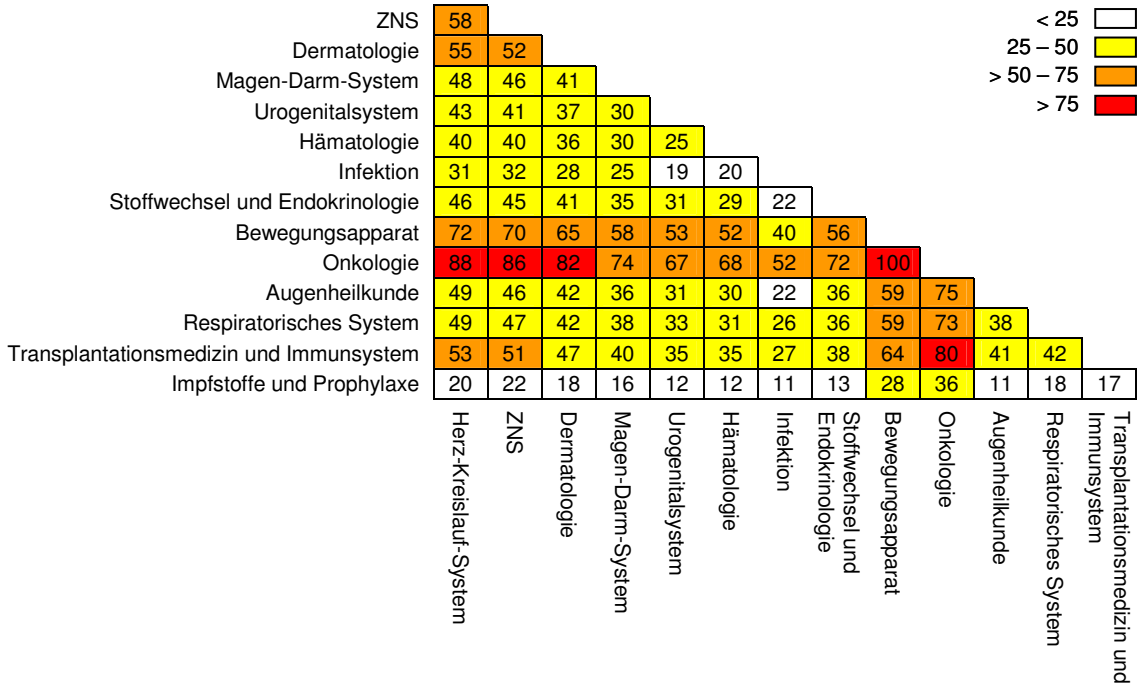


Abb. 7.26: Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit doppelter Durchschnittsbildung und Schwerpunkt auf Signalwegverbindungen, die durch präklinische/frühe klinische Projekte gestützt werden

Die Abbildung zeigt die Vernetzungsmatrix der therapeutischen Gebiete nach Gewichtung mit veränderten Gewichtungsfaktor und doppelter Durchschnittsbildung. Die Verbindungsstärke der Verbindung zwischen zwei therapeutischen Gebieten ist in den Feldern der Matrix als normierter Verbindungswert zwischen 0 und 100 eingetragen. Das bedeutet, dass die stärkste Verbindung den Wert 100 hat und die anderen in Verhältnis dazu eingestuft werden. Die Rottönung der Matrix ist an die Normierung angepasst (vergleiche Legende oben rechts in der Abbildung).

Der „Hot-Spot“ der Matrix, d.h. der Verbindungspunkt mit dem maximalen Verbindungswert, ist die Verbindung der Onkologie mit dem Bewegungsapparat. Die Matrix ähnelt der Verbindungsmatrix aus Abb. 7.22 sehr – offenbar werden die Verbindungen sowohl durch frühe als auch durch validierte Verbindungen stützt, so dass durch Veränderung des Gewichtungsfaktors nur geringe Unterschiede auftreten.

7.3 Analyse der Verbindungen von Indikationsgruppen durch gemeinsame Signalwege

Nach Analyse der Vernetzung der therapeutischen Gebiete, wurde im zweiten Schritt eine detaillierte Untersuchung einzelner Verbindungspunkte zwischen zwei therapeutischen Gebieten durchgeführt, um folgende Fragen zu klären:

- Wird die Verbindung der beiden therapeutischen Gebiete in gleichem Maß durch alle Indikationsgruppen innerhalb der Gebiete bestimmt oder gibt es Indikationsgruppen, die besonders eng verbunden sind?
- Welches sind die Kombinationen von Indikationsgruppen, bei denen die Chance am größten ist, indikationsübergreifende Zielmoleküle zu finden?

Um diese Fragen zu beantworten, wurde eine Vernetzungsanalyse analog zur Untersuchung von therapeutischen Gebieten durchgeführt. Dabei wurde zunächst die Anzahl der verbindenden Signalwege zwischen zwei Indikationsgruppen bestimmt. Im nächsten Schritt wurde die Gewichtung appliziert. Im Rahmen der Indikationsgruppen-Analysen wurde Gewichtung mit einfacher und doppelter Durchschnittsbildung verwendet, um die Ergebnisse miteinander vergleichen zu können.

Insgesamt wurden die folgenden fünf Verbindungspunkte von therapeutischen Gebieten im Detail analysiert:

- Onkologie mit Bewegungsapparat
- Zentralnervensystem mit Herz-Kreislauf-System
- Zentralnervensystem mit Transplantationsmedizin und Immunsystem
- Bewegungsapparat mit Herz-Kreislauf-System
- Onkologie mit Herz-Kreislauf-System

Die Untersuchung der Verbindungspunkte wurde durch Verwendung von Makros in der Datenbank automatisiert, so dass bei Bedarf weitere Verbindungspunkte analysiert werden können.

7.3.1 Verbindungspunkt Onkologie mit Bewegungsapparat

Abb. 7.27 zeigt die detaillierte Untersuchung der Verbindung der Gebiete Onkologie mit Bewegungsapparat. Diese Verbindung ist besonders interessant, da sie in der Untersuchung auf Ebene der therapeutischen Gebiete als „Hot-Spot“, das heißt als insgesamt stärkste Verbindung, aufgefallen war. Abb. 7.27 zeigt sowohl die ungewichtete Verbindungsmatrix, aus der die Anzahl der gemeinsamen Signalwege hervorgeht, als auch die beiden gewichteten Matrizen mit einfacher und doppelter Durchschnittsbildung.

Bei Betrachtung der ungewichteten Verbindungsmatrix fällt auf, dass die enge Verbindung zwischen den beiden therapeutischen Gebieten nicht durch alle Verbindungen von Indikationsgruppen gleichermaßen verursacht ist, sondern vor allem durch Verbindungen mit der Indikationsgruppe rheumatoide Arthritis getrieben wird. Rheumatoide Arthritis ist mit allen Indikationsgruppen mit Ausnahme der unterstützenden Therapien der Onkologie eng verbunden. Bei der Indikationsgruppe der unterstützenden Therapien handelt es sich allerdings um eine sehr heterogene und für die Onkologie eher untypische Indikationsgruppe, die die Indikationen Kachexie, krebsbedingte Müdigkeit sowie die Behandlung chemotherapieinduzierte Nebenwirkungen (zum Beispiel Übelkeit, Schleimhautentzündungen) zusammenfasst. Eine geringere Verbindung mit dieser Gruppe ist daher nicht überraschend.

Ein ähnliches Phänomen lässt sich bei der Indikationsgruppe Anästhesie beobachten. Während alle anderen Indikationsgruppen mit jedem möglichen Partner mindestens zehn gemeinsame Signalwege haben, teilt sich die Anästhesie nur zwei bzw. drei gemeinsame Signalwege mit den Verbindungspartnern. Dies ist aufgrund der der spezifischen Eigenschaften der Anästhesie, die sehr zielgerichtet nur bestimmte Unterbereiche des Nervensystems beeinflusst, nachvollziehbar.

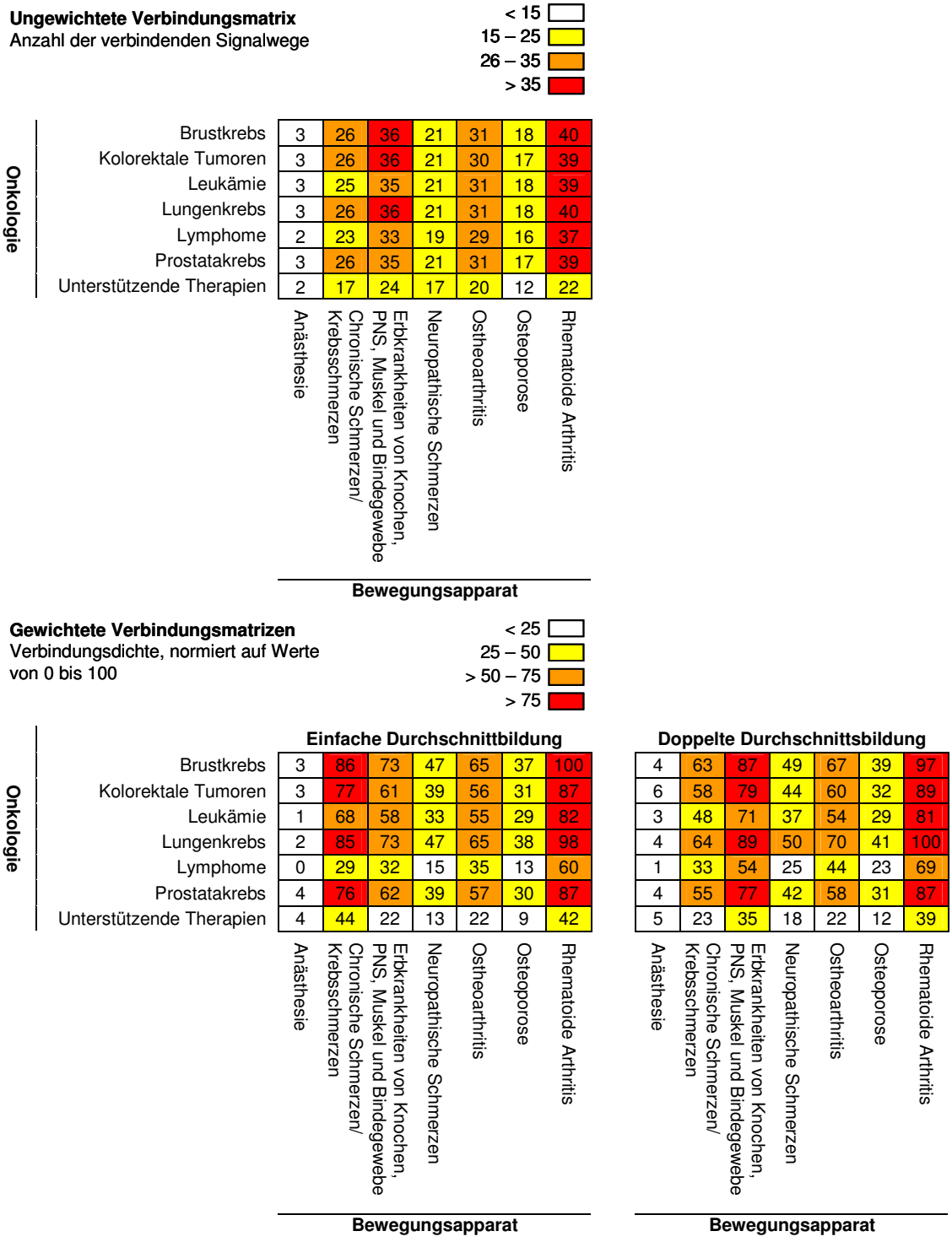


Abb. 7.27: Vernetzungsmatrix der Indikationsgruppen für die Verbindung Onkologie mit Bewegungsapparat

Die obere Matrix zeigt die Anzahl der verbindenden Signalwege zwischen den Indikationsgruppen. Je mehr Signalwege, desto intensiver ist die Rottönung. Die beiden unteren Matrizen zeigen die Ergebnisse nach Gewichtung mit einfacher (links) und doppelter Durchschnittsbildung (rechts). Die Verbindungsstärke ist als normierter Verbindungswert zwischen 0 (keine Verbindung) und 100 (maximale Verbindung) eingetragen. Je intensiver die Rottönung eines Feldes ist, desto höher ist der normierte Verbindungswert.

(Fortsetzung der Bildunterschrift auf der nächsten Seite)

*(Fortsetzung der Bildunterschrift von **Abb. 7.27**.)*
Die Abbildung zeigt unter anderem, dass die enge Verbindung der beiden therapeutischen Gebiete vor allem durch Verbindungen mit der Indikationsgruppe rheumatoide Arthritis getrieben wird. Bei der Gruppe der Erbkrankheiten des Bewegungsapparats kann durch Vergleich der Gewichtungen gezeigt werden, dass zwar viele gemeinsame Signalwege vorliegen, jedoch nur wenige Substanzen diese Verbindungen stützen (niedrige Verbindungswerte bei einfacher Durchschnittsbildung).

Durch Applikation der Gewichtung mit doppelter Durchschnittsbildung verändert sich das Bild im Vergleich zur ungewichteten Matrix nur wenig. Der Verbindungspunkt zwischen rheumatoider Arthritis und Lungenkrebs, der bereits durch 40 gemeinsame Signalwege aufgefallen ist, bestätigt sich auch nach Gewichtung als am stärksten vernetzter Verbindungspunkt. Die Tendenz zur stärkeren Vernetzung aller Gebiete mit den Indikationsgruppen rheumatoide Arthritis und der Erbkrankheiten von Knochen, PNS, Muskel und Bindegewebe ist auch mit doppelter Durchschnittsbildung zu beobachten.

Im Gegensatz dazu sind bei Gewichtung mit einfacher Durchschnittsbildung einige deutliche Unterschiede zur ungewichteten Matrix zu erkennen: Während ohne Gewichtung und mit doppelter Durchschnittsbildung die Gruppe der Erbkrankheiten des Bewegungsapparats stark vernetzt erscheint, ist bei einfacher Durchschnittsbildung eine deutlich geringere Verbindungsstärke zu sehen. Dafür sind allerdings einige onkologische Indikationsgruppen stärker mit der Indikationsgruppe chronische Schmerzen/Krebsschmerzen verbunden. Darüber hinaus ist der am stärksten verbundene Punkt der Matrix nicht mehr zwischen rheumatoider Arthritis und Lungenkrebs sondern zwischen rheumatoider Arthritis und Brustkrebs. Alle drei Unterschiede stehen mit dem stärkeren Einfluss der Substanzanzahl pro Zielmolekül bei der Gewichtung mit einfacher Durchschnittsbildung im Zusammenhang. Offenbar gibt es mehr Substanzen in Forschung und Entwicklung gegen chronische Schmerzen/Krebsschmerzen als gegen Erbkrankheiten des Bewegungsapparats. Das gleiche gilt für Brustkrebs versus Lungenkrebs. Es kann nicht ausgeschlossen werden, dass dies unter anderem durch Unterschiede der wirtschaftlichen Attraktivität der einzelnen Indikationsgruppen für die pharmazeutische Industrie beeinflusst wird.

7.3.2 Verbindungspunkt Zentralnervensystem mit Herz-Kreislauf-System

In Abb. 7.28 ist die Verbindungsmatrix der therapeutischen Gebiete Zentralnervensystem (ZNS) und Herz-Kreislauf-System zu erkennen. Dieser Verbindungspunkt gehört zum „Mittelfeld“ der Matrix der therapeutischen Gebiete und hat in der doppelten Durchschnittsbildung eine normierte Verbindungsstärke von 65. Das bedeutet, seine Verbindungsstärke entspricht 65% des „Hot-Spots“ zwischen Onkologie und Bewegungsapparat.

Bei Betrachtung der ungewichteten Matrix fällt besonders die starke Variation der Verbindungsdichte auf. Während zwischen dem Großteil der Indikationsgruppen nur wenige gemeinsame Signalwege vorliegen, gibt es zwei Verbindungspunkte mit einer deutlich erhöhten Anzahl an gemeinsamen Signalwegen: Thrombose mit multipler Sklerose sowie Thrombose mit Alzheimer und Demenz.

Durch Gewichtung verändert sich das Vernetzungsbild. Vor allem bei Gewichtung mit einfacher Durchschnittsbildung lässt sich eine starke Betonung aller Verbindungen mit der Indikationsgruppe Bluthochdruck feststellen. Die Vermutung, dass dies durch den großen Markt sowie die zahlreichen Substanzen getrieben ist, liegt nahe. Auch bei doppelter Durchschnittsbildung sind die Verbindungen mit Bluthochdruck im Verhältnis zur ungewichteten Matrix stärker. Das bedeutet, dass offenbar nicht nur zahlreiche Substanzen existieren, sondern auch pro Signalweg eine hohe Anzahl an Zielmolekülen die Verbindung stützt. Der „Hot-Spot“ der ungewichteten Matrix (Alzheimer und Demenz mit Thrombose) erhält auch bei doppelter Durchschnittsbildung die größte Verbindungsstärke und scheint daher der „Hot-Spot“ mit den meisten verbindenden Zielmolekülen zu sein. Bei stärkerer Betonung der Anzahl der Substanzen pro Zielmolekül (einfache Durchschnittsbildung) bildet sich ein neuer „Hot-Spot“ zwischen den Indikationsgruppen Angstzustände und Bluthochdruck.

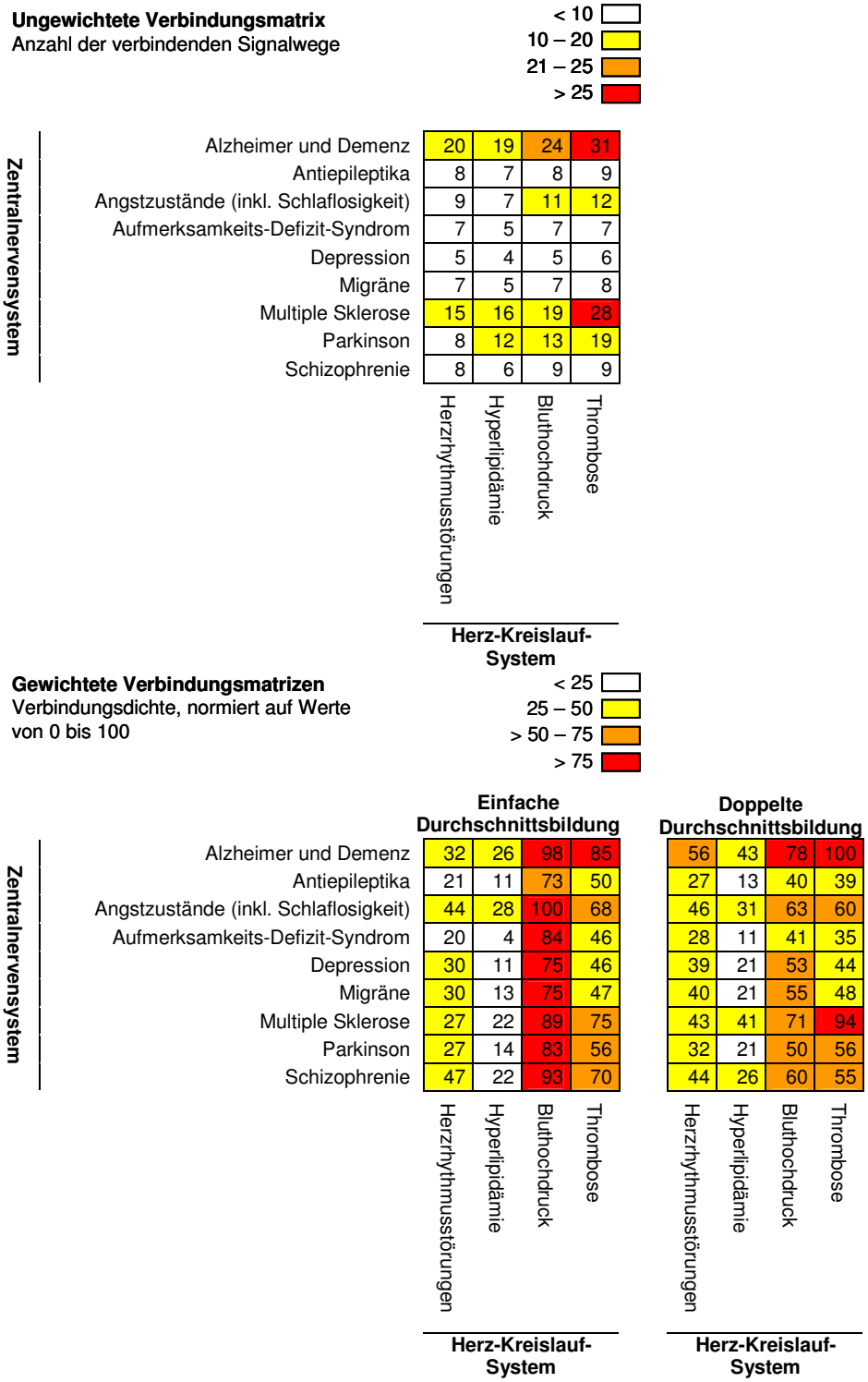


Abb. 7.28: Vernetzungsmatrix der Indikationsgruppen für die Verbindung Zentralnervensystem (ZNS) mit Herz-Kreislauf-System

Die obere Matrix zeigt die Anzahl der verbindenden Signalwege zwischen den Indikationsgruppen. Je mehr Signalwege, desto intensiver ist die Rottönung. Die beiden unteren Matrizen zeigen die Ergebnisse nach Gewichtung mit einfacher (links) und doppelter Durchschnittsbildung (rechts). Die Verbindungsstärke ist als normierter Verbindungswert zwischen 0 (keine Verbindung) und 100 (maximale Verbindung) eingetragen. Je intensiver die Rottönung eines Feldes ist, desto höher ist der normierte Verbindungswert.

(Fortsetzung der Bildunterschrift auf der nächsten Seite)

(Fortsetzung der Bildunterschrift von Abb. 7.28.)
Die Anzahl der gemeinsamen Signalwege zwischen den Indikationsgruppen variiert sehr stark. Nur bei zwei Verbindungspunkten können mehr als 25 gemeinsame Signalwege gezählt werden. Durch Gewichtung wird deutlich, dass vor allem Verbindungen mit den Indikationsgruppen Bluthochdruck und Thrombose für die Vernetzung verantwortlich sind.

7.3.3 Verbindungspunkt Zentralnervensystem mit Transplantationsmedizin und Immunsystem

Abb. 7.29 zeigt die Vernetzungsmatrix der therapeutischen Gebiete Zentralnervensystem (ZNS) mit Transplantationsmedizin und Immunsystem. Auch diese Verbindung gehört mit einer normierten gewichteten Verbindungsstärke von 53 zum „Mittelfeld“ der Verbindungsmatrix. Die ungewichtete Vernetzungsmatrix zeigt ein heterogenes Bild von sehr eng und kaum verbundenen Indikationsgruppen. Die Vernetzung des ZNS mit dem therapeutischen Gebiet Transplantationsmedizin und Immunsystem ist offenbar vor allem durch die Verbindungen der Indikationsgruppen Alzheimer und Demenz, multiple Sklerose, generalisierte Immunerkrankungen inkl. Autoimmunerkrankungen und immunvermittelte Transplantatabstoßung (Organtransplantationen) verursacht. Die Kombinationen dieser vier Indikationsgruppen teilen sich jeweils zahlreiche Signalwege.

Durch Anwendung der Gewichtung verschiebt sich der Fokus insgesamt leicht auf die Indikationsgruppe der generalisierten Immunerkrankungen. Da die Veränderung des Vernetzungsbilds sowohl bei einfacher als auch bei doppelter Durchschnittsbildung sichtbar ist, kann vermutet werden, dass die meisten Verbindungen mit dieser Indikationsgruppe sowohl durch zahlreiche Zielmoleküle als auch durch zahlreiche Substanzen gestützt sind. Je nachdem wie stark die Anzahl der Substanzen pro Zielmolekül in die Gewichtung eingeht, verschiebt sich der stärkste Verbindungspunkt („Hot-Spot“) von der Verbindung von Alzheimer und Demenz mit generalisierten Immunerkrankungen zu multipler Sklerose mit generalisierten Immunerkrankungen. Offenbar gibt es zwar in beiden Fällen zahlreiche gemeinsame Signalwege, jedoch ist die Verbindung mit multipler Sklerose durch eine höhere Anzahl von Zielmolekülen gestützt, auch wenn es in der Verbindung mit Alzheimer und Demenz für die dort vorhandenen Zielmoleküle anscheinend mehr Substanzen gibt.

Ungewichtete Verbindungsmatrix
Anzahl der verbindenden Signalwege

< 10
10 – 20
21 – 25
> 25

Zentralnervensystem	Generalisierte Immunerkrankungen inkl. Autoimmunerkrankungen	Graft-versus-Host-Reaktion (Stammzelltransplantationen)	Immunvermittelte Transplantatabstoßung (Organtransplantationen)
Alzheimer und Demenz	34	12	30
Antiepileptika	9	4	8
Angstzustände (inkl. Schlaflosigkeit)	12	6	11
Aufmerksamkeits-Defizit-Syndrom	7	2	7
Depression	7	1	7
Migräne	8	1	7
Multiple Sklerose	31	15	27
Parkinson	19	11	17
Schizophrenie	9	5	9

Transplantationsmedizin und Immunsystem

Gewichtete Verbindungsmatrizen
Verbindungsdichte, normiert auf Werte von 0 bis 100

< 25
25 – 50
> 50 – 75
> 75

Einfache Durchschnittsbildung

Zentralnervensystem	Generalisierte Immunerkrankungen inkl. Autoimmunerkrankungen	Graft-versus-Host-Reaktion (Stammzelltransplantationen)	Immunvermittelte Transplantatabstoßung (Organtransplantationen)
Alzheimer und Demenz	100	16	55
Antiepileptika	34	13	20
Angstzustände (inkl. Schlaflosigkeit)	94	48	69
Aufmerksamkeits-Defizit-Syndrom	22	3	9
Depression	44	16	33
Migräne	59	17	32
Multiple Sklerose	98	37	67
Parkinson	68	15	36
Schizophrenie	93	33	73

Transplantations-medizin und Immunsystem

Doppelte Durchschnittsbildung

Generalisierte Immunerkrankungen inkl. Autoimmunerkrankungen	Graft-versus-Host-Reaktion (Stammzelltransplantationen)	Immunvermittelte Transplantatabstoßung (Organtransplantationen)
95	19	68
24	6	14
63	20	39
21	4	13
32	12	27
33	10	26
100	36	70
58	13	32
54	16	32

Transplantations-medizin und Immunsystem

(Bildunterschrift auf der nächsten Seite)

Abb. 7.29: Vernetzungsmatrix der Indikationsgruppen für die Verbindung Zentralnervensystem (ZNS) mit Transplantationsmedizin und Immunsystem
Die obere Matrix zeigt die Anzahl der verbindenden Signalwege zwischen den Indikationsgruppen. Je mehr Signalwege, desto intensiver ist die Rottönung. Die beiden unteren Matrizen zeigen die Ergebnisse nach Gewichtung mit einfacher (links) und doppelter Durchschnittsbildung (rechts). Die Verbindungsstärke ist als normierter Verbindungswert zwischen 0 (keine Verbindung) und 100 (maximale Verbindung) eingetragen. Je intensiver die Rottönung eines Feldes ist, desto höher ist der normierte Verbindungswert.
Die ungewichtete Verbindungsmatrix zeigt, dass nur wenige Verbindungen von Indikationsgruppen eine hohe Anzahl gemeinsamer Signalwege haben. Werden durch Gewichtung die Anzahl der Zielmoleküle und der Substanzen pro gemeinsamen Signalweg berücksichtigt, wird deutlich, dass vor allem zwischen den Indikationsgruppen Alzheimer und Demenz und generalisierten Immunerkrankungen sowie multipler Sklerose und generalisierten Immunerkrankungen eine enge Verbindung besteht.

7.3.4 Verbindungspunkt Bewegungsapparat mit Herz-Kreislauf-System

Die Verbindung des Bewegungsapparats mit dem Herz-Kreislauf-System gehört mit einer normierten gewichteten Verbindungsstärke (doppelte Durchschnittsbildung) von 78 zu den stärkeren Verbindungspunkten der Vernetzungsmatrix der therapeutischen Gebiete. Die enge Verbindung der therapeutischen Gebiete spiegelt sich in der ungewichteten Vernetzungsmatrix wider. Beinahe alle Indikationsgruppen haben mindestens zehn gemeinsame Signalwege mit ihren Verbindungspartnern. Einzige Ausnahme stellt die Anästhesie dar, die durch ihre spezifischen Eigenschaften kaum mit anderen Gebieten in Verbindung steht (vergleiche Abschnitt 7.3.1). Zwei Verbindungspunkte dominieren die ungewichtete Matrix: Thrombose mit rheumatoider Arthritis und Thrombose mit Erbkrankheiten von Knochen, PNS, Muskel und Bindegewebe. Beide Verbindungen sind durch 35 gemeinsame Signalwege gekennzeichnet.

Während die Vernetzungsmatrix, die mit doppelter Durchschnittsbildung gewichtet wurde, der ungewichteten Matrix sehr stark ähnelt, führt die Gewichtung mit einfacher Durchschnittsbildung zu einem anderen Bild. Die Indikationsgruppen Bluthochdruck und chronische Schmerzen/Krebsschmerzen bekommen bei einfacher Durchschnittsbildung einen deutlich höheren Verbindungswert.

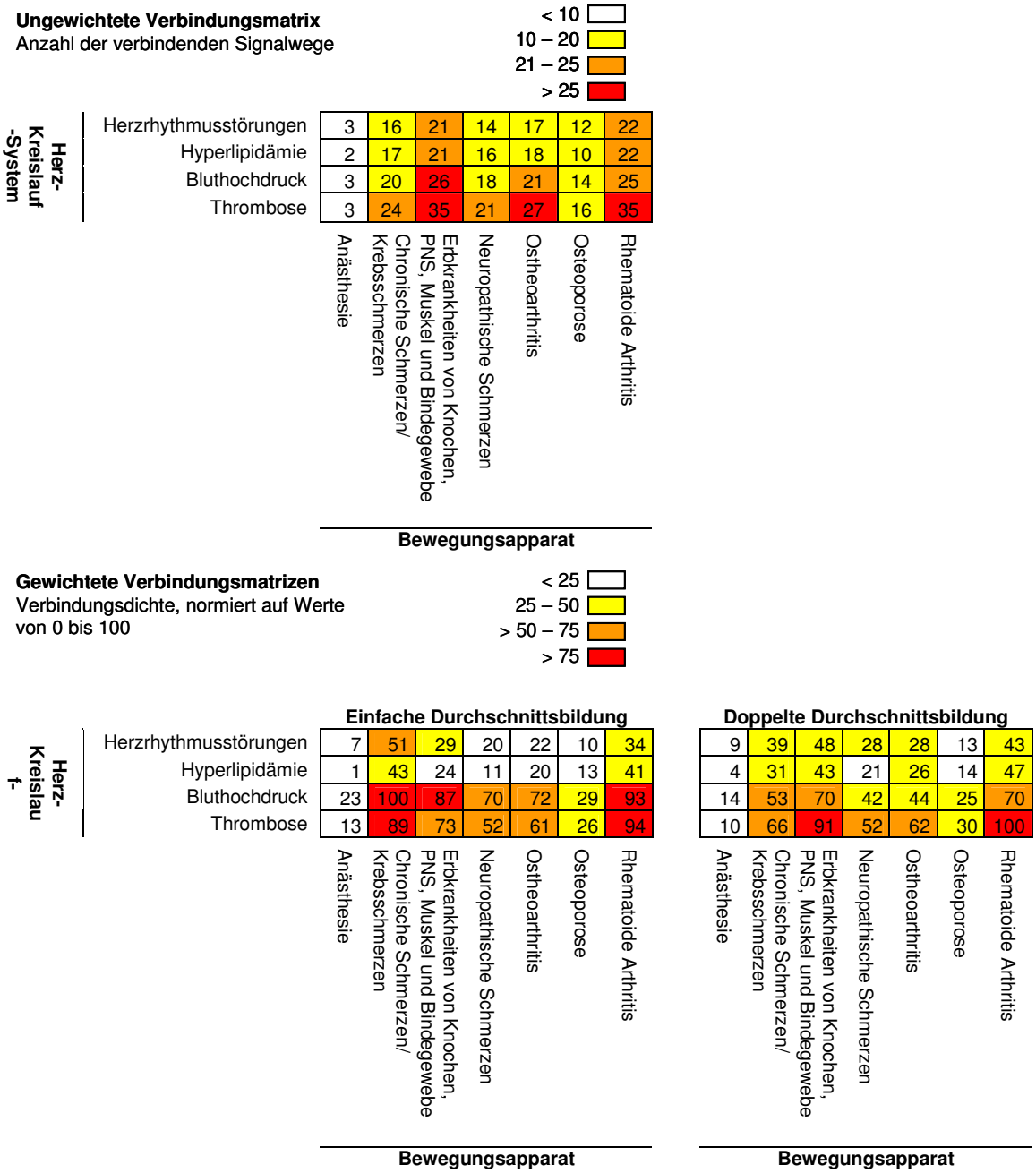


Abb. 7.30: Vernetzungsmatrix der Indikationsgruppen für die Verbindung Herz-Kreislauf-System mit Bewegungsapparat

Die obere Matrix zeigt die Anzahl der verbindenden Signalwege zwischen den Indikationsgruppen. Je mehr Signalwege, desto intensiver ist die Rottönung. Die beiden unteren Matrizen zeigen die Ergebnisse nach Gewichtung mit einfacher (links) und doppelter Durchschnittsbildung (rechts). Die Verbindungsstärke ist als normierter Verbindungswert zwischen 0 (keine Verbindung) und 100 (maximale Verbindung) eingetragen. Je intensiver die Rottönung eines Feldes ist, desto höher ist der normierte Verbindungswert.

Die ungewichtete Matrix zeigt, dass sich mit Ausnahme der Anästhesie alle Indikationsgruppen eine verhältnismäßig hohe Anzahl von Signalwegen mit ihren Partnern teilen. Durch Gewichtung wird deutlich, dass insbesondere die Indikationsgruppen Bluthochdruck und Thrombose sowie rheumatoide Arthritis und chronische Schmerzen/Krebsschmerzen eng untereinander verbunden sind.

Im Gegensatz dazu wird die Stärke der Verbindungen mit der Indikationsgruppe Erbkrankheiten des Bewegungsapparats geringer bewertet und der „Hot-Spot“ verschiebt sich von der Verbindung von rheumatoider Arthritis mit Thrombose zur Verbindung von Bluthochdruck mit chronischen Schmerzen/Krebsschmerzen.

Diese Veränderung im Vernetzungsbild ist durch die stärke Betonung der Anzahl an Substanzen bei der einfachen Durchschnittsbildung verursacht und legt einen Zusammenhang zur wirtschaftlichen Attraktivität der Indikationsgruppen nahe. Während Bluthochdruck und chronische Schmerzen/Krebsschmerzen durch hohe Prävalenz umsatzstarke Optionen für Pharmafirmen bieten, scheint das Interesse an der Entwicklung von Medikamenten gegen Erbkrankheiten des Bewegungsapparats deutlich geringer zu sein.

Einen weiteren Beleg für den unterschiedlichen Schwerpunkt der Gewichtungsarten lässt sich am Verbindungspunkt Bluthochdruck mit rheumatoider Arthritis beobachten. Bei doppelter Durchschnittsbildung kann die Verbindung 70 % der Verbindungsstärke des „Hot-Spots“ erreichen; bei einfacher Durchschnittsbildung stützen offenbar so viele Substanzen die gemeinsamen Signalwege, dass es für 93 % ausreicht.

7.3.5 Verbindungspunkt Onkologie mit Herz-Kreislauf-System

Abb. 7.31 zeigt die Analyse des Verbindungspunkt Onkologie mit Herz-Kreislauf-System. Diese Verbindung stellt nach dem „Hot-Spot“ Onkologie mit Bewegungsapparat die zweitstärkste Verbindung von therapeutischen Gebieten dar (gewichtet, doppelte Durchschnittsbildung). Bei der Detailanalyse lassen sich – im Verhältnis zu den zuvor beschriebenen Verbindungspunkten – verhältnismäßig wenige Unterschiede der Verbindungsmatrizen vor und nach Gewichtung erkennen. Insgesamt tragen alle Indikationsgruppen zur Verbindung der therapeutischen Gebiete bei. Verbindungen mit der Indikationsgruppe Thrombose erhalten die höchste Anzahl gemeinsamer Signalwege sowie die höchsten normierten Verbindungswerte. Die Indikationsgruppe der unterstützenden Therapien hat tendenziell etwas weniger Verbindungen zu anderen Indikationsgruppen. Dies entspricht der Beobachtung, die bereits bei der

Vernetzungsanalyse der Gebiete Onkologie mit Bewegungsapparat gemacht wurde (vergleiche Abschnitt 7.3.1). Bei Gewichtung mit einfacher Durchschnittsbildung erhalten Verbindungen mit der Indikationsgruppe Bluthochdruck neben den Verbindungen mit Thrombose die höchsten normierten Verbindungswerte. In der ungewichteten Vernetzungsmatrix erscheinen zwei Verbindungspunkte als „Hot-Spots“, die beide durch 35 gemeinsame Signalwege gekennzeichnet sind (Brustkrebs mit Thrombose und Lungenkrebs mit Thrombose).

Durch Gewichtung ist eine Differenzierung zwischen den beiden „Hot-Spots“ möglich. Die Verbindung von Lungenkrebs mit Thrombose ist offenbar am besten durch Zielmoleküle gestützt und erscheint als „Hot-Spot“ bei doppelter Durchschnittsbildung. Die Verbindung von Brustkrebs mit Thrombose hingegen hat offenbar mehr Substanzen in Forschung und Entwicklung, was das Vertrauen der Industrie anzeigt und dazu führt, dass der Verbindungspunkt als „Hot-Spot“ bei einfacher Durchschnittsbildung auftritt.

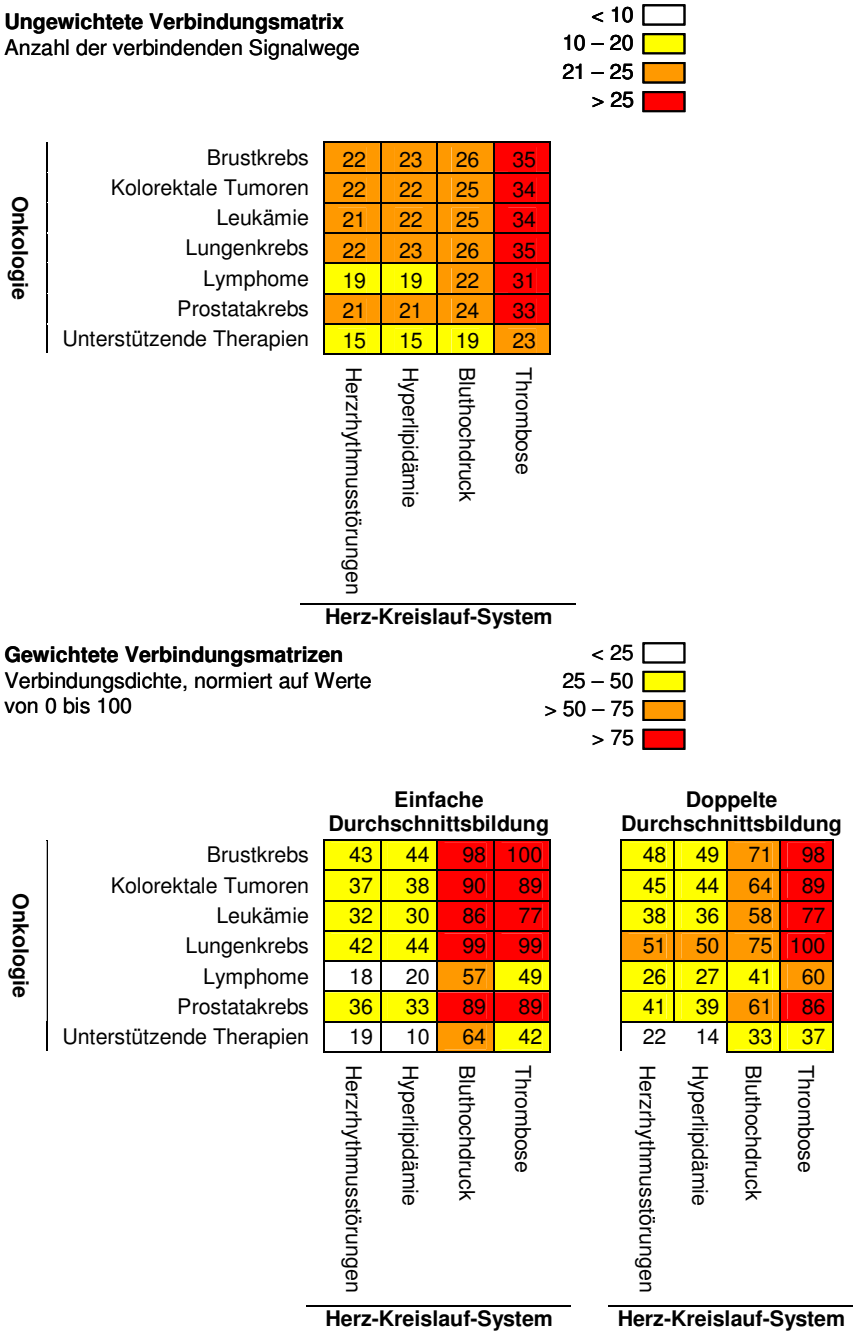


Abb. 7.31: Vernetzungsmatrix der Indikationsgruppen für die Verbindung Onkologie mit Herz-Kreislauf-System

Die obere Matrix zeigt die Anzahl der verbindenden Signalwege zwischen den Indikationsgruppen. Je mehr Signalwege, desto intensiver ist die Rottönung. Die beiden unteren Matrizen zeigen die Ergebnisse nach Gewichtung mit einfacher (links) und doppelter Durchschnittsbildung (rechts). Die Verbindungsstärke ist als normierter Verbindungswert zwischen 0 (keine Verbindung) und 100 (maximale Verbindung) eingetragen. Je intensiver die Rottönung eines Feldes ist, desto höher ist der normierte Verbindungswert.

Dieser Verbindungspunkt ist die zweitstärkste Verbindung von zwei therapeutischen Gebieten. Dies spiegelt sich in der insgesamt hohen Anzahl von gemeinsamen Signalwegen wider. Besonders die Indikationsgruppen Thrombose und Bluthochdruck sind mit den onkologischen Indikationsgruppen eng verbunden.

7.4 Identifikation und Priorisierung von gemeinsamen Signalwegen

Die bisher dargestellten Analysen erlauben die Frage nach besonders eng vernetzten therapeutischen Gebieten bzw. Indikationsgruppen zu beantworten und entsprechen den ersten beiden Schritten der fünfschrittigen Analyseabfolge (vergleiche Abschnitt 7.1). Im folgenden Abschnitt werden die Untersuchungen des dritten Analyseschritts erläutert: die Identifikation und Priorisierung von verbindenden Signalwegen.

7.4.1 Methodik und Anwendung

Jeder Verbindung von zwei therapeutischen Gebieten liegen gemeinsame Signalwege zu Grunde. Jede Signalwegverbindung ist durch Zielmoleküle gestützt, die in einem oder beiden therapeutischen Gebieten eine Rolle spielen. Wie im Zusammenhang mit der Gewichtungslogik beschrieben wurde, ist ein Signalweg dann besonders wichtig für die Verbindung, wenn er über zahlreiche, möglichst in beiden therapeutischen Gebieten relevante Zielmoleküle verfügt. Zusätzlich kann – je nach verwendetem Gewichtungstyp – auch die Anzahl der Substanzen, die für die Zielmoleküle in dem therapeutischen Gebiet entwickelt werden, berücksichtigt werden (vergleiche Abschnitt 7.2.3.2). Entsprechend ihrer Bedeutung für eine bestimmte Verbindung von therapeutischen Gebieten können die verbindenden Signalwege in einer Rangliste sortiert werden. Dies ist zum Beispiel dann sinnvoll, wenn für eine bestimmte Verbindung eine Priorisierung der Forschungs- und Entwicklungsprojekte stattfinden soll. Forschungsprojekte, die sich auf besonders stark verbindende Signalwege konzentrieren, haben eine erhöhte Chance zur Entwicklung von Medikamenten zu führen, die in beiden therapeutischen Gebieten nützlich sind.

Um die wichtigsten Signalwege für eine Verbindung von therapeutischen Gebieten identifizieren zu können, wurde in der Datenbank eine Folge von Abfragen aufgebaut, die für jede Verbindung von therapeutischen Gebieten eine

Rangliste der verbindenden Signalwege erstellt. Dabei werden die folgenden Aspekte berücksichtigt:

- Die Verbindungsstärke eines Signalwegs wird durch seinen Anteil an der Gesamtverbindungsstärke ausgedrückt.
- Der Verbindungsanteil wird sowohl mit doppelter als auch mit einfacher Gewichtung berechnet, um mögliche Einflüsse durch hohe Substanzanzahl erkennen zu können.
- Für jeden verbindenden Signalweg wird die Zahl der stützenden Zielmoleküle ermittelt. Dabei wird zwischen Zielmolekülen, die direkt verbindend, das heißt bereits heute für beide therapeutischen Gebiete relevant sind, und indirekt verbindend, das heißt bisher nur für eines der beiden Gebiete relevant sind, unterschieden. Dies ist sinnvoll, um eine Unterscheidung zwischen Verbindungen machen zu können, die vor allem direkt oder indirekt gestützt sind. Signalwegverbindungen mit vielen direkt verbindenden Molekülen sind bereits verhältnismäßig validiert und die Rolle der direkt indikationsübergreifenden Zielmoleküle ist in der Regel bekannt. Signalwege, die bisher nur indirekt verbindende Zielmoleküle haben, sind weniger validiert, bieten aber dafür einen höheren Neuheitswert.
- Pro Signalweg wird angegeben, welcher Anteil des Signalwegs für die Verbindung von Bedeutung ist. Dieser sogenannte relevante Anteil berechnet sich aus der Anzahl der Zielmoleküle, die für die Verbindung direkt und indirekt stützend sind im Verhältnis zur Gesamtzahl von Zielmolekülen des Signalwegs. Eine Berücksichtigung dieses Anteils ist wichtig, um bei der Priorisierung große Signalwege von kleinen unterscheiden zu können. Ein großer Signalweg hat durch seine hohe Anzahl an Molekülen generell eine große Chance, eine hohe Anzahl von verbindenden Molekülen zu stellen. Ein Signalweg, der jedoch einen hohen relevanten Anteil hat, ist spezifisch für die Verbindung wichtig. In der Auswahl von geeigneten Signalwegen für Forschung und Entwicklung sollte darauf geachtet werden, dass neben genügend

Zielmolekülkandidaten auch ein möglichst hoher relevanter Anteil des Signalwegs vorliegt.

7.4.2 Identifikation und Priorisierung von Signalwegen in der Verbindung von therapeutischen Gebieten: Beispiel Hämatologie und Magen-Darm-System

Abb. 7.32 zeigt ein Beispiel der Signalwegidentifizierung und –priorisierung für die Verbindung der therapeutischen Gebiete Hämatologie und Magen-Darm-System. Diese beiden therapeutischen Gebiete sind durch 30 gemeinsame Signalwege verbunden (vergleiche ungewichtete Vernetzungsmatrix Abb. 7.6). Bei doppelter Gewichtung erreicht die Verbindung einen normierten Verbindungswert von 40 (vergleiche gewichtete Verbindungsmatrix Abb. 7.22).

Die Rangliste der verbindenden Signalwege (Abb. 7.32) macht deutlich, dass von den 30 Signalwegen vor allem eine Gruppe von etwa sieben bis zehn Signalwegen für die Verbindung wichtig sind. Bei Gewichtung mit doppelter Durchschnittsbildung machen die zehn am stärksten verbindenden Signalwege bereits 68 % der Gesamtverbindung aus. Berücksichtigt man darüber hinaus noch die Anzahl der Substanzen und damit die kommerzielle Attraktivität (einfache Durchschnittsbildung), reichen bereits die sieben wichtigsten Signalwege aus, um 69 % der Gesamtverbindungsstärke zu erreichen. Vor allem der Signalweg „Neuroactive ligand-receptor interaction“ dominiert die Verbindung von Hämatologie mit dem Magen-Darm-System mit mehr als 15 % Anteil an der Verbindungsstärke unabhängig von der Gewichtungsart.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Neuroactive ligand-receptor interaction**	**16**	**17**	**4**	**46**	**20**
2.	**Cytokine-cytokine receptor interaction**	**9**	**11**	**1**	**38**	**16**
3.	**Calcium signaling pathway**	**9**	**10**	**2**	**26**	**16**
4.	**Complement and coagulation cascades**	**8**	**8**	**2**	**19**	**30**
5.	**Jak-STAT signaling pathway**	**6**	**9**	**0**	**22**	**14**
6.	**Regulation of actin cytoskeleton**	**6**	**6**	**1**	**17**	**9**
7.	**Insulin signaling pathway**	**5**	**9**	**1**	**21**	**16**
8.	**Focal adhesion**	**4**	**3**	**0**	**18**	**9**
9.	**Toll-like receptor signaling pathway**	**3**	**2**	**2**	**10**	**12**
10.	**MAPK signaling pathway**	**3**	**3**	**2**	**11**	**5**
11.	Natural killer cell mediated cytotoxicity	3	2	0	11	8
12.	Cell adhesion molecules (CAMs)	3	1	0	14	11
13.	T cell receptor signaling pathway	2	2	1	9	11
14.	ECM-receptor interaction	2	2	0	9	10
15.	Apoptosis	2	2	1	8	11
16.	Leukocyte transendothelial migration	2	1	1	7	7
17.	Adipocytokine signaling pathway	2	2	1	8	13
18.	Gap junction	2	2	0	8	8
19.	Fc epsilon RI signaling pathway	2	1	1	5	8
20.	Cell cycle	2	1	2	3	4
21.	VEGF signaling pathway	2	3	1	4	7
22.	GnRH signaling pathway	1	1	1	3	4
23.	Notch signaling pathway	1	1	2	1	6
24.	ErbB signaling pathway	1	1	0	5	6
25.	Axon guidance	1	1	0	6	5
26.	B cell receptor signaling pathway	1	1	1	4	8
27.	Insulin signaling pathway	1	0	0	4	3
28.	p53	1	0	0	4	6
29.	Adherens junction	1	0	0	3	4
30.	Wnt signaling pathway	0	0	0	2	1

Abb. 7.32: Signalwegrangliste für die Verbindung der Hämatologie mit dem Magen-Darm-System

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Hämatologie und Magen-Darm-System wichtig sind. „Indirekt“ zählt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eines der beiden Gebiete genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Es fällt auf, dass vor allem der Signalweg „Neuroactive ligand-receptor interaction“ wichtig für die Verbindung ist. Er hat einen großen Anteil an der Gesamtverbindungsstärke, viele relevante Zielmoleküle und mit 20 % einen verhältnismäßig hohen relevanten Anteil. Mit einem relevanten Anteil von 30 % ist auch der Signalweg „Complement and coagulation cascades“ interessant für indikationsübergreifende Projekte.

Auf den ersten Blick ist dies nicht unbedingt überraschend, da es sich bei diesem Signalweg nicht nur um eine sehr große funktionelle Gruppe von Molekülen handelt, sondern auch um einen relativ heterogen zusammengesetzten Signalweg, der eine Vielzahl von unterschiedlichen Neurotransmittern und Rezeptoren beinhaltet. Bei näherer Betrachtung lassen sich aber mehrere Hinweise dafür

finden, dass der Signalweg tatsächlich von besonderer Bedeutung ist: 50 Moleküle innerhalb des „Neuroactive ligand-receptor interaction“ Signalwegs spielen für die Verbindung eine Rolle; vier davon sind direkt verbindend. Im Vergleich zu den anderen Signalwegen ist diese Anzahl an direkten Verbindungen deutlich am größten. Insgesamt sind 20 % des Signalwegs für die Verbindung von Bedeutung. Dieser sogenannte relevante Anteil ist lediglich bei der Gerinnungskaskade („Complement and coagulation cascades“) höher. Neben dem „Neuroactive ligand-receptor interaction“ Signalweg sollte auch dieser Signalweg für Projekte, die beide therapeutischen Gebiete überspannen, beachtet werden.

7.4.3 Identifikation und Priorisierung von Signalwegen in der Verbindung von Indikationsgruppen: Beispiel Lungenkrebs und rheumatoide Arthritis

Die Identifikation und Priorisierung von Signalwegen erfolgt für Verbindungen zwischen Indikationsgruppen genauso wie für Verbindungen von therapeutischen Gebieten. Grundsätzlich gilt, dass bei einer möglichen Nutzung zur Priorisierung von Forschungs- und Entwicklungsprojekten immer die Signalwegrangliste auf der tiefstmöglichen Detailebene verwendet werden sollte. Dies ist deshalb wichtig, da die Signalwegranglisten auf Ebene von therapeutischen Gebieten zwar einen guten Indikator für verbindende Signalwege bieten, jedoch stets eine gewisse Unschärfe enthalten, da in der Einheit des therapeutischen Gebiets eine verhältnismäßig große Anzahl von Indikationen zusammengefasst werden.

Zur Priorisierung der Signalwege auf Indikationsgruppenebene wurde aus jeder Verbindungsmatrix, die im Abschnitt 7.3 beschrieben wurde, ausgewählte Verbindungspunkte herausgegriffen, für die eine detaillierte Signalwegrangliste erstellt wurde. Im Folgenden soll nur die Verbindung zwischen Lungenkrebs und rheumatoider Arthritis im Detail dargestellt werden, um die Übersichtlichkeit des Kapitels zu erhöhen. Alle anderen Verbindungsanalysen können im Anhang nachgelesen werden.

Abb. 7.27 konnte zeigen, dass vor allem die Verbindung zwischen Lungenkrebs und rheumatoider Arthritis wichtig für die Verbindung der therapeutischen Gebiete ist. Abb. 7.33 zeigt die Signalwegrangliste für diesen „Hot-Spot“.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Cytokine-cytokine receptor interaction**	**11**	**12**	**9**	**44**	**21**
2.	**MAPK signaling pathway**	**6**	**7**	**7**	**21**	**11**
3.	**Focal adhesion**	**5**	**7**	**4**	**25**	**15**
4.	**Jak-STAT signaling pathway**	**5**	**3**	**5**	**17**	**14**
5.	**Hematopoietic cell lineage**	**5**	**5**	**6**	**17**	**26**
6.	**Toll-like receptor signaling pathway**	**4**	**4**	**5**	**16**	**21**
7.	**Natural killer cell mediated cytotoxicity**	**4**	**3**	**6**	**12**	**14**
8.	**ErbB signaling pathway**	**4**	**4**	**4**	**13**	**20**
9.	**Apoptosis**	**3**	**4**	**5**	**12**	**20**
10.	**Adipocytokine signaling pathway**	**3**	**3**	**7**	**6**	**18**
11.	Regulation of actin cytoskeleton	3	3	3	13	8
12.	T cell receptor signaling pathway	3	2	4	8	13
13.	Gap junction	3	4	2	11	13
14.	VEGF signaling pathway	3	7	2	10	17
15.	Fc epsilon RI signaling pathway	3	3	4	7	15
16.	Neuroactive ligand-receptor interaction	2	2	0	14	6
17.	Cell adhesion molecules (CAMs)	2	1	1	11	9
18.	p53	2	1	2	10	18
19.	Cell cycle	2	1	2	9	10
20.	GnRH signaling pathway	2	3	3	7	10
21.	Calcium signaling pathway	2	4	0	12	7
22.	Leukocyte transendothelial migration	2	2	2	7	8
23.	Long-term depression	2	1	3	5	11
24.	Insulin signaling pathway	2	1	3	4	5
25.	Melanogenesis	2	1	3	5	8
26.	B cell receptor signaling pathway	2	1	2	6	13
27.	ECM-receptor interaction	2	1	1	9	11
28.	Axon guidance	1	1	2	4	5
29.	PPAR signaling pathway	1	1	1	4	7
30.	Adherens junction	1	2	1	6	9
31.	Long-term potentiation	1	1	2	3	7
32.	mTOR signaling pathway	1	1	1	4	11
33.	Tight junction	1	1	0	6	5
34.	Dorso-ventral axis formation	1	2	1	4	18
35.	TGF-beta signaling pathway	1	1	1	4	6
36.	Wnt signaling pathway	1	1	1	3	3
37.	Notch signaling pathway	1	0	3	0	6
38.	Antigen processing and presentation	1	0	1	2	4
39.	Phosphatidylinositol signaling system	1	0	0	3	4
40.	Complement and coagulation cascades	0	0	0	3	4

Abb. 7.33: Signalwegrangliste für die Verbindung von Lungenkrebs mit rheumatoider Arthritis

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Lungenkrebs und rheumatoide Arthritis wichtig sind. „Indirekt“ zählt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Die Signalwegrangliste zeigt unter anderem, dass vor allem Signalwege grundlegender Zellfunktionen für die Verbindung von Lungenkrebs und rheumatoider Arthritis eine Rolle spielen (z.B. „Cytokine-cytokine interaction“).

Die Verbindung zwischen Lungenkrebs und rheumatoider Arthritis wird vor allem durch den Signalweg „Cytokine-cytokine interaction“ bestimmt. Neun Zielmoleküle sind innerhalb dieses Signalwegs direkt verbindend, d.h. heute bereits für Lungenkrebs und rheumatoide Arthritis in Forschung und Entwicklung. 44 weitere sind indirekt für die Verbindung von Bedeutung, das heißt Teil des verbindenden Signalwegs, aber bisher nur entweder für Lungenkrebs oder rheumatoide Arthritis genutzt. Der „Cytokine-cytokine interaction“ Signalweg gehört zu den größeren Signalwegen und hat für die Verbindung von Lungenkrebs zu rheumatoider Arthritis einen relevanten Anteil von 21 %. Das bedeutet, dass mehr als ein Fünftel aller Zielmoleküle aus dem Signalweg direkt oder indirekt für die Verbindung eine Rolle spielen. Neben dem „Cytokine-cytokine interaction“ Signalweg sind vor allem der MAP-Kinase Signalweg („MAPK signaling pathway“) und der Jak-STAT Signalweg („Jak-STAT signaling pathway“) für die Verbindung der Indikationsgruppen von Bedeutung. Sie verfügen über eine hohe Anzahl direkt und indirekt verbindender Zielmoleküle und haben einen relevanten Anteil von mehr als 10 %. Eine Funktion von Signalwegen, die an grundlegenden Zellfunktionen wie der Signaltransduktion beteiligt sind, macht insbesondere bei rheumatoider Arthritis und Lungenkrebs Sinn – beides sind Erkrankungen, die durch fehlgesteuertes Wachstum gekennzeichnet sind. In diesem Zusammenhang ist auch die Beobachtung interessant, dass weitere grundlegende Zellmechanismen in beiden Erkrankungen eine Rolle spielen. Sowohl der „Focal adhesion“ Signalweg, der für die Integrität des Gewebes und mögliche Metastasierung von außerordentlicher Bedeutung ist, als auch der Signalweg des programmierten Zelltods („Apoptosis“) erscheinen unter den zehn wichtigsten Signalwegen.

7.5 Identifikation von möglichen Zielmolekülkandidaten innerhalb eines Signalwegs

7.5.1 Zielsetzung des vierten Analyseschritts

Die in diesem Abschnitt beschriebenen Analysen stellen den vierten Analyseschritt der vorliegenden Arbeit dar: die Identifizierung von möglichen Zielmolekülkandidaten innerhalb eines Signalwegs. Nachdem in den ersten drei Analyseschritten besonders eng verbundene Kombinationen von therapeutischen Gebieten und Indikationsgruppen sowie die wichtigsten Signalwege dieser Verbindungen identifiziert wurden, geht es im vierten Schritt darum, konkrete Kandidaten für mögliche neue indikationsübergreifende Zielmoleküle zu ermitteln. Zur Eingrenzung der Analyse werden drei Angaben benötigt: die beiden Indikationsgruppen, für die das Zielmolekül nutzbar sein soll, sowie der Signalweg, aus dem die Zielmolekülkandidaten stammen sollen. Die Indikationsgruppen sind in der Regel durch die strategische Ausrichtung einer Firma oder einer Forschungsabteilung bestimmt. Der zu untersuchende Signalweg sollte basierend auf der Signalwegrangliste der Verbindung ausgewählt werden. Dabei gilt, dass neben dem Anteil an der Gesamtverbindungsstärke auch die Höhe des relevanten Anteils des Signalwegs sowie die Anzahl der Zielmoleküle, die die Verbindung stützen, berücksichtigt werden sollte (vergleiche 7.4.3).

Da es im Rahmen der vorliegenden Arbeit primär um die Entwicklung einer Methodik zur Unterstützung von Forschung und Entwicklung, nicht aber um die detaillierte Untersuchung aller Signalwege geht, wurden die Analysen des vierten Schritts exemplarisch für den „Hot-Spot“ des „Hot-Spots“ durchgeführt. Dabei handelt es sich um die am stärksten verbundene Kombination von Indikationsgruppen (Lungenkrebs mit rheumatoider Arthritis) innerhalb der am stärksten verbundenen Kombination von therapeutischen Gebieten (Onkologie mit Bewegungsapparat). Für die Analysen wurde der Signalweg „Apoptosis“ als Beispiel verwendet, der die funktionellen Zusammenhänge des sogenannten programmierten Zelltods oder Apoptose beschreibt. Dieser Signalweg kommt in

der Priorisierung der Verbindungswege auf den neunten Platz und zeichnet sich in dieser Verbindung durch fünf direkt verbindende und zwölf indirekt verbindende Zielmoleküle aus (vergleiche Abb. 7.33).

Der vierte Schritt der Analyse bewegt sich auf der tiefsten Detailebene der Datenbank, den Zielmolekülen und ihren Indikationen. Aufgrund dieser Tatsache kann im Rahmen des vierten Schritts neben der Identifikation von Zielmolekülkandidaten auch eine Überprüfung der Datenbankfunktion auf tiefster Detailebene erfolgen.

7.5.2 Apoptose als Verbindung zwischen Lungenkrebs und rheumatoider Arthritis

Im ersten Schritt der Analyse werden alle Zielmoleküle, die die Verbindung zwischen zwei Indikationsgruppen innerhalb eines Signalwegs unterstützen, namentlich aufgelistet und in drei Klassen unterteilt. Die erste Klasse der Zielmoleküle enthält die Moleküle, die bereits zum jetzigen Zeitpunkt für beide Indikationsgruppen genutzt werden, und damit direkt indikationsübergreifend sind. Die Anzahl der Zielmoleküle dieser Klasse wurde im Rahmen der Signalwegrangliste als „direkt“ verbindend angegeben. Die zweite Klasse fasst diejenigen Moleküle zusammen, die bisher nur in einer der beiden Indikationsgruppen pharmazeutisch genutzt werden. Da sie aber ebenfalls eine funktionelle Rolle in dem verbindenden Signalweg spielen, haben sie eine gute Chance in Zukunft, auch noch für die zweite der beiden Indikationsgruppen eine Rolle zu spielen und damit in die erste Zielmolekülklasse zu rücken. Die Zielmoleküle dieser zweiten Klasse wurden in der Signalwegrangliste als „indirekt verbindend“ bezeichnet und werden im Folgenden auch Kandidatenmoleküle für indikationsübergreifende Funktion genannt. Die dritte Klasse enthält die Moleküle, die zum heutigen Zeitpunkt noch in keiner der beiden Indikationsgruppen aktiv sind, jedoch ebenfalls zum verbindenden Signalweg gehören. Grundsätzlich sind auch diese Moleküle Kandidaten für indikationsübergreifende Funktionen, allerdings ist die Chance dafür etwas geringer, da beide Indikationsgruppen noch neu erschlossen werden müssen. Die Summe der Zielmoleküle der ersten und zweiten Klasse entspricht dem im

Rahmen der Signalwegranglisten ermittelten relevanten Anteils eines Signalwegs für eine Verbindung. Die dritte Gruppe der Zielmoleküle enthält alle die Moleküle, die bisher nicht in den relevanten Anteil fallen.

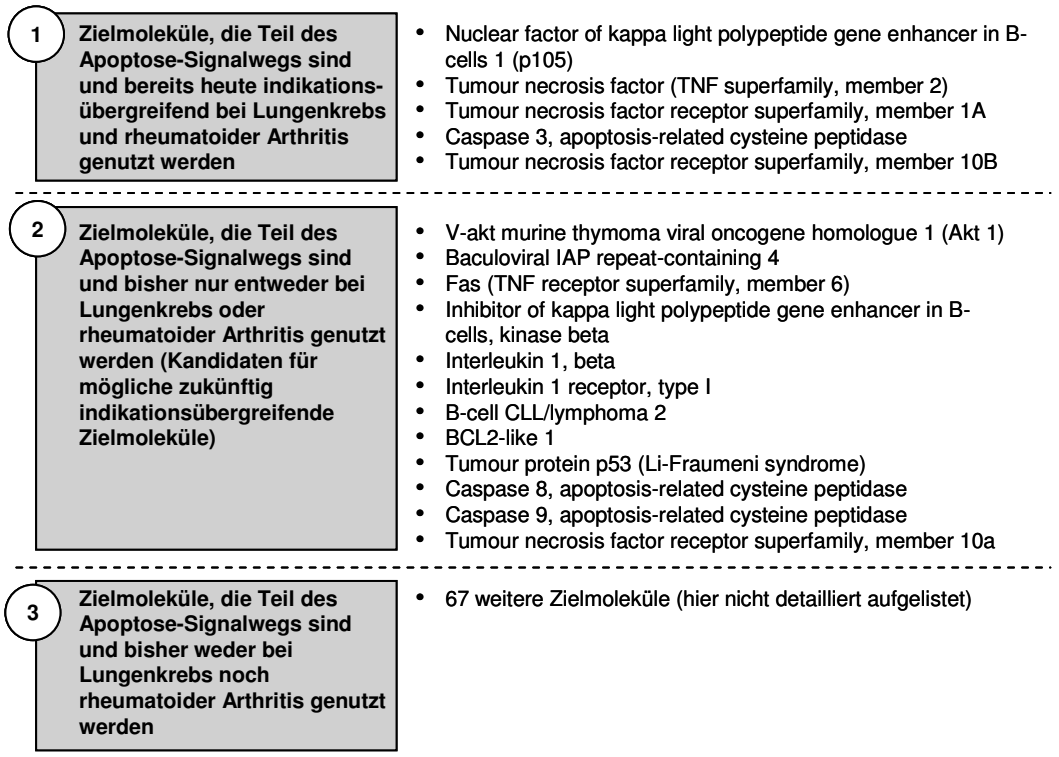


Abb. 7.34: Identifizierung und Klassifizierung der Zielmoleküle des Apoptose-Signalwegs für die Verbindung von Lungenkrebs mit rheumatoider Arthritis
Für die Verbindung der Indikationsgruppen Lungenkrebs und rheumatoide Arthritis wurden alle Zielmoleküle namentlich identifiziert und entsprechend ihrer bisherigen pharmazeutischen Nutzung für eine oder beide Indikationsgruppen in drei Gruppen eingeteilt. Die erste Gruppe enthält die Zielmoleküle, die bereits heute indikationsübergreifend sind, das bedeutet, in beiden Indikationsgruppen genutzt werden. In die zweite Klasse fallen die Zielmoleküle, die bisher nur für eine der beiden Indikationsgruppen genutzt werden. Zur dritten Gruppe gehören die Zielmoleküle, die zwar ebenfalls zum verbindenden Apoptose-Signalweg gehören, bisher aber weder für die Behandlung von Lungenkrebs noch für die Behandlung von rheumatoider Arthritis genutzt werden.
Bei der oben dargestellten Verbindung von Lungenkrebs und rheumatoider Arthritis durch den Apoptose-Signalweg sind fünf Zielmoleküle direkt indikationsübergreifend. Zwölf weitere spielen in einer der beiden Indikationsgruppen eine Rolle und 67 sind weder für Lungenkrebs noch für rheumatoide Arthritis pharmazeutisch genutzt.

Abb. 7.34 zeigt die Klassifizierung der Zielmoleküle für die Verbindung der Indikationsgruppen Lungenkrebs und rheumatoide Arthritis durch den Apoptose-Signalweg. Wie in der Signalwegrangliste für diese Verbindung (Abb. 7.33) bereits ermittelt wurde, sind im Apoptose-Signalweg fünf Zielmoleküle bereits heute sowohl in der Indikationsgruppe Lungenkrebs als auch in der Indikationsgruppe rheumatoide Arthritis aktiv. Zwölf Zielmoleküle werden für eine der beiden Indikationsgruppen genutzt und 67 weitere sind zwar Teil des

Apoptose-Signalwegs, sind aber bisher noch für keine der beiden Indikationsgruppen von Bedeutung.

Nach der namentlichen Identifizierung und Klassifizierung der einzelnen Moleküle wurde eine detaillierte Überprüfung der Ergebnisse durchgeführt. Das bedeutet, dass für jedes ermittelte Zielmolekül geprüft wurde, ob die Angabe von Forschungsprojekten in den Rohdaten mit der Einteilung in die unterschiedlichen Verbindungsklassen übereinstimmt. Eine korrekte Funktion der Datenbankabfragen konnte so auf tiefster Detailebene bestätigt werden.

7.6 Priorisierung von Zielmolekülkandidaten

7.6.1 Zielsetzung der Priorisierung

Endprodukt des im Abschnitt 7.5 beschriebenen vierten Analyseschritts ist eine Liste aller Zielmoleküle innerhalb des ausgewählten Signalwegs, die entsprechend ihrer heutigen pharmazeutischen Nutzung für beide, eine oder keine der untersuchten Indikationsgruppen in drei Klassen eingeteilt wurden. Der fünfte Schritt der Analysen, der im folgenden Abschnitt beschrieben wird, beschäftigt sich nun damit, diese Listen von Molekülen möglichst unmittelbar für pharmazeutische Forschung und Entwicklung nutzbar zu machen.

Abb. 7.35 zeigt eine Zusammenfassung der Ergebnisse der Molekülanalyse mit Überlegungen zur möglichen Anwendung. Die Liste der bereits heute indikationsübergreifenden Zielmoleküle kann direkt, ohne weitere spezifische Analysen von forschenden Einrichtungen zum Abgleich mit der eigenen Forschungs- und Entwicklungspipeline genutzt werden. Forschende Unternehmen haben in der Regel Informationen über bisher vorhandenen Indikationen von Substanzen für ein in ihrem Portfolio vorhandenes Zielmolekül. Eine systematische Analyse über die gesamte Industriepipeline kann jedoch hilfreich sein, neue Indikationen für ein Zielmolekül zu entdecken. Nach firmenspezifischen strategischen Abwägungen kann ein solcher Abgleich zu einer möglichen Erweiterung des Indikationsspektrums von bereits vorhandenen Zielmolekülen führen. Die Erfolgssicherheit dieser neuen Indikationen ist – je nach momentanen Forschungs- und Entwicklungsstatus der Referenzsubstanz – verhältnismäßig hoch, denn ein anderes Unternehmen hat bereits bewiesen, dass es möglich ist, für dieses Zielmolekül ein Medikament mit dieser Indikation zu entwickeln. Entsprechend dieser Tatsache geht die hohe Erfolgssicherheit daher mit einem geringen Neuheitswert einher.

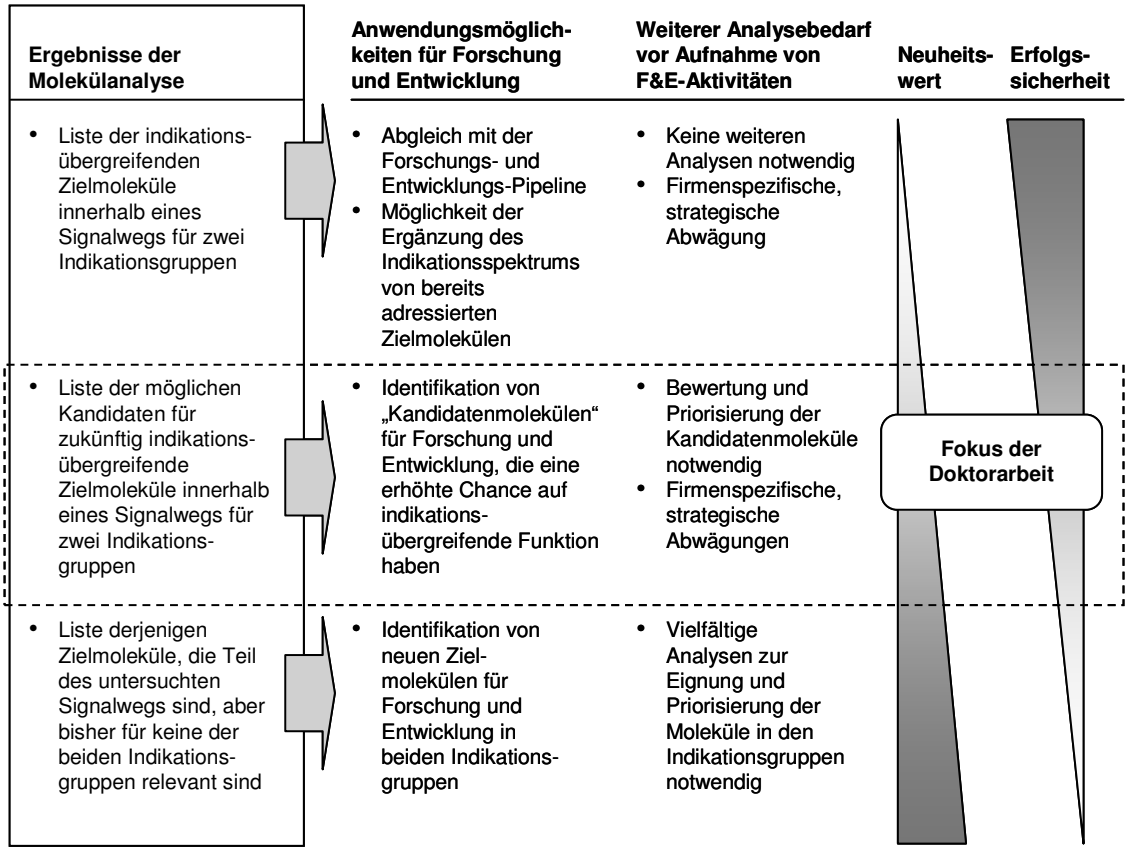


Abb. 7.35: Mögliche Nutzung der Zielmoleküllisten für Forschung und Entwicklung
Die Abbildung zeigt Nutzungsmöglichkeiten für die drei Listen von Zielmolekülen, die im Rahmen der Analyse eines verbindenden Signalwegs für zwei Indikationsgruppen aufgestellt wurden.
Bereits bekannte indikationsübergreifende Zielmoleküle können mit wenig Aufwand und hoher Erfolgssicherheit für Forschung und Entwicklung genutzt werden; allerdings bietet dies wenig Neuheitswert. Die Zielmoleküle, die in der zweiten Liste als Kandidaten für indikationsübergreifende Funktion ermittelt wurden, bieten bei mittlerer Erfolgssicherheit und moderatem Neuheitswert eine interessante Option für Forschung und Entwicklung. Um zu entscheiden, welche Kandidatenmoleküle am besten nutzbar sind, wird eine Priorisierung angeschlossen. Eine indikationsübergreifende Funktion der Zielmoleküle, der dritten Gruppe ist eher unwahrscheinlich (niedrige Erfolgssicherheit aber hoher Neuheitswert).

Mit den Kandidaten für mögliche zukünftige indikationsübergreifende Zielmoleküle verhält es sich anders: Keines der Kandidatenmoleküle wurde bisher indikationsübergreifend genutzt, jedoch konnte für alle gezeigt werden, dass eine pharmazeutische Nutzung in einer der beiden Indikationsgruppen möglich ist. Darüber hinaus sind alle Kandidaten Teil des Signalwegs, der für die Verbindung eine wichtige Rolle spielt und in dem sich auch die bereits heute indikationsübergreifenden Zielmoleküle befinden. Zum „Aufrücken“ in die Klasse der indikationsübergreifenden Zielmoleküle fehlt daher nur die pharmazeutische Aktivität in der Indikationsgruppe, für die das Molekül bisher nicht entwickelt

wird („fehlende Indikationsgruppe“). Die Sicherheit, dass sich ein Kandidatenmolekül indikationsübergreifend nutzen lässt, ist daher geringer für die erste Klasse von Molekülen, jedoch deutlich höher als für die dritte Gruppe derjenigen Moleküle, die bisher für keine der beiden Indikationsgruppen von Bedeutung sind. Im Gegenzug ist der Neuheitswert einer indikationsübergreifenden Funktion der Kandidatenmoleküle höher. Da die Liste der Kandidatenmoleküle je nach Signalweg verhältnismäßig lang und heterogen ist, ist es sinnvoll, eine weitere Priorisierung der Kandidaten durchzuführen, um eine Empfehlung für Forschungs- und Entwicklungskandidaten machen zu können. Eine solche Priorisierung muss außerhalb der Datenbank stattfinden und die Zielmoleküle nach allgemeinen und firmenspezifischen Gesichtspunkten bewerten. Eine Priorisierungslogik wurde im Rahmen der vorliegenden Arbeit entwickelt und wird im Abschnitt 7.6.2 dargestellt.

Die dritte Klasse der Zielmoleküle bietet weitere Kandidaten für Forschung und Entwicklung, jedoch hat keines der Moleküle bisher seine Bedeutung für eine der beiden Indikationsgruppen bewiesen. Bevor Kandidaten aus dieser Gruppe in Forschung und Entwicklung adressiert werden können, müssen vielfältige Analysen zur Bewertung und Priorisierung durchgeführt werden. Da dies den Rahmen dieser Doktorarbeit übersteigt, wurde sich im Folgenden auf die zweite Klasse der Kandidatenmoleküle konzentriert.

7.6.2 Priorisierungslogik

Nach Identifikation der Kandidatenmoleküle für eine mögliche indikationsübergreifende Funktion in den beiden Indikationsgruppen wurde im fünften Schritt eine Priorisierungslogik entwickelt, mit der diejenigen Kandidatenmoleküle priorisiert werden können, die die besten Chancen für indikationsübergreifende Funktion haben.

Zu diesem Zweck werden die Kandidatenmoleküle einzeln bewertet und dann zur Priorisierung miteinander verglichen. Bei der Bewertung der Kandidatenmoleküle ist es wichtig, alle Bereiche abzudecken, die für Erfolg oder Misserfolg der Forschung und Entwicklung an einem Zielmolekül von Bedeutung sein können. Aus diesem Grund erfolgt die Bewertung entlang der folgenden Bereiche:

1. Biologische Evidenz für indikationsübergreifende Funktion – Wie groß ist die Chance für eine Funktion in beiden Indikationsgruppen?
2. Pharmazeutische Machbarkeit – Wie leicht kann für das Zielmolekül ein Medikament entwickelt werden?
3. Wettbewerbsintensität – Ist die Entwicklung eines Medikaments in Anbetracht der Wettbewerbssituation empfehlenswert?
4. Strategische Eignung – Passt das Zielmolekül strategisch ins Portfolio?

Während die ersten drei Bereiche unabhängig von der Firma, die eine mögliche Nutzung des Zielmoleküls erwägen könnte, allgemein bewertet werden können, ist dies für den Bereich der strategischen Eignung nicht der Fall. In wie weit ein Zielmolekül strategisch für Forschung und Entwicklung in einer Firma geeignet ist, kann nicht ohne firmenspezifische Überlegungen bewertet werden.

Für jeden Bereich wurden mir Hilfe von Kernfragen mehrere Bewertungskriterien entwickelt. Für den Bereich der biologischen Evidenz für indikationsübergreifende Funktion wurden die folgenden Bewertungskriterien verwendet:

- Position im Signalweg – Ist das Kandidatenmolekül aufgrund seiner Position im Signalweg geeignet in beiden Indikationsgruppen aktiv zu sein?
- Anhaltspunkte für Funktion in der fehlenden Indikation – Gibt es in der Fachliteratur Hinweise für eine Funktion in der fehlenden Indikation? Wie verlässlich sind diese Hinweise?
- Vernetzungseigenschaft des Zielmoleküls über Signalwege – Hat das Kandidatenmolekül aufgrund seiner Eigenschaften (Knotenpunkt versus spezifischer Teil eines einzelnen Signalwegs) Chancen in mehr als einer Indikation von Bedeutung zu sein?

Für den Bereich der pharmazeutischen Machbarkeit wurden zwei Bewertungskriterien definiert:

- Biologische Eignung des Zielmoleküls für pharmazeutische Intervention – Gibt es biologische Aspekte, die für oder gegen die Eignung des Kandidatenmoleküls als pharmazeutischen Angriffspunkt sprechen?
- Evidenz der Entwicklungsfähigkeit von Medikamenten gegen das Zielmolekül – Gibt es Belege dafür, dass die Entwicklung von Medikamenten gegen das Kandidatenmolekül erfolgreich sein kann?

Die Wettbewerbsintensität wurde mit Hilfe folgender Kriterien bewertet:

- Anzahl und Entwicklungsstatus der bereits existierenden Substanzen für das Zielmolekül – Wie viele Substanzen gibt es bereits, die das Kandidatenmolekül adressieren und wie weit sind diese in Forschung und Entwicklung fortgeschritten?
- Wettbewerberstruktur – Gibt es viele oder wenige Wettbewerber? Wie ist das Portfolio an Substanzen der Wettbewerber für das Zielmolekül?

Für die Bereiche der pharmazeutischen Machbarkeit und der Wettbewerbsintensität kann es dazu kommen, dass abwägende Entscheidungen notwendig werden. Eine gute pharmazeutische Machbarkeit kann zum Beispiel dazu führen, dass viele Firmen Substanzen entwickeln und daher die Wettbewerbssituation eher ungünstig ist.

Auch wenn der vierte Bereich (strategische Eignung) im Gegensatz zu den ersten drei Bereichen nicht allgemein von außen bewertet werden kann, wurden dennoch Bewertungskriterien definiert, die bei möglicher Anwendung der Priorisierungslogik als Richtlinie dienen können:

- Erfahrung und Methodenkompetenz für Signalweg – Kann die Firma Forschungsprojekte im Signalweg erfolgreich bewältigen? Sind die notwendigen Methoden etabliert?
- Erfahrung und Methodenkompetenz für Molekülklasse – Welche Kompetenzen hat die Firma in der Forschung und Entwicklung für die Molekülklasse? Sind die notwendigen Methoden etabliert?

- Existierende Pipeline – Gibt es bereits Substanzen in der Pipeline, die eventuell für das Zielmolekül aktiv sein könnten?
- Einordnung Gesamtstrategie – Passt die Entwicklung von Substanzen für das Kandidatenmolekül zu den strategischen Prioritäten der Firma?

Die überwiegende Anzahl der Bewertungskriterien kann nicht numerisch sondern nur qualitativ bewertet werden. Trotzdem muss die Priorisierungslogik sicherstellen, dass alle Kandidatenmoleküle mit den gleichen Maßstäben bewertet werden. Dies ist von außerordentlicher Bedeutung, damit auch Kandidaten, die aus unterschiedlichen Signalwegen stammen und daher nicht unbedingt im unmittelbaren Vergleich bewertet wurden, miteinander verglichen werden können. Mit anderen Worten heißt das, dass Kandidaten beispielsweise für die übergreifende Funktion in Lungenkrebs und rheumatoider Arthritis eine vergleichbare Bewertung haben müssen, egal ob sie Teil des Signalwegs „Apoptosis“ oder „Calcium signaling pathway“ sind. Um diese vergleichbare Bewertung zu realisieren, wurde für alle Ausprägungen der Bewertungskriterien eine Beschreibung definiert. Dieser Beschreibung muss das Kandidatenmolekül entsprechen, um die Ausprägung zugeordnet zu bekommen. Bei dem Kriterium Anhaltspunkte für Funktion in der fehlenden Indikation zum Beispiel bedeutet dies, dass die Ausprägung „sehr gut“ nur dann vergeben wird, wenn es *in vivo* Evidenz für die Funktion des Zielmoleküls in der fehlenden Indikationsgruppe gibt (zum Beispiel im Tiermodell, vergleiche Abschnitt 6.6.1).

7.6.3 Exemplarische Priorisierung von fünf Molekülen für die Verbindung von Lungenkrebs und rheumatoider Arthritis durch den Apoptose-Signalweg

Um die Funktion der Priorisierungslogik und ihre Hilfe zur Fokussierung auf die am Besten geeigneten Kandidatenmoleküle zu überprüfen, wurden exemplarisch die folgenden fünf Zielmoleküle dem Bewertungsprozess unterzogen:

- Akt1/Proteinkinase B (Originalbezeichnung in der Datenbank: V-akt murine thymoma viral oncogene homologue 1)

- Interleukin 1 β (Originalbezeichnung in der Datenbank: Interleukin 1, beta)
- Interleukin 1 Rezeptor Typ I (Originalbezeichnung in der Datenbank: Interleukin 1 receptor, type I)
- p53 (Originalbezeichnung in der Datenbank: Tumour protein p53, Li-Fraumeni syndrome)
- Death receptor 4/TNFRSF10A (Originalbezeichnung in der Datenbank: Tumour necrosis factor receptor superfamily member 10a)

Die fünf Moleküle waren im Rahmen der Analyse des Apoptose-Signalwegs als Kandidaten für mögliche zukünftige indikationsübergreifende Funktion in rheumatoider Arthritis und Lungenkrebs identifiziert worden. Alle Moleküle sind heute entweder für Lungenkrebs oder rheumatoide Arthritis in Forschung und Entwicklung bzw. zugelassen. Aufgrund der Art der Indikationsbezeichung wurden in der Indikationsgruppe Lungenkrebs die Indikationen „lung cancer“, „small cell lung cancer“ und „non small cell lung cancer“ zusammengefasst. Bei der detaillierten Analyse auf Molekülebene wurde, wo immer möglich, eine weitere Unterscheidung der Lungenkrebsart vorgenommen.

Nach der detaillierten Bewertung jedes Moleküls entlang der Bewertungsmatrizen (siehe Abschnitte 7.6.3.1 bis 7.6.3.3) wurde eine zusammenfassende Gesamtbewertung erstellt (siehe Abschnitt 7.6.3.4), die die untersuchten Moleküle vergleicht. Aufgrund der notwendigen firmenspezifischen Aspekte wurde der Bereich der strategischen Eignung nicht bewertet.

7.6.3.1 Detaillierte Bewertung der biologischen Evidenz für indikations-übergreifende Funktion

Position im Signalweg

Zur Bewertung der Position im Signalweg wird der Frage nachgegangen, ob die Zielmoleküle aufgrund ihrer Position im Signalweg eine erhöhte oder erniedrigte Chance haben, in beiden Indikationsgruppen relevant zu sein. Wie in Abschnitt 6.6.1 erläutert ist, wurde dazu untersucht, ob das Zielmolekül in funktioneller Nähe (zum Beispiel „upstream“ oder „downstream“) eines Moleküls liegt, das

bereits heute in beiden Indikationen eine Rolle spielt oder für die fehlende Indikation relevant ist. Ein solcher funktioneller Zusammenhang könnte darauf hindeuten, dass das Kandidatenmolekül ebenfalls eine gute Chance hat, als Vorläufer oder Effektor des bereits wichtigen Moleküls in der fehlenden Indikation eine Rolle zu spielen. Im Gegensatz dazu könnte die Position in einem separaten Unterbereich des Signalwegs nahelegen, dass das Zielmolekül zwar Teil des Signalwegs ist, aber nicht unbedingt eine indikationsübergreifende Funktion haben muss. Abb. 7.36 zeigt das Diagramm des Apoptose-Signalwegs, so wie er in der KEGG Datenbank (Kanehisa Laboratories, 2008) dargestellt wird.

Zielmoleküle, die bereits heute in den Indikationsgruppen Lungenkrebs und rheumatoider Arthritis relevant sind, wurden mit einem roten Oval gekennzeichnet. Kandidatenmoleküle, die bisher nur für Lungenkrebs genutzt werden, haben ein blaues Oval und Moleküle, deren Funktion nur für rheumatoide Arthritis bekannt ist, sind gelb gekennzeichnet.

Für die Kandidatenmoleküle wurde die Position im Signalweg bewertet und entsprechend der Bewertungsmatrix wurde die Ausprägung des Bewertungskriteriums in Form eines „Mondes“ vergeben (vergleiche Abschnitt 6.6.1). Ein voller Mond stellt eine optimale Ausprägung entlang des Bewertungskriteriums für indikationsübergreifende Funktion dar, während ein leerer Mond indiziert, dass das Kandidatenmolekül das Kriterium gar nicht erfüllt und damit nur geringe Chancen für indikationsübergreifende Funktion bestehen.

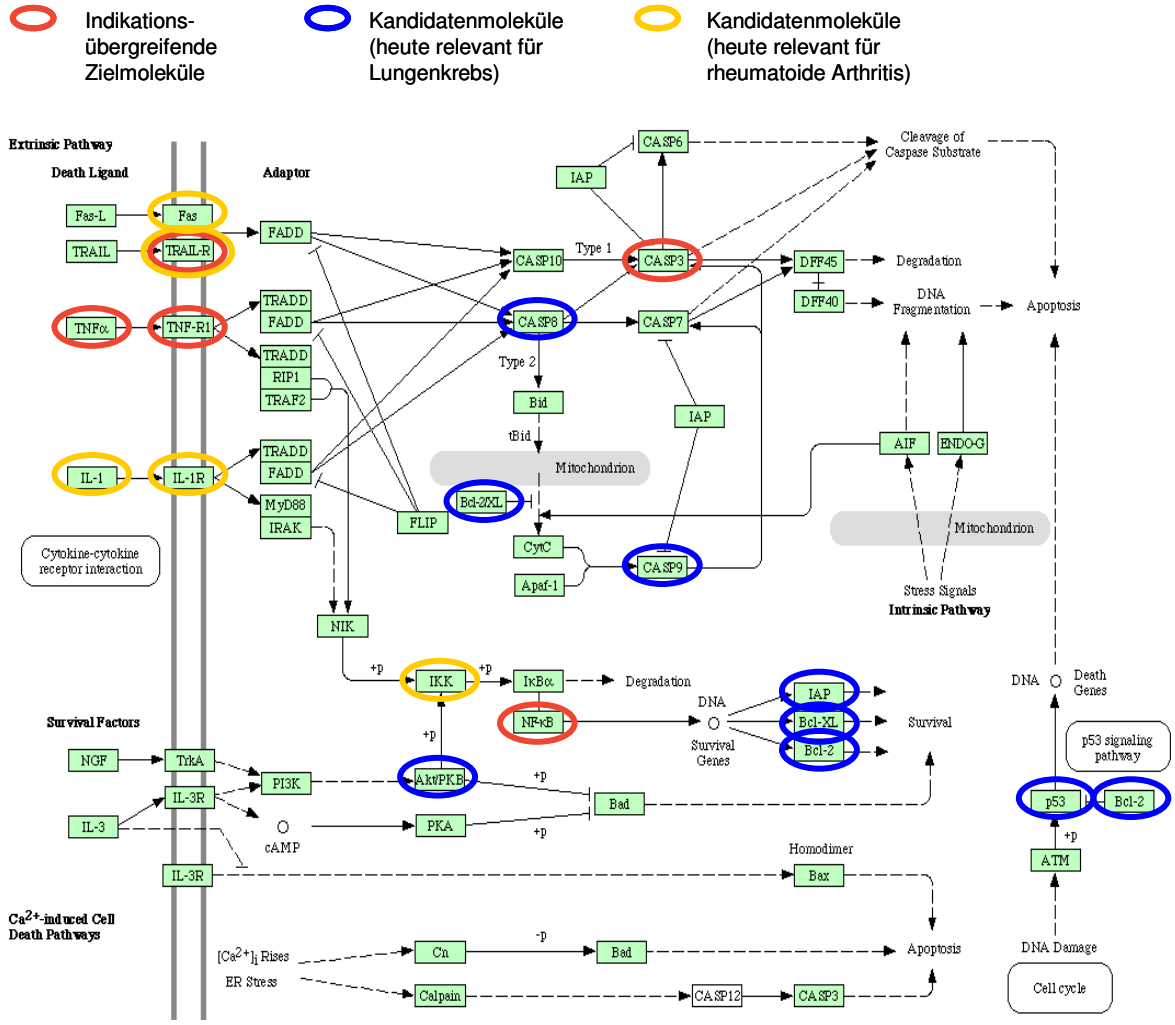


Abb. 7.36: Apoptose-Signalweg in der Darstellung der KEGG Datenbank (Kanehisa Laboratories, 2008)
Die Abbildung zeigt die Darstellung des Apoptose-Signalwegs der KEGG Datenbank. Moleküle, die sowohl für Lungenkrebs als auch rheumatoide Arthritis eine Rolle spielen, sind mit einem roten Oval gekennzeichnet. Kandidaten für indikationsübergreifende Funktion sind mit einem blauen oder gelben Oval umrandet, je nachdem für welche der beiden Indikationsgruppen sie bisher pharmazeutisch adressiert werden (siehe Legende der Abbildung).

- **Akt1**:

  Akt1 ist indirekt mit NFκB verbunden, welches ein indikationsübergreifendes Zielmolekül ist. Darüber hinaus steht Akt1 in direktem funktionellen Zusammenhang mit dem IKK (Inhibitor of kappaB kinase beta) Komplex, der bereits heute für rheumatoide Arthritis eine Rolle spielt. Durch die Nähe zu indikationsübergreifenden Molekülen sowie solchen, die in der fehlenden Indikation eine Rolle spielen, hat Akt1 eine gute Chance für indikationsübergreifende Funktion.

- **Interleukin 1 β**:

  Interleukin 1 β ist genau wie TNFα, TNF-R1 und TRAIL (alle drei bereits heute indikationsübergreifende Zielmoleküle) Teil des ligandenvermittelten, extrinsischen Apoptose-Signalwegs. Interleukin 1 β hat keine direkte funktionelle Beziehung zu heute schon indikationsübergreifenden Molekülen oder Molekülen, die für Lungenkrebs relevant sind. Interleukin 1 β kann durch den Interleukin 1 Rezeptor Typ I „downstream" TRADD und FADD aktivieren und verwendet daher einen zum indikationsübergreifenden TNF-Weg analogen Signalprozess.

- **Interleukin 1 Rezeptor Typ I:**

  Für den Interleukin 1 Rezeptor Typ I gilt aufgrund der direkten Liganden-Rezeptor-Beziehung zwischen Interleukin 1 β und Interleukin 1 Rezeptor Typ I die gleiche Bewertung. Wie Interleukin 1 β hat der Interleukin 1 Rezeptor Typ I keinen unmittelbaren funktionellen Zusammenhang mit indikationsübergreifenden Zielmolekülen und/oder einem für Lungenkrebs relevanten Molekül. Allerdings gilt auch hier, dass sich durch die zum TNF-Weg analoge Kaskade (FADD, TRADD) eine Chance für indikationsübergreifende Funktion bietet.

- **p53:**

  p53 ist Kernbestandteil des durch DNA-Schaden ausgelösten Apoptose-Signalwegs. Es hat jedoch keinen funktionellen Zusammenhang mit Unterbereichen des Apoptose-Signalwegs, die indikationsübergreifende Zielmoleküle enthalten. Unmittelbarer funktioneller Zusammenhang findet sich nur mit Molekülen, die nur für Lungenkrebs, nicht aber für rheumatoide Arthritis relevant sind (zum Beispiel Bcl-2).

- **Death receptor 4:**

  Der Tumor necrosis factor receptor superfamily member 10A (TNFRSF10A, alias Death receptor 4) ist einer der beiden bekannten funktionellen TRAIL Rezeptoren und steht zu Beginn des ligandenvermittelten Apoptose-Signalwegs, in dem bereits einige indikationsübergreifende Zielmoleküle vorhanden sind. Der zweite bekannte funktionelle Todesrezeptor

Death receptor 5 (DR5 oder TNFRSF10B) ist bereits heute für Lungenkrebs und rheumatoide Arthritis in Forschung und Entwicklung und steht an der gleichen Stelle des Signalwegs. Die Isoform Death receptor 4 hat also – basierend auf der Position im Signalweg – eine gut begründete Chancen auf eine indikationsübergreifende Funktion.

Anhaltspunkte für Funktion in der fehlenden Indikation

Das zweite Bewertungskriterium für den Bereich der biologischen Evidenz für indikationsübergreifende Funktion sind Anhaltspunkte für die Funktion in der fehlenden Indikation. Bei der Bewertung entlang dieses Kriteriums wird vor allem danach gesucht, ob es in der wissenschaftlichen Literatur bereits Veröffentlichungen zu einer Rolle des Moleküls in der fehlenden Indikation gibt. Dabei wird unterschieden, ob es sich um theoretische Betrachtungen oder experimentelle Ergebnisse handelt. Gibt es experimentelle Evidenz für eine Funktion, so wird diese unterschiedlich bewertet, je nachdem ob es sich um ein *in vitro* Experiment oder zum Beispiel ein Tiermodell der entsprechenden Erkrankung handelt. Insgesamt spielt für die Bewertung der Ausprägung dieses Kriteriums auch die Anzahl an vorhandener Literatur eine Rolle. Mit anderen Worten: gibt es nur einen Autor, der in einer Veröffentlichung eine Funktion des Moleküls postuliert, wird die Evidenz für eine Funktion in der fehlenden Indikation weniger positiv bewertet, wie wenn zahlreiche Veröffentlichungen dies belegen. Selbstverständlich ist keine quantitative Beurteilung der Literatur möglich, jedoch kann durch Verwendung einer standardisierten Suchautomatik ein vergleichbares Ergebnis für alle untersuchten Moleküle erzielt werden (vergleiche Abschnitt 6.6.1).

- **Akt1 in rheumatoider Arthritis:**

  Akt1 ist bereits für Lungenkrebs in Forschung und Entwicklung, daher wurden die Anzeichen für eine Funktion bei rheumatoider Arthritis als fehlende Indikation untersucht. Es konnten 13 relevante Veröffentlichungen zur Funktion von Akt1/PKB im Zusammenhang mit rheumatoider Arthritis gefunden werden, davon zwei Reviews (Liu und Pope, 2003; Marone u. a., 2008). Zahlreiche *in vitro* Assays weisen auf eine mögliche Funktion von

Akt1 im Zusammenhang mit rheumatoider Arthritis hin. Diese wurden vor allem mit Zellkulturen von humanen synovialen Fibroblasten („human fibriblast-like synoviocytes“, FLS), die von Patienten mit rheumatoider Arthritis gewonnen wurden, durchgeführt (Migita u. a., 2000; Zhang u. a., 2001; Kim u. a., 2002; Miyashita u. a., 2003; Hwang u. a., 2004; Morel u. a., 2005; Chen u. a., 2006).

Folgende deskriptive Erkenntnisse wurden bisher beschrieben: Phosphorylierte Akt1 ist in synovialem Gewebe mit rheumatoider Arthritis überexprimiert (Liu und Pope, 2003) und FLS aus rheumatoidem Gewebe weisen eine spezifische, Akt-vermittelte Inaktivierung von Transkriptionsfaktoren der FoxO Familie auf (Ludikhuize u. a., 2007).

Veröffentlichungen mit funktionellen Ergebnissen postulieren, dass Akt1 durch mehrere Mechanismen mitverantwortlich für mangelnde Apoptose bei rheumatoider Arthritis ist. Eine der beschriebenen Ursachen ist die Beobachtung, dass Akt1 die TNFα vermittelte Inhibition der Apoptose in rheumatoiden FLS steuert (Zhang u. a., 2001; Liu und Pope, 2003). Darüber hinaus schützt Akt1 vor der durch TGFβ/Fas ausgelösten Apoptose in rheumatoiden FLS (Kim u. a., 2002; Liu und Pope, 2003). Diese Effekte führen zu einer verminderten Apoptose von FLS und damit zu einem der Kernsymptome der rheumatoiden Arthritis. Zusätzlich verhindert Akt1 die durch TRAIL vermittelte Apoptose (Miyashita u. a., 2003; Morel u. a., 2005). Im Zusammenhang mit dem in der Literatur beschriebenen proliferativen Effekt von TRAIL auf die apoptose-resisitenten Zellen stellen die Autoren die Hypothese auf, dass Akt1 nicht nur das Auslösen der Apoptose hemmt, sondern auch den nachfolgenden proliferativen Effekt von TRAIL auf die überlebenden Zellen vermittelt (Morel u. a., 2005). Eine andere Publikation konnte belegen, dass C2-Ceramide Akt1 und Erk1/2 inhibieren und für Apoptose bei rheumatoiden FLS sorgen können (Migita u. a., 2000). Genauso wurde eine Apoptose-Induktion durch Inhibition von Akt1 durch Thalidomid beschrieben (Gockel u. a., 2004). Auch dies lässt eine Beteiligung von Akt in der FLS-Apoptose vermuten.

Neben dem anti-apoptotischen Effekt wurde eine Beteiligung von Akt1 an der Überexpression der Interleukine IL17, IL6 und IL8 beschrieben (Hwang u. a., 2004; Kim u. a., 2005). Diese sind an der inflammatorischen Situation der rheumatoiden Arthritis beteiligt. Während die oben beschriebenen funktionellen Erkenntnisse für Akt1 in Zellkultur gewonnen wurden, konnte die Bedeutung der PI3K (Akt1 aktivierende Kinase) für rheumatoide Arthritis bereits im Tiermodell gezeigt werden (Ludikhuize u. a., 2007; Marone u. a., 2008).

- **Interleukin 1 β und Interleukin 1 Rezeptor Typ I in Lungenkrebs**:

  Interleukin 1 β und der Interleukin 1 Rezeptor Typ I werden heute bereits therapeutisch zur Behandlung von rheumatoider Arthritis eingesetzt. Daher wurden die Chancen einer Funktion in der Indikationsgruppe Lungenkrebs untersucht. Beide Moleküle wurden aufgrund ihrer direkten Liganden-Rezeptor-Beziehung gemeinsam behandelt. Es konnten 14 relevante Veröffentlichungen für die mögliche Funktion des IL-1 Signalwegs (Interleukin 1 β, Interleukin 1 Rezeptor Typ I oder auch Interleukin 1 Rezeptor Antagonist) bei Lungenkrebs identifiziert werden; darunter befindet sich ein Review (Lewis u. a., 2006). Es wurde nach Hinweisen für alle Lungenkrebsformen gesucht, jedoch wo immer möglich eine Unterscheidung zwischen kleinzelligem Lungenkarzinom („small cell lung cancer“, SCLC) und nicht-kleinzelligem Lungenkarzinom („non small cell lung cancer“, NSCLC) gemacht. Die Untersuchungen der Veröffentlichungen beziehen sich auf verschiedene Tiermodelle *in vivo* (Lewis Lung Cancer, B16 Melanom Lungenmetastasen, DA/3 Mammaadenokarzinom mit Lungenmetastasen, verschiedene KO-Mäuse) (Voronov u. a., 2003; Weinreich u. a., 2003; Yano u. a., 2003; Nakao u. a., 2005; Elaraj u. a., 2006; Bunt u. a., 2007; Lavi u. a., 2007), *in vitro* Zellkultur (LLC, A549 Zelllinien) (Song u. a., 2003; Voronov u. a., 2003; Yano u. a., 2003; Nakao u. a., 2005; Zhu u. a., 2006), sowie auf Proben von Patienten mit NSCLC und SCLC (Zienolddiny u. a., 2004; Lind u. a., 2005; Hu u. a., 2006; Engels u. a., 2007).

Im Tiermodell und in Zellkultur wurde beobachtet, dass Interleukin 1 β invasives Tumorwachstum, Neovaskularisierung und Metastasierung fördert (Lewis u. a., 2006). Im Einzelnen konnten die Arbeitsgruppen feststellen, dass Mäuse mit überexprimiertem Interleukin 1 β oder Interleukin 1-Rezeptor-Antagonist-KO-Mäuse ein stärkeres Tumorwachstum haben (Voronov u. a., 2003; Weinreich u. a., 2003; Bunt u. a., 2007). Im Gegensatz dazu ist das Tumorwachstum bei Interleukin 1-Rezeptor-Typ I-KO-Mäusen verringert (Bunt u. a., 2007). In ähnlicher Weise führt die Inhibition der Interleukin 1 β Wirkung durch Applikation des Interleukin 1 Rezeptor Antagonisten zu Hemmung von Tumorwachstum und der Vaskularisierung im Mausmodell (B16 Melanom Lungenmetastasen und humane NSCLC Xenografts und humane Melanom Xenografts in Nacktmäusen) (Voronov u. a., 2003; Weinreich u. a., 2003; Elaraj u. a., 2006; Lavi u. a., 2007). In Zellkultur führt eine Überexpression von Interleukin 1 β zu verschiedenen Effekten auf Adhäsions- und Invasionsfähigkeit der Tumorzellen, die in unmittelbarem Zusammenhang zu Metastasierung und Neovaskularisierung stehen (Yano u. a., 2003). Die Arbeitsgruppe vermutet, dass eine indirekte Wirkung auf die Mikroumgebung von Tumoren für dieses Phänomen verantwortlich ist (Nakao u. a., 2005): Um den Tumor herum kommt es im Mausmodell zu Akkumulation von MDSC („myeloid-derived suppressor cells“) (Bunt u. a., 2007). Dies führt in SCLC-Zellkultur zur Produktion des angiogenetischen Faktors GCP-2 („granulocyte-chemotactic protein 2“), der die Neovaskularisierung fördert (Zhu u. a., 2006). Zusätzlich wird die Rekrutierung von COX-2 exprimierenden Makrophagen berichtet (Nakao u. a., 2005). In den Veröffentlichungen wird eine unterschiedliche Wirkung von Interleukin 1 β und Interleukin 1 α betont. Während Interleukin 1 β sezerniert wird und in Reaktion auf externe Stimuli tumorfördernde Effekte hat, ist Interleukin 1 α membranständig oder zytosolisch und hat eine tumorprotektive Wirkung (Song u. a., 2003; Voronov u. a., 2003).

Untersuchungen von Proben von Lungenkrebspatienten und Zellkultur-Versuche liefern deskriptive Erkenntnisse: Bei Patienten mit NSCLC mit inflammatorischem Zustand konnte ein erhöhter Interleukin 1 β Spiegel

nachgewiesen werden (Ilonidis u. a., 2006). Darüber hinaus war die Interleukin 1 β mRNA in mehr als der Hälfte aller von einer Arbeitsgruppe untersuchten metastasierenden humanen Tumoren (inkl. NSCLC) hoch exprimiert (Elaraj u. a., 2006). Zum Teil konnte dieselbe Arbeitsgruppe Tumorzellen mit konstitutiv aktivem Interleukin 1 nachweisen. Vor allem für die entzündliche Komponente des Lungenkrebs scheint Interleukin 1 β durch die Aktivierung anderer entzündlicher Mediatoren (G-CSF und PTHrP, „parathyroid hormone related protein“) mitverantwortlich zu sein. Diese Beobachtungen wurden an Zelllinien von Lungenkrebspatienten mit Plattenepithelkarzinomen gemacht (Zellinie OK-C-1) (Uemura u. a., 2000).

Genetische Untersuchungen konnten Polymorphismen im Interleukin 1 β Gen und im Interleukin 1 Rezeptor Antagonist (IL-1RA) Gen feststellen, die einen Einfluss auf die Wahrscheinlichkeit an NSCLC zu erkranken haben. Insgesamt bringen drei Veröffentlichungen Interleukin 1 β Polymorphismen mit erhöhter Wahrscheinlichkeit für NSCLC in Zusammenhang (Zienolddiny u. a., 2004; Lind u. a., 2005; Engels u. a., 2007). Im Bezug auf das Interleukin 1 Rezeptor Antagonist Gen sind die Ergebnisse nicht einheitlich: Eine Veröffentlichung zeigt eine Reduktion der Lungenkrebswahrscheinlichkeit bei Vorhandensein eines bestimmten Allels des Interleukin 1 Rezeptor Antagonist Gens vor allem für Raucher mit Veränderungen des Plattenepithels und Nichtraucher (Hu u. a., 2006). Bei einem anderen Polymorphismus wird eine erhöhte Wahrscheinlichkeit für die Erkrankung an Lungenkrebs beschrieben (Lind u. a., 2005). Darüber hinaus gibt es Veröffentlichungen, die keinen Zusammenhang zum Erkrankungsrisiko sehen (Engels u. a., 2007). Unabhängig davon gibt es Beobachtungen, dass Patienten mit Polymorphismen im Interleukin 1 β Gen eine erhöhte Wahrscheinlichkeit für p53 Mutationen in Tumorzellen haben (Zienolddiny u. a., 2004).

- **p53**:

Das Tumorsuppressorprotein p53 ist bisher in einigen Projekten unter anderem für Lungenkrebs (SCLC und NSCLC) in Forschung und Entwicklung. Für die

Bewertung der indikationsübergreifenden Funktion wurde untersucht, ob es Anhaltspunkte dafür gibt, dass p53 auch für rheumatoide Arthritis eine therapeutische Rolle spielen könnte. Insgesamt gibt es sehr viel Literatur (85 Treffer in PubMed ab 2000), aus der 35 relevante Veröffentlichungen identifiziert werden konnten. Darunter befinden sich einige Reviews (Michael und Alisa, 2000; Tak u. a., 2000; Yamanishi und Firestein, 2001; Sun und Cheung, 2002; Baier u. a., 2003; Liu und Pope, 2003; Yamasaki u. a., 2007). Im Vergleich zur Literatur zu Akt1, Interleukin 1 β und zum Interleukin 1 Rezeptor Typ I wird in den einzelnen Veröffentlichung deutlich mehr über mögliche Therapien diskutiert. Im Vergleich zu den anderen Zielmolekülen scheint die Forschung für p53 bei rheumatoider Arthritis bereits deutlich näher an der praktischen Umsetzung zu sein. Für funktionelle Untersuchungen wurden vor allem Kaninchen- und Mausmodelle verwendet. Darüber hinaus wurden Patientenproben für deskriptive Untersuchungen sowie für *in vitro* Zellkulturuntersuchungen benutzt.

Zusammenfassend lässt sich folgende Forschungshypothese formulieren: p53 ist nicht Auslöser für rheumatoide Arthritis, spielt aber im Verlauf eine wichtige Rolle und könnte daher ein therapeutisches Ziel sein (Baier u. a., 2003). Rheumatoide Arthritis stellt eine chronische Entzündung mit oxidativem Stress dar, der im Verlauf der Erkrankung zu genotoxischen Schäden – unter anderem an p53 – führt (Michael und Alisa, 2000; Tak u. a., 2000). Diese Schäden leiten den Übergang von Entzündung zu Autoimmunkrankheit ein (Simelyte u. a., 2005). Zahlreiche Arbeitsgruppen zeigen, dass p53 in synovialen Fibroblasten entweder mutiert oder nicht ausreichend aktiv ist (Sun und Cheung, 2002; Baier u. a., 2003; Liu und Pope, 2003). Zwei mögliche Therapieansätze direkt an p53 werden daher diskutiert: Gentransfer von funktionellem p53 (Pap u. a., 2001; Yao u. a., 2001; Sun und Cheung, 2002; Yao u. a., 2003) oder Aktivierung von endogenem p53 durch Proteasomeninhibition (Migita u. a., 2001; Yang u. a., 2005; Brun, 2008).

Die oben beschriebene Forschungshypothese wird vor allem durch funktionelle Erkenntnisse aus Kaninchen- und Mausmodellen gestützt (Yao u. a., 2001; Yamanishi u. a., 2002a; Leech u. a., 2003; Simelyte u. a., 2005;

Leech u. a., 2008): In allen Tiermodellen haben p53-/- Tiere einen schwereren Krankheitsverlauf als Vergleichstiere (Yamanishi u. a., 2002a; Leech u. a., 2008). Ähnlich wird bei Patienten mit rheumatoider Arthritis eine Korrelation der Krankheitsschwere mit Stärke der Überexpression von p53 beobachtet (Yatsugi u. a., 2000; Salvador u. a., 2005). Vorhandenes p53 lässt sich durch Proteasomeninhibition aktivieren. Dies leitet die Apoptose im Synovialgewebe ein (Migita u. a., 2001). In ähnlicher Weise führt die intraartikuläre Injektion von Adenoviren, die p53 in die Zellen bringen, im Kaninchen-Modell für rheumatoide Arthritis *in vivo* zu Apoptose und Minderung der Krankheitssymptome (Yao u. a., 2001). Umgekehrt führen humane Papilloma (HP)-Viren über Inhibition von p53 via E6-Protein im Tiermodell zur Verschlimmerung der Krankheitssymptome (Pap u. a., 2001).

Deskriptive Erkenntnisse durch Untersuchungen an Synovialgewebe von Patienten mit rheumatoider Arthritis sind nicht ganz so eindeutig. Vor allem durch Untersuchung des p53-Spiegels in peripheren Blutzellen von Patienten mit rheumatoider Arthritis gibt es widersprüchliche Daten (El-Sayed u. a., 2003; Abou-Shousha u. a., 2005). Im Gegensatz dazu konnten verschiedene Arbeitsgruppen konsistent einen erhöhten p53-Spiegel in aktiven FLS, beziehungsweise in Synovialflüssigkeit (Sun und Cheung, 2002; El-Sayed u. a., 2003; Abou-Shousha u. a., 2005; Maas u. a., 2005; Salvador u. a., 2005) sowie ein reduziertes Level von Mdm2 (Taranto u. a., 2005) in FLS von Patienten mit rheumatoider Arthritis feststellen. Zur Verteilung von p53 im Synovialgewebe gibt es unterschiedliche Meinungen: Zum Teil wird eine Überexpression vor allem in den Stromazellen unter dem Synovialepithel, dem sogenannten „sub-lining“, beschrieben (Chou u. a., 2001); zum Teil beschränkt sich die Beobachtung von konstitutiver Überexpression auf die oberste Zellschicht der Synovia, das sogenannte „synovial lining“ und dehnt sich nur im akuten Proliferationsschub bis ins Stroma aus (Helmchen u. a., 2005). Untersuchungen durch Implantationen von Fibroblasten aus Patienten mit rheumatoider Arthritis ins Mausmodell zeigen ein erhöhtes p53-Niveau an den Stellen, in denen das Synovialgewebe in den Knorpel invasiv einwächst (Seemayer u. a., 2003). Untersuchungen mit Bestrahlung von Zellen von

Patienten mit rheumatoider Arthritis mit γ-Strahlen konnten demonstrieren, dass die Aktivierung von p53 gestört ist (Maas u. a., 2005).

Genetische Untersuchungen bei Patienten mit rheumatoider Arthritis konnten Mutationen im p53 Gen nachweisen (Inazuka u. a., 2000; Yamanishi und Firestein, 2001; Sun und Cheung, 2002; Yamanishi u. a., 2002a; Yamanishi u. a., 2002b) – zum Teil sind diese übereinstimmend mit den Mutationen, die bei Tumoren detektiert werden können (Yamanishi und Firestein, 2001). Polymorphismen, insbesondere der Polymorphismus „72“ im p53-Gen, stehen allerdings nicht im Zusammenhang mit dem Auftreten von rheumatoider Arthritis (Lee u. a., 2001), wohl aber mit der Schwere der Erkrankung (Lee u. a., 2001; Macchioni u. a., 2007). Mutationen im p53-Gen treten vor allem im „synovial lining“ und seltener in tieferen Zellschichten auf (Yamanishi u. a., 2002b). Zellkulturuntersuchungen konnten keinen Unterschied in p53 Mutationen bei rheumatoider Arthritis mit oder ohne Gelenkerosion feststellen (Yamanishi u. a., 2005), so dass ein Zusammenhang mit dem invasiven Wachstum in den Gelenkknorpel unklar bleibt.

Neben p53 Mutationen und Unter- bzw. Überexpression legen einige Veröffentlichungen auch Störungen in der Regulation von p53 im Zusammenhang mit rheumatoider Arthritis nahe. Eine Hypothese ist, dass eine verstärkte Aktivierung der Ubiquitinligase Synoviolin, die bei „ER-Stress“, (zum Beispiel Überfüllung des endoplamatischen Retikulums) aktiviert wird und unter anderem p53 für den Abbau im Proteasom vorbereitet, eine Rolle spielt. Die Autoren vermuten einen Zusammenhang („Crosstalk“) zwischen der Überfüllung des ERs und p53 (Yamasaki u. a., 2007). Andere Veröffentlichungen finden eine Überexpression des Tumor-Protein S100A4 bei Patienten mit rheumatoider Arthritis und vermuten dadurch eine Stabilisierung von p53 (Klingelhofer u. a., 2007). Als möglicher Ansatz für eine indirekt p53 betreffende Therapie wird ein Angriff an CBP („cAMP response element binding protein binding protein“) diskutiert, welches unter anderem für die Acetylierierung von p53 zuständig ist (Nakazawa u. a., 2002; Nakajima u. a., 2004). Darüber hinaus werden funktionelle Zusammenhänge von TRAIL (Yang u. a., 2005) und Fas (Itoh u. a., 2004) mit p53 in FLS

gesehen. Beide induzieren p53-abhängig Apoptose in den FLS. Fas ist darüber hinaus noch im mitochondrialen Apoptoseweg beteiligt (Itoh u. a., 2004). Zuletzt postuliert eine weitere Arbeitsgruppe einen noch unklaren funktionellen Zusammenhang von p53 und dem für die FLS-Apoptose wichtigen MIF („macrophage migration inhibitory factor“) (Leech u. a., 2003).

Aufgrund der Fülle an Literatur und Untersuchungen muss die Evidenz der biologischen Funktion von p53 bei rheumatoider Arthritis sehr positiv bewertet werden.

- **Death receptor 4**:

  Der TRAIL-Rezeptor Death receptor 4 (DR4) ist bisher mit einem Projekt zur Behandlung von NSCLC in Phase II der klinischen Entwicklung. Für den Death receptor 4 wurden daher Anhaltspunkte zu einer Funktion in rheumatoider Arthritis gesucht. Dabei konnten 11 Veröffentlichungen gefunden werden, die TRAIL bzw. den Death receptor 4 und rheumatoide Arthritis unmittelbar in Zusammenhang bringen.

  Zusammenfassend gilt, dass es keine Hinweise darauf gibt, dass der Death receptor 4 als Zielmolekül zur Behandlung von rheumatoider Arthritis geeignet sein könnte. Die Rolle des Death receptor 4 ist sehr umstritten und existierende Therapievorschläge im Rahmen des TRAIL-Signalings betreffen entweder den Death receptor 5 oder TRAIL direkt.

  Untersuchungen mit funktionellen Beobachtungen gibt es vor allem zu TRAIL selbst und zum Death receptor 5: Das Einbringen von TRAIL in synovialen Fibroblasten führt bei einem Teil der FLS zu Apoptose (Morel u. a., 2005; Terzioglu u. a., 2007) und verbessert im Tiermodell von rheumatoider Arthritis bei intra-artikulärer Injektion das Krankheitsbild (Song u. a., 2000; Yao u. a., 2003). Ebenso wird durch Applikation von stimulierenden Antikörpern am Death receptor 5 der programmierte Zelltod eingeleitet (Ichikawa u. a., 2003). Neben den funktionellen TRAIL-Rezeptoren Death receptor 4 und Death receptor 5 gibt es noch sogenannte Decoy-Rezeptoren, deren cytoplasmatischen Domänen so verkürzt sind, dass sie zu

keiner Effektorstimulation führen (Kimberley und Screaton, 2004). Es konnte *in vitro* gezeigt werden, dass eine Zugabe von Decoy-Rezeptoren durch Abfangen von TRAIL protektiv auf FLS gegenüber TRAIL-vermittelter Apoptose wirkt (Miyashita u. a., 2004). Außerdem scheint die Expression der Decoy-Rezeptoren mit der Resistenz gegenüber TRAIL-vermittelter Apoptose zu korrelieren (Terzioglu u. a., 2007).

Auch deskriptive Erkenntnisse in Bezug auf die Rezeptorexpression sprechen nicht eindeutig für eine Bedeutung des Death receptor 4 im Zusammenhang mit rheumatoider Arthritis. Zum Teil wird der Death receptor 4 auf FLS bei Patienten mit rheumatoider Arthritis nachgewiesen (Ichikawa u. a., 2003; Park u. a., 2003), zum Teil berichten Arbeitsgruppen Negativ-Ergebnisse (Perlman u. a., 2003). Expressionsuntersuchungen für den Death receptor 5 sind genauso wenig eindeutig: Zum Teil wird eine Überexpression bei FLS-Zelllinien von Patienten mit rheumatoider Arthritis beschrieben (Ichikawa u. a., 2003; Miranda-Carus u. a., 2004), zum Teil kann der Rezeptor überhaupt nicht nachgewiesen werden (Park u. a., 2003; Perlman u. a., 2003).

Genetische Untersuchungen zeigen keinen Zusammenhang zwischen Auftreten von rheumatoider Arthritis und Polymorphismen im Death receptor 4-Gen (Yazdani-Biuki u. a., 2008).

Als therapeutische Ansätze für rheumatoide Arthritis werden Interventionen am Death receptor 5 oder TRAIL diskutiert: Agonistische Anti-DR5-Antikörper lösen *in vitro* Apoptose aus. In diesem Zusammenhang werden jedoch mögliche Nebenwirkungen (vor allem Hepatotoxizität und Störung der negativen T-Zell Selektion im Thymus) als Problem angesprochen (Ichikawa u. a., 2003; Tsokos und Tsokos, 2003; Miranda-Carus u. a., 2004). Alternativ könnte die Apoptose im rheumatoiden Gewebe durch Gentransfer von TRAIL induziert werden, sofern funktionelle DR4/5 Rezeptoren vorhanden sind (Song u. a., 2000; Yao u. a., 2003; Terzioglu u. a., 2007). Eine Inhibition von Decoy-Rezeptoren wird ebenfalls diskutiert und könnte mögliche Resistenzen von FLS gegen TRAIL-vermittelte Apoptose aufheben (Terzioglu u. a., 2007). Auch bei diesem therapeutischen Ansatz werden mögliche Nebenwirkungen

(vor allem TRAIL-stimulierte Proliferation der apoptose-resistenten FLS) als Problem gesehen (Morel u. a., 2005). Eine weitere Arbeitsgruppe diskutiert eine Gen-modifizierte Zelltherapie als Ansatz zur Behandlung von rheumatoider Arthritis (Liu u. a., 2003; Tsokos und Tsokos, 2003). Nach Gentransfer kann im Mausmodell mit Hilfe von Doxycyclin die Produktion von TRAIL selektiv in Kollagen-spezifischen dendritischen Zellen induziert werden. Dies führt im Modell zu verminderter Schwere der rheumatoiden Arthritis

Vernetzungseigenschaft des Zielmoleküls über Signalwege

Im Gegensatz zum vorherigen Bewertungskriterium kann die Vernetzungseigenschaft des Zielmoleküls mit Hilfe der Datenbank ermittelt werden. Die grundlegende Annahme für dieses Bewertungskriterium ist, dass ein Molekül, welches in mehreren Signalwegen vorkommt, eine größere Chance hat, in unterschiedlichen Indikationen Bedeutung zu haben.

- **Akt1:**

  Akt1 ist das im Vergleich der fünf untersuchten Zielmoleküle am stärksten vernetzende Molekül. Die Bewertung mit einem vollen Mond bezieht sich jedoch auf die Bewertungsmatrix und nicht auf den relativen Vergleich. Akt1 ist in 14 Signalwegen relevant und hat daher die Eigenschaften eines Netzwerkknotens. Akt1 verknüpft die folgenden Signalwege: „MAPK signaling pathway“, „ErbB signaling pathway“, „mTOR signaling pathway“, „Apoptosis“, „VEGF signaling pathway“,“ Focal adhesion“, „Tight junction“, „Toll-like receptor signaling pathway“, „Jak-STAT signaling pathway“, „T cell receptor signaling pathway,“ „B cell receptor signaling pathway“, „Fc epsilon RI signaling pathway,“ „Insulin signaling pathway“ und „Adipocytokine signaling pathway“.

- **Interleukin 1 β:**

  Interleukin 1 β ist in fünf Signalwegen („Toll-like receptor signaling pathway”, „MAPK signaling pathway”, „Hematopoietic cell lineage”,

„Cytokine-cytokine receptor interaction", „Apoptosis") relevant und damit ein moderat vernetzendes Protein.

- **Interleukin 1 Rezeptor Typ I:**

  Der Interleukin 1 Rezeptor Typ I ist Teil von vier Signalwegen („MAPK signaling pathway", „Hematopoietic cell lineage", „Cytokine-cytokine receptor interaction", „Apoptosis") und wird daher genauso wie sein Ligand Interleukin 1 β als moderat vernetzend eingestuft.

- **p53:**

  Das Tumorsuppressorprotei p53 spielt in fünf Signalwegen („Wnt signaling pathway", „p53", „MAPK signaling pathway", „Cell cycle", „Apoptosis") eine Rolle und ist daher ebenfalls ein moderater Knotenpunkt.

- **Death receptor 4:**

  Der TRAIL-Rezeptor Death receptor 4 bzw. TNFRSF10A kommt nur in drei Signalwegen („Natural killer cell mediated cytotoxicity", „Cytokine-cytokine receptor interaction", „Apoptosis") vor und ist daher nur gering vernetzend.

7.6.3.2 Detaillierte Bewertung der pharmazeutischen Machbarkeit

Schwierigkeiten bei der Medikamentenentwicklung entscheiden über Erfolg oder Scheitern eines Forschungsprojekts und bestimmen die Kosten und die Dauer der Medikamentenentwicklung. Bei Auswahl eines Zielmoleküls sollten daher absehbare Schwierigkeiten berücksichtigt werden. Allerdings muss eine schwierige Entwicklungssituation nicht zwangsweise gegen die Entwicklung sprechen, wenn die biologische Wirkung überzeugend ist und die Firma Zugang zu dem notwendigen Wissen und den benötigten Methoden hat. Häufig gehen Hürden bei der Machbarkeit mit einer vorteilhaften Wettbewerbssituation einher.

Biologische Eignung des Zielmoleküls für pharmazeutische Intervention

Dieses Bewertungskriterium berücksichtigt biologische Aspekte, die Erfolg oder Scheitern des Moleküls in Forschung und Entwicklung bestimmen könnten. Dabei werden die Art der Zielmolekülklasse, die Lage in der Zelle (zum Beispiel als

Rezeptor auf der Membran), mögliche Therapieansätze sowie die Menge des vorhandenen Wissens über das Zielmolekül berücksichtigt.

- **Akt1**:

  Akt1 ist seit 1991 bekannt (Jones u. a., 1991) und konnte als Protein nachgewiesen werden. Sowohl Sekundär- als auch Tertiärstruktur von Akt1 sind bekannt und mit Hilfe von mutiertem Akt1 wurden Kristallstrukturuntersuchungen durchgeführt. Die Proteindomänen des Moleküls wurden experimentell und über Ähnlichkeit charakterisiert. Es gibt zahlreiche experimentelle Erkenntnisse zu funktionellen und strukturellen Mutationen.

  Akt1 liegt im Zytoplasma vor und wird durch Plekstrin-Homologie-(PH)-Domäne an die Zellmembran rekrutiert. Durch Aktivierung ist eine Translokation in den Zellkern möglich. Diese Verteilung in der Zelle ermöglicht die Verwendung aller therapeutischen Ansätze, die zum Angreifen von intrazellulären Proteinen geeignet sind (unter anderem „small molecules“, RNAi, therapeutische Viren, Ceramide) (Yang u. a., 2004; Manning und Cantley, 2007).

  Akt1 ist eine Serin/Threonin-Kinase. Kinasen dieser Art sind häufig Zielmoleküle in der pharmazeutischen Industrie.

- **Interleukin 1 β:**

  Interleukin 1 β wurde bereits 1984 charakterisiert (Auron u. a., 1984). Auch für dieses Zielmolekül wurde die Existenz auf Proteinebene bewiesen (Dinarello, 1994). Die Sekundär- und Tertiärstruktur sind bekannt; die Kristallstruktur wurde vollständig experimentell bestimmt (Priestle u. a., 1988; Finzel u. a., 1989; Priestle u. a., 1989). Die Bindestelle des Zytokins für den Rezeptor wurde experimentell charakterisiert und die Spaltungsstelle des Vorläuferproteins zum funktionsfähigen Interleukin 1 β ebenfalls experimentell bestimmt. Am Aminoende des Vorläuferproteins befindet sich eine Sequenz, der regulatorische Aktivität zugesprochen wird. Allerdings ist unklar, um welche Funktion es sich dabei handelt.

Interleukin 1 β wird als Zytokin in den Extrazellulärraum sezerniert. Dadurch ist das Angreifen des Moleküls verhältnismäßig einfach durch verschiedene therapeutische Ansätze möglich (Antikörper, Chelatoren, Impfstoffe etc.). Nachteil ist, dass durch einen Angriff am extrazellulären Zytokin selbst ein zellspezifisches Ausrichten der Therapie (zum Beispiel an bestimmten Charakteristika eines Organs) nicht gut möglich ist. Eine weitere Option für therapeutische Intervention ist ein Eingriff in den stark regulierten Prozess des sogenannten „Processings" (Herausschneiden des funktionsfähigen Moleküls aus dem Vorläufer) und der Sekretion.

Zytokin-Signalwege sind häufig therapeutisches Ziel der pharmazeutischen Industrie, jedoch werden bisher vor allem die Rezeptoren durch Medikamente adressiert.

- **Interleukin 1 Rezeptor Typ I**:

Der Interleukin 1 Rezeptor Typ I wurde 1989 charakterisiert (Sims u. a., 1989) und konnte auf Proteinebene nachgewiesen werden. Im Komplex mit dem Liganden Interleukin 1 β wurde die vollständige Kristallstruktur experimentell bestimmt (Vigers u. a., 1997). Die Charakterisierung der funktionellen Domänen erfolgte überwiegend experimentell und zusätzlich über Homologie-Untersuchungen. Einige Mutationen und natürliche Varianten sind bekannt und in ihrer Funktion charakterisiert.

Als Zytokinrezeptor befindet sich das Protein in der Plasmamembran. Die Liganden-Bindedomäne ragt in den Extrazellulärraum und intrazellulär befindet sich eine Domäne zur Rekruitierung eines Adapterproteins für den weiteren Signaltransduktionsprozess (Huang u. a., 1997; Vigers u. a., 2000). Verschiedene therapeutische Ansätze sind möglich, unter anderem auch mit Substanzen, die aufgrund ihrer Größe und Struktur zum Teil nicht in Zellen eindringen können (zum Beispiel Antikörper, Rezeptorantagonisten).

Das Interleukin 1 Rezeptor Typ I Molekül ist ein typisches Zielmolekül für pharmazeutische Intervention. Es gibt mehr Erfahrung für den Angriff am Rezeptor als am Zytokin selbst und es ist möglich, bestimmte

Zellpopulationen gezielt zu erreichen (zum Beispiel durch spezifische Antikörperdomänen).

- **p53**:

  Das Tumorsuppressorprotein p53 wurden 1984 charakterisiert (Crawford u. a., 1984; Lamb und Crawford, 1986; Lane, 1992) und seine Existenz wurde auf Proteinebene bewiesen. Die dreidimensionale Struktur des Proteins und seine Effektordomänen, regulatorische Domänen und Strukturelemente wurden experimentell bestimmt und umfangreich charakterisiert. Unzählige somatische Mutationen im p53-Gen sind bei sporadisch auftretenden Tumoren beschrieben.

  p53 kann im Zellkern, im Zytosol und im Endoplasmatischen Retikulum nachgewiesen werden; die aktive Form befindet sich im Zellkern. Durch diese Lokalisation sind komplexe therapeutische Ansätze erforderlich, die auf den Zellkern zielen. Meist geht es bei der pharmazeutischen Intervention um das Einbringen einer funktionellen Kopie des p53-Gens in den Zellkern (zum Beispiel durch viralen Gentransfer, Impfstoffe, Gentherapie). Diese Ansätze sind durch die zu verwendenden Vektoren mit erheblichem Nebenwirkungsrisiko verbunden (zum Beispiel Immunreaktion auf viralen Vektor) (Wilson, 2002; Roth und Grammer, 2004).

  Transkriptionsfaktoren sind pharmazeutisch eher schwierige Zielmoleküle. Bisher gibt es nur wenig praktische Erfahrung mit erfolgreichem Gentransfer.

- **Death receptor 4:**

  Im Vergleich zu den vorherigen Zielmolekülen ist der Death receptor 4 deutlich weniger gut charakterisiert. Das Molekül wurde 1997 bekannt (Schneider u. a., 1997a; Schneider u. a., 1997b) und es gibt Hinweise für seine Existenz auf Proteinebene (Kimberley und Screaton, 2004). Bisher wurde keine Kristallstruktur veröffentlicht. Die funktionellen Regionen des Death receptor 4 wurden zum Teil experimentell und zum Teil durch Homologie-Untersuchungen bestimmt.

Als TRAIL-Rezeptor befindet sich der Death receptor 4 in der Plasmamembran. Er hat eine extrazelluläre Ligandenbindungsdomäne sowie eine intrazelluläre Domäne zur Rekrutierung von Effektorproteinen. Durch die Lokalisation sind neben intrazellulären Ansätzen, die die Effektordomäne betreffen, auch extrazelluläre Ansätze möglich (zum Beispiel Antikörper).

Insgesamt gibt es im Bereich der Pharmaindustrie viel Erfahrung mit Rezeptorproteinen im Allgemeinen – das Wissen über die TRAIL-Rezeptoren (insbesondere über Death receptor 4) ist jedoch begrenzt.

Evidenz der Entwicklungsfähigkeit von Medikamenten gegen das Zielmolekül

Bei diesem Bewertungskriterium wird untersucht, ob es Hinweise darauf gibt, dass eine Firma bereits erfolgreich Substanzen entwickelt hat, die das zu untersuchende Zielmolekül adressieren. Dabei ist es nicht so sehr von Bedeutung, in welcher Indikation diese anderen Substanzen aktiv sind. Im Vordergrund steht vielmehr die Frage, ob und wie weit ein Entwicklungsprozess bereits in der Praxis durchlaufen wurde. Hintergrund dieses Kriteriums ist die Tatsache, dass auch ein in der biologischen Theorie gut geeignetes Zielmolekül in der Praxis zu unterwarteten Komplikationen führen kann (zum Beispiel Problem mit der Toxizität in Phase I). Wenn eine Firma jedoch bereits gezeigt hat, dass es möglich ist, Substanzen präklinisch und klinisch zu entwickeln, stehen die Chancen verhältnismäßig gut, dass dies auch für andere Indikationen möglich ist.

- **Akt1:**

  In der Datenbank können 41 Forschungsprojekte in vielen verschiedenen Indikationsgruppen ermittelt werden, die Akt1 als Zielmolekül adressieren. Dabei erklärt sich die hohe Projektzahl vor allem dadurch, dass viele verschiedene Tumortypen in Projekten vertreten sind. Projekte außerhalb der Onkologie sind tendenziell in frühen Phasen (vor allem Präklinik).

  Es gibt keine zugelassenen Substanzen oder Projekte in Phase III.

  Phase II: Es gibt 27 Projekte in Phase II; davon 26 im therapeutischen Gebiet Onkologie in sechs verschiedenen Indikationsgruppen (Prostatakrebs, Brustkrebs, kolorektale Tumore, Leukämie, Lungenkrebs und andere

Tumore). Ein weiteres Phase II-Projekt liegt im Grenzgebiet zwischen Hämatologie und Onkologie (Myelodysplastisches Syndrom).

Phase I: Es gibt vier Projekte im therapeutischen Gebiet Onkologie in drei verschiedenen Indikationsgruppen (Leukämien, Lungenkrebs und andere Tumore).

Präklinik: Es lassen sich zehn präklinische Projekte identifizieren. Sieben davon sind im Bereich Onkologie (Brustkrebs, Prostatakrebs, Leukämien, kolorektale Tumoren und andere Tumoren); zwei weitere Projekte gehören zum therapeutischen Gebiet Stoffwechsel und Endokrinologie (Indikationsgruppen Diabetes und Fettleibigkeit) und ein Projekt lässt sich dem therapeutischen Gebiet ZNS (Alzheimer) zuordnen.

- **Interleukin 1 β**:

Mit Hilfe der Datenbank können neun Projekte in insgesamt sechs verschiedenen Indikationsgruppen ermittelt werden, die Interleukin 1 β adressieren und verteilt auf fünf therapeutische Gebiete sind. Die überwiegende Anzahl der Projekte befindet sich in frühen Entwicklungsstadien, das heißt noch in der Präklinik oder vor Phase II.

Es gibt keine zugelassenen Projekte.

Phase III: Es gibt ein Projekt im therapeutischen Gebiet Transplantationsmedizin und Immunsystem in der Indikationsgruppe der generalisierten Immunerkrankungen (Muckle-Wells Syndrom).

Phase II: Auch hier ist nur ein Projekt im Gebiet des Bewegungsapparats (Indikationsgruppe rheumatoide Arthritis) vorhanden.

Phase I: Drei Phase I-Projekte in drei verschiedenen therapeutischen Gebieten (Augenheilkunde: Indikationsgruppe AMD Behandlung; Stoffwechsel und Endokrinologie: Indikationsgruppe Diabetes sowie respiratorisches System: Indikationsgruppe COPD) erweitern das Indikationsspektrum von Interleukin 1 β.

Präklinik: Es gibt vier Projekte, die alle zum Bereich des Bewegungsapparats gehören (Indikationsgruppen rheumatoide Arthritis und Osteoarthritis).

- **Interleukin 1 Rezeptor Typ I**:

  Für den Interleukin 1 Rezeptor Typ I gibt es zwar in Summe weniger Projekte als für die bisher bewerteten Zielmoleküle, dafür sind die Projekte über alle Entwicklungsphasen verteilt und zugelassene Medikamente beweisen, dass die Entwicklung einer Substanz durch alle klinischen Phasen möglich ist. Insgesamt können in der Datenbank sieben Projekte in drei verschiedenen Indikationsgruppen in den therapeutischen Gebieten Bewegungsapparat und Stoffwechsel und Endokrinologie identifiziert werden.

  Ein Projekt ist bereits vollständig entwickelt und die entsprechende Substanz ist zugelassen. Es handelt sich dabei um ein Medikament gegen rheumatoide Arthritis (therapeutisches Gebiet Bewegungsapparat).

  Phase III: Es gibt keine Projekte in Phase III, jedoch zwei Projekte, die sich kurz vor der Zulassung („pre-registration“) befinden. Beide sind im therapeutischen Gebiet Transplantationsmedizin und Immunsystem einzuordnen und richten sich gegen Indikationen im Bereich generalisierte Immunerkrankungen („Familial cold autoinflammatory syndrome“ und Muckel-Wells Syndrom).

  Phase II: Hier gibt es zwei Projekte in den therapeutischen Gebieten Bewegungsapparat (Indikationsgruppe rheumatoide Arthritis) sowie Stoffwechsel und Endokrinologie (Indikationsgruppe „andere Stoffwechselerkrankungen“; hier Indikation Hyperuricämie).

  Phase I: Die Datenbank zeigt ein Phase I-Projekt im Gebiet Bewegungsapparat (Indikationsgruppe rheumatoide Arthritis).

  Präklinik: Es gibt ein Projekt im therapeutischen Gebiet Stoffwechsel und Endokrinologie (Indikationsgruppe andere Stoffwechselerkrankungen: hier Indikation Hyperuricämie).

- **p53:**

  31 Projekte richten sich gegen p53 und sind vor allem in der Onkologie sowie in angrenzenden Gebieten angesiedelt. Ein Projekt, welches sich noch in der Präklinik befindet, gehört zum therapeutischen Gebiet Urogenitalsystem.

Insgesamt sind alle Entwicklungsstadien vertreten und es gibt ein bereits zugelassenes Projekt.

Ein Projekt im Bereich Onkologie (Indikation: Kopf-Hals Karzinome) ist bereits zugelassen und hat die gesamte Entwicklung erfolgreich durchlaufen.

Phase III: Es gibt ein Projekt in Phase III im therapeutischen Gebiet Onkologie (Indikation: Kopf-Hals Karzinome) sowie ein weiteres Projekt, das sich vor der Zulassung („pre-registration“) gegen Li-Fraumeni Syndrom (Grenzgebiet Onkologie und Stoffwechsel und Endokrinologie) befindet.

Phase II: 14 Projekte befinden sich in dieser Phase. Damit ist die Mehrzahl aller p53 Projekte in Phase II. Bei allen Projekten handelt es ich um Projekte im therapeutischen Gebiet Onkologie in den Indikationsgruppen Brustkrebs, Lungenkrebs, Leukämie, Prostatakrebs und andere Tumoren.

Phase I: Es gibt zwölf Projekte in der Onkologie verteilt auf die Indikationsgruppen Brustkrebs, Lungenkrebs, kolorektale Tumore, Prostatakrebs und andere Tumore sowie ein Projekt im Grenzbereich zwischen Hämatologie und Onkologie (Myelodysplastisches Syndrom).

Präklinik: Zwei Projekte befinden sich in der Präklinik, eines davon im therapeutischen Bereich Onkologie (Leberkarzinom) und eines im therapeutischen Gebiet Urogenitalsystem (Nierenversagen).

- **Death receptor 4**:

Für den TRAIL-Rezeptor Death receptor 4 gibt es momentan zehn Projekte in Forschung und Entwicklung. Alle davon befinden sich im therapeutischen Bereich Onkologie und in frühen Entwicklungsstadien.

Es gibt noch keine zugelassenen Projekte oder Projekte in Phase III.

Phase II: Es gibt fünf Projekte in Phase II im therapeutischen Bereich Onkologie, die auf die folgenden Indikationsgruppen verteilt sind: Lymphome, Kolorektale Tumoren, Lungenkrebs (NSCLC) sowie andere Tumore (hier: Myelom).

Phase I: Mit Hilfe der Datenbank können fünf Phase I-Projekte aus dem Gebiet Onkologie in den Indikationsgruppen Prostatakrebs und anderen Tumoren (hier: Ösophaguskarzinom, Pankreastumore, Nierentumore und Schilddrüsetumore) identifiziert werden.

Für den Death receptor 4 gibt es keine präklinischen Projekte.

7.6.3.3 Detaillierte Bewertung der Wettbewerbsintensität

Anzahl und Status der bereits existierenden Substanzen für das Zielmolekül

Im Rahmen dieses Bewertungskriteriums wird untersucht, ob und wie viele Substanzen es bereits für das Zielmolekül gibt. Eine hohe Substanzzahl deutet an, dass eine mögliche neue Substanz eventuell starken Wettbewerb ausgesetzt sein könnte, sofern sie sich nicht deutlich von den anderen Substanzen differenziert.

- **Akt1**:

  Bisher gibt es zehn verschiedene Substanzen, die für Akt1 aktiv sind. Fünf davon sind noch im präklinischem Stadium für diverse Indikationen auch außerhalb von Lungenkrebs. Die anderen fünf sind mindestens für eine Indikation schon in Phase II (Behandlung von kolorektalen Tumoren, Prostatakrebs, Brustkrebs, Lungenkrebs, Leukämien und anderen Tumoren).

- **Interleukin 1 β:**

  Das Zytokin wird erst von drei verschiedenen Substanzen angegriffen, von denen eine noch ausschließlich in präklinischer Entwicklung für nur eine Indikation (rheumatoide Arthritis) ist. Die zweite Substanz befindet sich für Diabetes bereits in Phase I und in präklinischer Entwicklung für verschiedene Indikationen im Gebiet Bewegungsapparat. Die dritte Substanz ist weiter in der Entwicklung fortgeschritten und hat Indikationen in den Phasen III, II und I und ein breites Indikationsspektrum in verschiedenen therapeutischen Gebieten (Bewegungsapparat, Stoffwechsel und Endokrinologie, Augenheilkunde und respiratorisches System).

- **Interleukin 1 Rezeptor Typ I:**

  Im Gegensatz zum Liganden Interleukin 1 β ist die Wettbewerbsituation für den Rezeptor Interleukin 1 Rezeptor Typ I weniger positiv. Es sind fünf verschiedene Substanzen bekannt; davon ist eine bereits für rheumatoide Arthritis zugelassen. Zwei sind in Phase II bzw. Phase I der klinischen Entwicklung für rheumatoide Arthritis. Eine Substanz steht kurz vor der Zulassung für verschiedene Autoimmunerkrankungen und befindet sich in Phase II für Hyperuricämie. Die letzte Substanz ist bisher ausschließlich in präklinischer Entwicklung für Hyperuricämie.

- **p53:**

  Für das Tumorsuppressorprotein gibt es heute schon acht aktive Substanzen. Eine davon ist zur Behandlung von Kopf-Hals Tumoren zugelassen. Eine weitere ist für verschiedene Indikationen im Gebiet Onkologie in klinischer Entwicklung in den Phasen I, II und III (bzw. kurz vor Zulassung „pre-registration"). Drei Substanzen befinden sich in klinischer Entwicklung mit mindestens einer Indikation im Bereich Onkologie oder in ähnlichen Indikationen in Phase II und eine Substanz ist für eine einzelne Indikation in Phase I (ebenfalls im Bereich Onkologie). Die verbleibenden zwei Substanzen sind bisher nur in präklinischer Entwicklung in den therapeutischen Gebieten Onkologie und Urogenitalsystem.

- **Death receptor 4:**

  Die Wettbewerbssituation für den Death receptor 4 ist sehr positiv, da es bisher nur zwei Substanzen gibt. Eine davon ist in Phase II zur Behandlung des Non-Hodgkin-Lymphoms, die andere ist mit zahlreichen onkologischen Indikationen in den Phasen I und II.

Wettbewerberstruktur

Im Rahmen der Bewertung von außen („outside-in") wird die Wettbewerberstruktur dann für einen neuen Teilnehmer als vorteilhaft eingestuft, wenn keiner der Wettbewerber über mehr als eine Konkurrenzsubstanz verfügt, es noch keine Lizenzaktivitäten gibt und keine großen Pharmakonzerne mit

dementsprechend großem Forschungs- und Entwicklungsbudget involviert sind (vergleiche Abschnitt 6.6.3). Diese Einstufung stellt nur eine Annäherung dar und kann – je nach Firma oder Forschungsgruppe – anders beurteilt werden. Für bestimmte Firmen kann ein aktiver Pharmakonzern eine Bedrohung sein, für andere aber auch ein strategischer Entwicklungspartner.

- **Akt1:**

  Neun Firmen sind unabhängig an der Entwicklung von Substanzen beteiligt. Außer einem größeren Pharmakonzern handelt es sich primär um kleinere Biotech- und Pharmaunternehmen. Nur eine Firma hat zwei verschiedene Substanzen im Portfolio; alle anderen haben nur eine Substanz. Für zwei Phase II-Substanzen gibt es Lizenzaktivitäten.

- **Interleukin 1 β:**

  Drei Firmen haben jeweils eine Substanz in ihrem Portfolio. Bei einer der Firmen handelt es sich um einen großen Pharmakonzern. Dieses Unternehmen hat darüber hinaus die stärkste Wettbewerbsposition, da seine Substanz für mehrere Indikationen in klinischer Entwicklung ist. Es gibt bisher keine Lizenzaktivitäten.

- **Interleukin 1 Rezeptor Typ I:**

  Vier Firmen sind in der Entwicklung von Substanzen aktiv; davon sind drei größere Konzerne. Einer hat zwei Substanzen in der Pipeline, alle anderen haben nur eine. Eine der Substanzen wurde unter Lizenz bereits an zwei weitere Firmen gegeben.

- **p53:**

  Für p53 sind acht Firmen sind als Entwickler aktiv, von denen jede nur eine Substanz hat. Es sind keine großen Pharmakonzerne involviert. Für zwei der präklinischen Substanzen gibt es Lizenzaktivitäten.

- **Death receptor 4:**

  Zwar entwickeln bisher nur zwei Firmen Substanzen, jedoch handelt es sich bei einer um einen großen Pharmakonzern und es gibt bereits Lizenzaktivitäten zu drei weiteren großen Firmen.

7.6.3.4 Gesamtbewertung

Nachdem in den vorherigen Abschnitten die fünf Moleküle, die als Beispiele zur Anwendung der Priorisierungslogik ausgewählt worden waren, im Einzelnen bewertet wurden, gilt es nun eine Gesamtbewertung zu erstellen. Die Gesamtbewertung soll dazu dienen, aus den fünf Kandidatenmolekülen des Apoptose-Signalwegs diejenigen auszuwählen, die die größte Chance haben, in den beiden Indikationsgruppen Lungenkrebs und rheumatoide Arthritis von Bedeutung zu sein. Es muss erwähnt werden, dass – wie bereits einleitend beschrieben – im Rahmen dieser Arbeit nur die von außen bewertbaren Bereiche zu einer Entscheidung zu Rate gezogen werden können („outside-in"-Perspektive). Für eine finale Priorisierung der Kandidatenmoleküle muss die anwendende Firma oder Forschungsgruppe nicht nur den vierten Bereich der strategischen Eignung berücksichtigen, sondern auch eine etwaige Abwägung zwischen den Bereichen pharmazeutische Machbarkeit und Wettbewerbsintensität vornehmen.

Abb. 7.37 zeigt den Vergleich der Bewertung der Zielmoleküle Interleukin 1 β und Interleukin 1 Rezeptor Typ I, die bisher nur in rheumatoider Arthritis aktiv sind und für die die Chancen für eine Relevanz im Indikationsbereich Lungenkrebs bewertet wurden. Insgesamt lässt sich feststellen, dass sowohl Interleukin 1 β als auch der Interleukin 1 Rezeptor Typ I eine gute Chance haben, sowohl bei kleinzelligem Lungenkarzinom (SCLC) als auch bei nicht-kleinzelligem Lungenkarzinom (NSCLC) therapeutisch genutzt zu werden. Die biologische Evidenz für eine Funktion wird bei beiden Molekülen identisch bewertet, da eine Unterteilung aufgrund ihrer unmittelbaren Liganden-Rezeptor-Beziehung mit Hilfe der vorhandenen Literatur nicht möglich war. Im Bereich der pharmazeutischen Machbarkeit schneidet der Rezeptor besser als der Ligand ab – dafür gibt es zumindest momentan noch weniger Wettbewerb in der direkten

Behandlung des Liganden Interleukin 1 β. Die Entscheidung, ob eher der Rezeptor oder der Ligand in ein Forschungs- und Entwicklungsportfolio aufgenommen werden sollte, muss daher mit Hilfe von strategischen Überlegungen und einem Abwägen der Bereiche pharmazeutische Machbarkeit versus Wettbewerbsintensität getroffen werden.

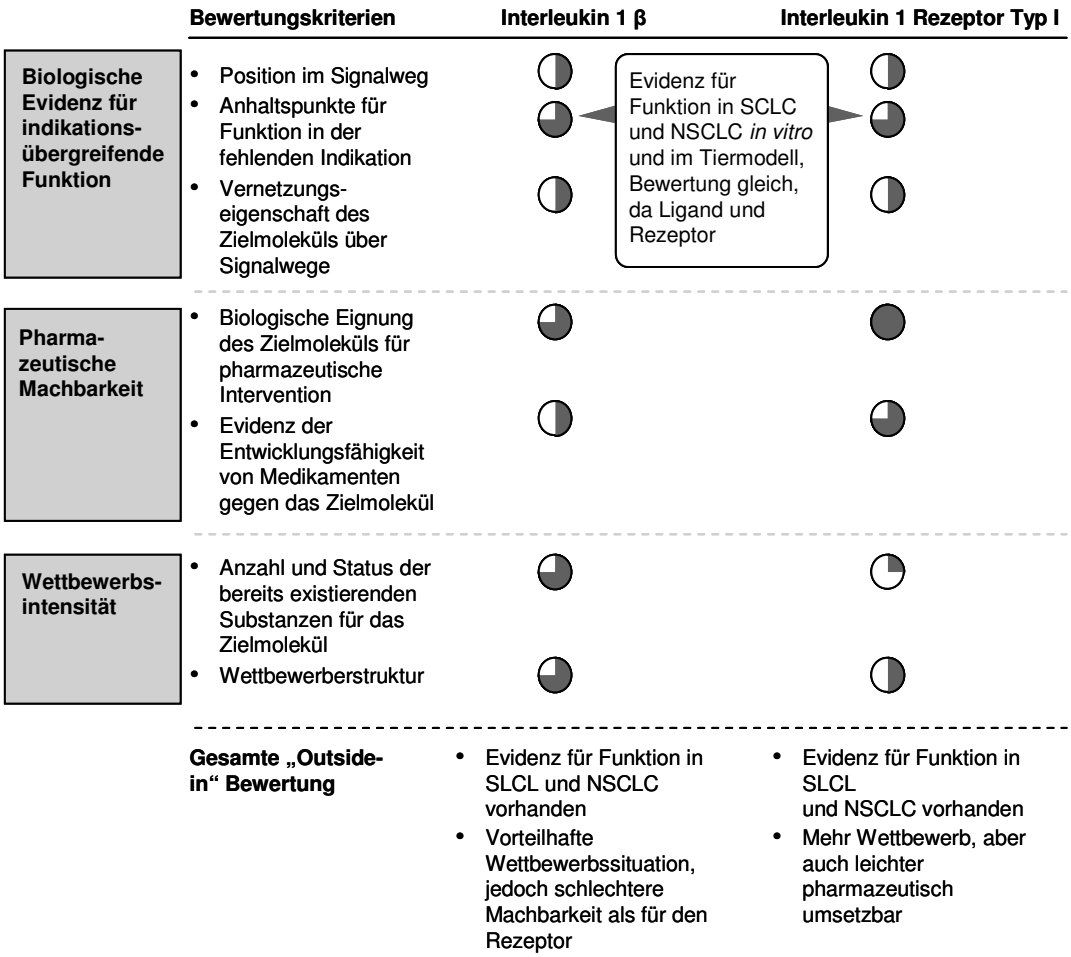


Abb. 7.37: Vergleichende Bewertung der Zielmoleküle Interleukin 1 β und Interleukin 1 Rezeptor Typ I

Die Abbildung zeigt einen Vergleich der Bewertung der Zielmoleküle Interleukin 1 β und Interleukin 1 Rezeptor Typ I für die indikationsübergreifende Funktion in den Indikationsgruppen Lungenkrebs und rheumatoider Arthritis. Beide Moleküle sind bereits für die Behandlung von rheumatoider Arthritis in Forschung und Entwicklung, daher wurde hier ihre Chance auf eine Funktion im Bereich Lungenkrebs bewertet.
Für beide Zielmoleküle gibt es Hinweise für eine Funktion bei Lungenkrebs. Medikamente, die sich gegen Interleukin 1 β richten, sind schwerer zu entwickeln als Medikamente, die den Interleukin 1 Rezeptor Typ I als Zielmolekül haben. Dafür ist die Wettbewerbssituation bei Interleukin 1 β vorteilhafter als beim Interleukin 1 Rezeptor Typ I.

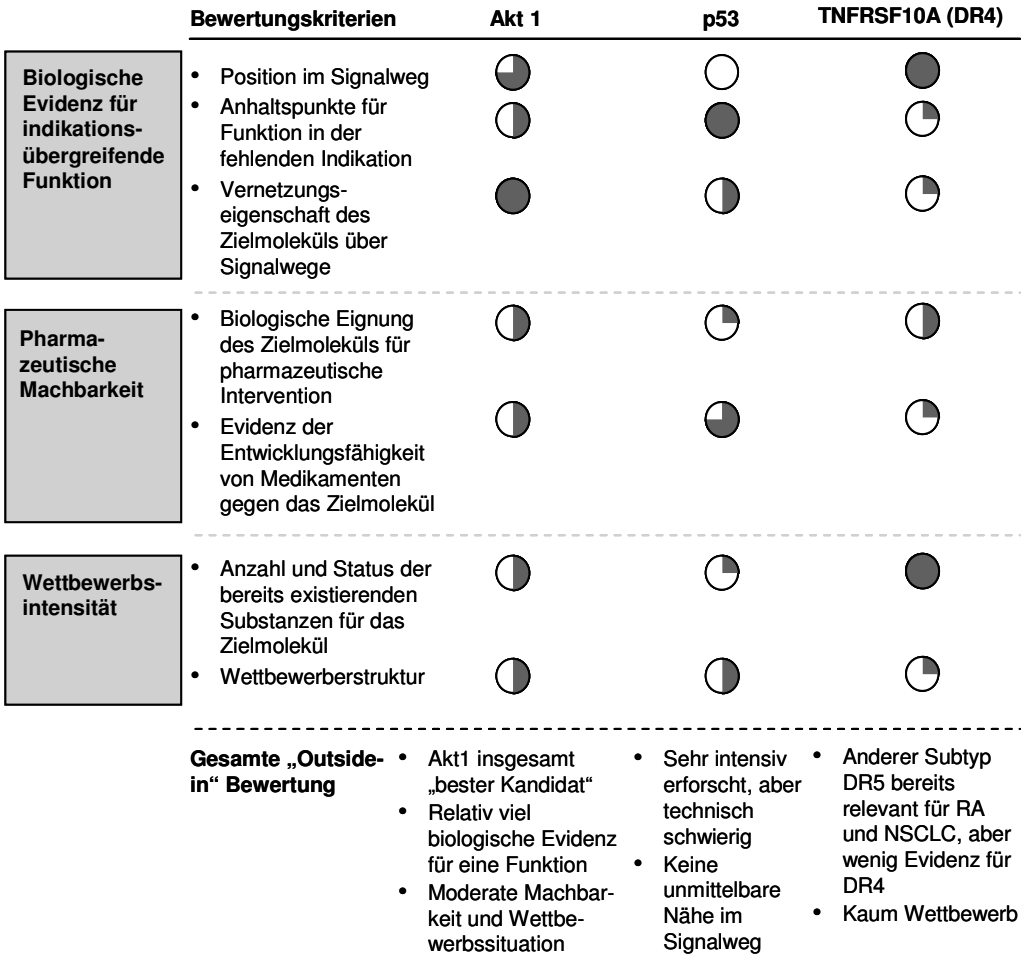


Abb. 7.38: Vergleichende Bewertung der Zielmoleküle Akt1, p53 und Death receptor 4 I
Die Abbildung zeigt einen Vergleich der Bewertung der Zielmoleküle Akt1, p53 und Death receptor 4 für die indikationsübergreifende Funktion in den Indikationsgruppen Lungenkrebs und rheumatoider Arthritis. Alle drei Moleküle sind bereits für die Behandlung von Lungenkrebs in Forschung und Entwicklung, daher wurde hier ihre Chance auf eine Funktion im Bereich rheumatoide Arthritis bewertet.
Akt1 erscheint im Vergleich als bester Kandidat für Forschung und Entwicklung. Zwar gibt es bei p53 mehr Hinweise für eine Funktion und eine breitere Literatur, dafür hat Akt1 im Bereich Machbarkeit einen großen Vorteil gegenüber p53. Der Death receptor 4 sollte depriorisiert werden, da es kaum biologische Evidenz für eine Funktion bei rheumatoider Arthritis gibt.

Abb. 7.38 zeigt den direkten Vergleich der Bewertung der Moleküle Akt1, p53 und Death receptor 4, für die die Chancen einer therapeutischen Bedeutung im Bereich rheumatoide Arthritis untersucht wurden. Zusammenfassend lässt sich sagen, dass bei „outside-in"-Bewertung Akt1 einen leichten Vorteil als Kandidatenmolekül gegenüber p53 und Death receptor 4 hat. Sowohl Akt1 als auch p53 haben deutlich mehr Belege für eine Funktion bei rheumatoider Arthritis als der Death receptor 4 und daher basierend auf dem momentanen Wissenstand eine größere biologische Erfolgschance. Der Death receptor 4 erscheint nicht als

geeigneter Kandidat für rheumatoide Arthritis und sollte aufgrund der geringen biologischen Evidenz einer therapeutischen Funktion depriorisiert werden. Im Bereich der pharmazeutischen Machbarkeit kann Akt1 einen Vorsprung gegenüber p53 gewinnen und erscheint daher als bester Kandidat. Für eine finale Priorisierung aus Sicht eines Unternehmens bzw. einer Forschungsgruppe müssen aber auch in diesem Fall strategische Abwägungen mit in die Entscheidung eingebunden werden.

8 Diskussion

8.1 Diskussion der Analyseschritte am Beispiel ausgewählter Ergebnisse

Forschung und Entwicklung haben in den vergangenen zwanzig Jahren eine Fülle von Informationen verfügbar gemacht, die es der pharmazeutischen Industrie erlauben, die Medikamentenentwicklung auf einem tieferen Verständnis der biologischen und medizinischen Zusammenhänge aufzubauen (Butcher u. a., 2004; Apic u. a., 2005). Die Integration der große Datenmengen stellt die Industrie aber auch vor neue Herausforderungen (Nicholson und Wilson, 2003; Nikolsky u. a., 2005). Informationen über Genom und Proteom (sogenannte „Omics"-Daten) sind ebenso öffentlich verfügbar, wie detaillierte Informationen über biologische Signalwege (Ashburner u. a., 2000; Kanehisa und Goto, 2000; Kanehisa u. a., 2006; Kanehisa u. a., 2008). Darüber hinaus gibt es eine große Anzahl pharmazeutischer und biotechnologischer Unternehmen, die unter hohem finanziellem Aufwand Forschung und Entwicklung von Medikamenten betreiben. Informationen über diese Forschungsprojekte sind in kommerziell erhältlichen Datenbanken, wie zum Beispiel Pharmaprojects, zugänglich (Informa UK Ltd, 2008).

Die in dieser Arbeit entwickelte Methodik bietet eine Möglichkeit, das verfügbare biologische Wissen über funktionelle Zusammenhänge und die Erfahrungen der Industrie mit der Entwicklung von Medikamenten zu kombinieren und für neue Forschungs- und Entwicklungsprojekte nutzbar zu machen. Die Zielsetzung bei der Entwicklung der Methodik war dabei, eine Möglichkeit zu finden, Zusammenhänge zwischen unterschiedlichen therapeutischen Gebieten und Indikationsgruppen durch gemeinsame, sogenannte indikationsübergreifende, Zielmoleküle und Signalwege automatisiert zu identifizieren. Zu diesem Zweck wurde eine Access-basierte relationale Datenbank aufgebaut und ein Bewertungssystem zur detaillierten Priorisierung der Ergebnisse angeschlossen.

Abb. 8.1 zeigt eine schematische Darstellung der Methodik mit den einzelnen Analyseschritten von der Identifikation von gemeinsamen Signalwegen zwischen therapeutischen Gebieten bis zur Priorisierung von Kandidatenmolekülen für indikationsübergreifende Forschung und Entwicklung.

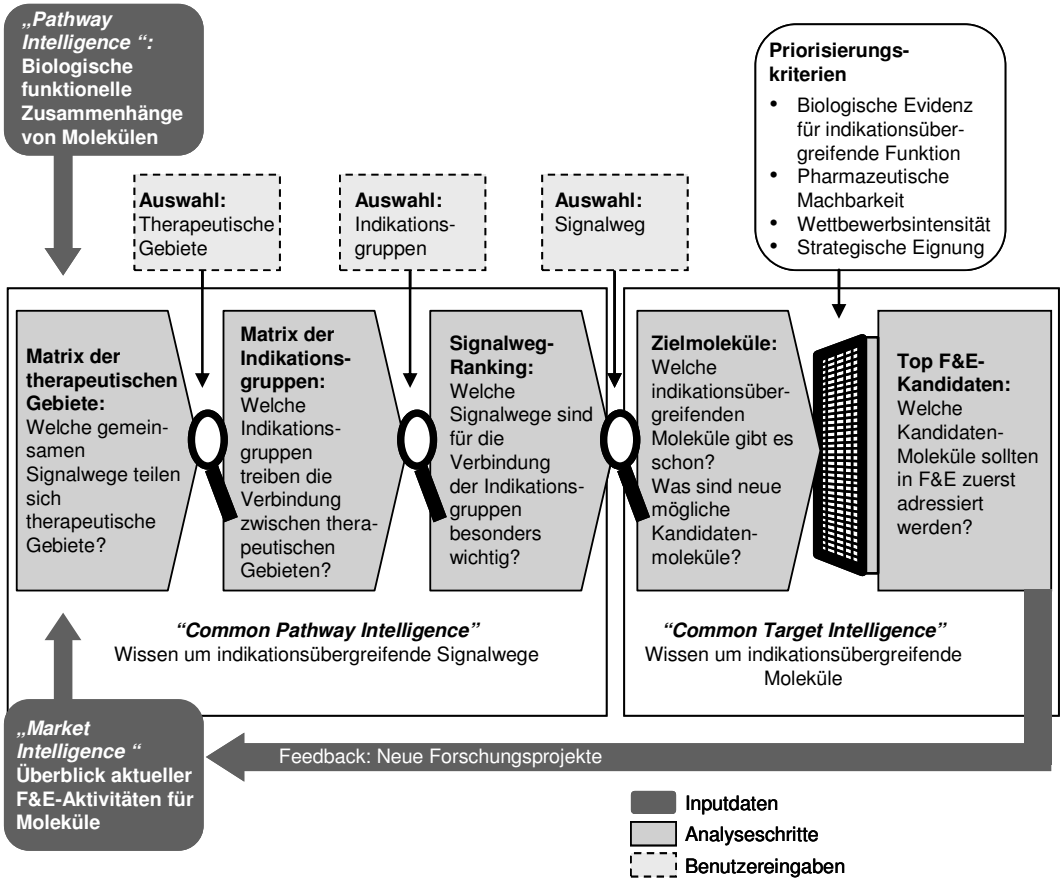


Abb. 8.1: Methodik zur Identifikation von indikationsübergreifenden Zusammenhängen
Die Abbildung zeigt schematisch die in dieser Arbeit entwickelte Methodik zur Identifikation von indikationsübergreifenden Zusammenhängen.
Inputdaten für die Datenbank sind ein Überblick über aktuelle Forschungs- und Entwicklungsaktivitäten („Market Intelligence") und biologische Informationen über Signalwege („Pathway Intelligence"). Im ersten Analyseschritt werden die Verbindungen von therapeutischen Gebieten durch gemeinsame Signalwege untersucht. Nach Auswahl von bestimmten Gebieten (Benutzereingabe) wird die Verbindung dieser Gebiete auf Ebene von Indikationsgruppen untersucht. Im nächsten Schritt kann nach Auswahl einer bestimmten Indikationsgruppenverbindung gezeigt werden, welche Signalwege für diese Verbindung besonders von Bedeutung sind. Im vierten Schritt werden dann Zielmoleküle für einen ausgewählten Signalweg in der Verbindung der Indikationsgruppen untersucht. Zuletzt folgt dann die Priorisierung der neuen Kandidatenmoleküle.
Neue, daraus entstehende, Forschungsprojekte müssen dann in die Inputdatenbasis eingefügt werden („Feedback").

Bevor im Abschnitt 8.2 übergreifend die Anwendung der Methodik für pharmazeutische Forschung und Entwicklung diskutiert wird, werden im Folgenden die einzelnen Analyseschritte am Beispiel von ausgewählten

Ergebnissen diskutiert. In diesem Zusammenhang sollen insbesondere die folgenden Fragen für jeden Analyseschritt beantwortet werden:

- Konnte das in der Zielsetzung der Arbeit definierte Ziel erreicht werden?
- Wie können die Ergebnisse des Analyseschritts interpretiert werden?
- Welche Einschränkungen müssen für den Analyseschritt beachtet werden?

8.1.1 Kombinationen von therapeutischen Gebieten mit erhöhter Erfolgswahrscheinlichkeit für indikationsübergreifende Medikamentenentwicklung

Erreichen der Zielsetzung

Zielsetzung des ersten Analyseschritts war es, Kombinationen von therapeutischen Gebieten zu identifizieren, für die es besonders gute Chancen gibt, erfolgreich indikationsübergreifende Medikamente zu entwickeln. Zu diesem Zweck wurde untersucht, wie groß die Schnittmenge der biologischen Zusammenhänge zwischen zwei therapeutischen Gebieten ist. Grundannahme war dabei, dass ein Medikament seine indikationsübergreifende Funktion dadurch erhält, dass es in einen biologischen Zusammenhang eingreift, der für mehrere therapeutische Gebiete von Bedeutung ist. Teilen sich zwei therapeutische Gebiete besonders viele biologische Zusammenhänge ist daher die Wahrscheinlichkeit groß, erfolgreich ein Medikament entwickeln zu können, das in beiden Gebieten therapeutisch wirksam ist.

Um therapeutische Gebiete mit einer besonders großen Schnittmenge von biologischen Zusammenhängen zu identifizieren, wurde zunächst die Anzahl gemeinsamer Signalwege von zwei therapeutischen Gebieten bestimmt. Diese wurde dann für alle Kombinationen von Gebieten verglichen (vergleiche Vernetzungsmatrix, Abb. 7.6). Mit Hilfe des Hilfsparameters Vernetzungsgrad (vergleiche Abschnitt 7.2.1) konnte eine zusammenfassende Aussage über die Vernetzung eines therapeutischen Gebiets mit anderen therapeutischen Gebieten gemacht werden. Durch Einführung einer Gewichtungslogik war es möglich, die

einzelnen verbindenden Signalwege entsprechend ihrer jeweiligen Verbindungsstärke differenziert zu berücksichtigen.

Zusammenfassend kann gesagt werden, dass es mit Hilfe der Methodik möglich ist, Kombinationen von therapeutischen Gebieten zu identifizieren, die eine erhöhte Erfolgschance für die Entwicklung indikationsübergreifender Medikamente haben.

Interpretation der Ergebnisse

Die Ergebnisse des ersten Analyseschritts werden durch die starke Vernetzung der Onkologie dominiert. Die therapeutischen Gebiete Onkologie und Bewegungsapparat teilen sich die höchste Anzahl gemeinsamer Signalwege und stellen damit die engste Verbindung zwischen zwei therapeutischen Gebieten dar. Durch diese große Schnittmenge von biologischen Zusammenhängen ist die Kombination von Onkologie und Bewegungsapparat besonders für die Entwicklung von indikationsübergreifenden Medikamenten geeignet. Auch für andere therapeutische Gebiete ist die Onkologie der Verbindungspartner mit der höchsten Anzahl an gemeinsamen Signalwegen. Dieses Ergebnis überrascht nicht, da die Onkologie ein sehr großes therapeutisches Gebiet ist (vergleiche Abb. 7.3), in dem vor allem grundlegende zelluläre Mechanismen, die viele Organsysteme betreffen, durch Medikamente adressiert werden.

Genauso erwartungsgemäß ist die geringe Vernetzung des therapeutischen Gebiets Infektion und des therapeutischen Gebiets Impfstoffe und Prophylaxe mit anderen Gebieten. Zielmoleküle aus diesen therapeutischen Gebieten sind häufig Teil des pathogenen Organismus und damit nicht Teil eines humanen Signalwegs. Diese Übereinstimmung der Ergebnisse mit den Erwartungen bestätigt die korrekte Funktion des Analyseschritts.

Darüber hinaus konnten auch überraschende Beobachtungen gemacht werden. So weist zum Beispiel das therapeutische Gebiet der Transplantationsmedizin und des Immunsystems einen hohen Vernetzungsgrad auf. Dieses, vergleichsweise kleine therapeutische Gebiet (vergleiche Abb. 7.3), teilt sich offenbar viele Signalwege mit anderen Gebieten.

Mit Hilfe der Gewichtung konnten stärkere von schwächeren Signalwegverbindungen getrennt werden (vergleiche Abschnitt 7.2.3). Der Vergleich unterschiedlicher Gewichtungen erlaubte außerdem zu beurteilen, wie stark eine Signalwegverbindung durch relevante Zielmoleküle und pharmazeutisch aktive Substanzen gestützt wird. Verbindungen mit dem therapeutischen Gebiet Transplantationsmedizin und Immunsystem erscheinen nach Gewichtung beispielsweise deutlich abgeschwächt gegenüber dem Vernetzungsbild vor Gewichtung (vergleiche Abb. 7.6, Abb. 7.21 und Abb. 7.22). Das therapeutische Gebiet hat offenbar zahlreiche gemeinsame Signalwege mit anderen Gebieten. Innerhalb derer sind aber bisher nur wenige Zielmoleküle für die Verbindung von Bedeutung und diese werden noch kaum durch pharmazeutisch aktive Substanzen adressiert. Eine mögliche Erklärung für diese Beobachtung wäre, dass es sich bei dem therapeutischen Gebiet noch um ein verhältnismäßig junges Forschungsfeld handelt, das bisher noch nicht im Mittelpunkt des pharmazeutischen Interesses steht. Aufgrund seiner zahlreichen biologischen Verbindungen zu anderen Gebieten könnte es aber Potential für indikationsübergreifende Mechanismen bieten.

Ein Beispiel für eine mögliche Differenzierung von treibenden Faktoren einer Verbindung kann anhand der therapeutischen Gebiete Herz-Kreislauf-System und Bewegungsapparat diskutiert werden. Verbindungen mit dem Herz-Kreislauf-System tendieren zu hohen Verbindungswerten bei einfacher Durchschnittsbildung und erscheinen bei doppelter Durchschnittsbildung deutlich schwächer. Der umgekehrte Effekt gilt für Verbindungen mit dem Bewegungsapparat (vergleiche Abb. 7.6, Abb. 7.21 und Abb. 7.22). Offenbar werden Herz-Kreislauf-Indikationen bereits durch viele pharmazeutisch aktive Substanzen in weit fortgeschrittenen Entwicklungsstadien adressiert. Diese Substanzen richten sich jedoch stets gegen eine sehr begrenzte Anzahl von Zielmolekülen. Es kann vermutet werden, dass die hohe Substanzanzahl pro Zielmolekül im Herz-Kreislauf-System auch durch die kommerzielle Attraktivität des Gebiets beeinflusst wird. Im Gegensatz dazu haben die gemeinsamen Signalwege mit dem Bewegungsapparat viele für die Verbindung relevante Zielmoleküle, die aber bisher nur durch wenige Substanzen adressiert werden.

Diese Verbindungen könnten also ein besonders interessanter Forschungsschwerpunkt sein, da sie noch nicht so stark kommerziell ausgeschöpft sind.

Einschränkungen

Ein Teil der Ergebnisse des ersten Analyseschritts ist naheliegend und entspricht den Erwartungen. Ursache hierfür sind vor allem die natürlichen Eigenschaften der untersuchten therapeutischen Gebiete. Die Größe eines therapeutischen Gebiets, seine Verteilung über verschiedene Organsysteme und der Wissenstand der Forschung auf dem Gebiet spielen eine Rolle für den Vernetzungsgrad. Diese Eigenschaften müssen bei der Interpretation der Ergebnisse berücksichtigt werden.

8.1.2 Kombinationen von Indikationsgruppen mit erhöhter Erfolgswahrscheinlichkeit für indikationsübergreifende Medikamentenentwicklung

Erreichen der Zielsetzung

Nach der Identifizierung der Kombinationen von therapeutischen Gebieten mit einer besonders großen Schnittmenge von biologischen Zusammenhängen, richtete sich die zweite Fragestellung auf die Vernetzung von Indikationsgruppen. Ziel war es, innerhalb einer gewählten Kombination von therapeutischen Gebieten, diejenigen Kombinationen von Indikationsgruppen zu identifizieren, die besonders eng durch biologische Zusammenhänge miteinander verbunden sind. Hintergrund dieser Fokussierung der Analyse auf bestimmte Kombinationen von therapeutischen Gebieten war die Tatsache, dass viele pharmazeutische und biotechnologische Unternehmen bereits über einen etablierten Forschungs- und Entwicklungsschwerpunkt in bestimmten therapeutischen Gebieten verfügen. Es sollte daher durch die Methodik möglich gemacht werden, innerhalb dieser bestehenden Gebiete, nach besonders geeigneten Kombinationen von Indikationsgruppen zu suchen (vergleiche Abb. 8.1).

Zur Lösung dieser Aufgabe wurde untersucht, welche Indikationsgruppen die größte Schnittmenge von biologischen Zusammenhängen haben.

Ausgangshypothese war dabei – analog zur Vernetzungsanalyse der therapeutischen Gebiete – dass mit der Größe des gemeinsamen biologischen Kontexts auch die Möglichkeiten für indikationsübergreifende therapeutische Interventionen steigen.

Um eine differenzierte Auswertung der Vernetzung zu ermöglichen, wurden sowohl die absolute Anzahl der gemeinsamen Signalwege, als auch die gewichteten Verbindungsstärken zwischen den einzelnen Indikationsgruppen ermittelt. Methodisch gesehen stellt der zweite Analyseschritt damit einen „Zoom“ auf eine Kombination von therapeutischen Gebieten dar (vergleiche Abb. 8.1).

Interpretation der Ergebnisse

Es wurden insgesamt fünf Verbindungen von therapeutischen Gebieten im Detail analysiert (vergleiche Abschnitt 7.3). Im Folgenden werden generelle Aspekte der Ergebnisse an ausgewählten Beispielen exemplarisch diskutiert.

Insgesamt zeigen die Ergebnisse des zweiten Analyseschritts, dass die Verbindung von zwei therapeutischen Gebieten in der Regel nicht gleichermaßen durch Verbindung aller Indikationsgruppen verursacht wird. Sowohl die Anzahl der gemeinsamen Signalwege als auch die gewichtete relative Verbindungsstärke variiert deutlich zwischen den Verbindungen der Indikationsgruppen von zwei ausgewählten therapeutischen Gebieten. Es ist daher möglich, Kombinationen von Indikationsgruppen zu benennen, die enger miteinander verbunden sind als andere und damit bessere Chancen für die erfolgreiche Entwicklung indikationsübergreifender Medikamente bieten. Der Vergleich der gewichteten Verbindungsstärken bietet außerdem Rückschlüsse darauf, in wie weit die Zielmoleküle des verbindenden Signalwegs bereits von kommerziellem Interesse sind.

Ähnlich wie bei der Vernetzungsanalyse der therapeutischen Gebiete sind auch die Ergebnisse des zweiten Analyseschritts eine Mischung aus erwartungsgemäßen und überraschenden Beobachtungen. Ein Beispiel für eine erwartungsgemäße Beobachtung kann anhand der Vernetzungsmatrix der therapeutischen Gebiete Onkologie mit Bewegungsapparat (vergleiche Abb. 7.27)

diskutiert werden. Die Indikationsgruppe der Erbkrankheiten von Knochen, peripherem Nervensystem, Muskel und Bindegewebe teilt sich eine hohe Anzahl von Signalwegen mit den Indikationsgruppen der Onkologie. Da in dieser Indikationsgruppe viele heterogene Krankheitsbilder zusammengefasst sind, ist es nicht verwunderlich, dass sich viele Möglichkeiten zur Verbindung mit anderen Indikationsgruppen finden. Durch Gewichtung können zusätzlich Informationen über die treibenden Faktoren der Verbindung gewonnen werden: Bei doppelter Durchschnittsbildung erscheinen die Verbindungen zu den Erbkrankheiten des Bewegungsapparats deutlich stärker als bei einfacher Durchschnittsbildung. Offenbar stützen zahlreiche Zielmoleküle die verbindenden Signalwege, jedoch werden diese nur wenig durch Substanzen in Forschung und Entwicklung adressiert. Eine mögliche Ursache hierfür könnte in der geringen wirtschaftlichen Attraktivität der Indikationsgruppe der Erbkrankheiten des Bewegungsapparats liegen.

Ein anderes erwartungsgemäßes Ergebnis ist die isolierte Stellung der Indikationsgruppe Anästhesie. Die spezifische Ausrichtung der Anästhesie auf selektive Modulation einzelner Vorgänge des Zentralnervensystems erklärt die verschwindend geringe Anzahl von gemeinsamen Signalwegen mit anderen Indikationsgruppen (vergleiche Abb. 7.27 und Abb. 7.30).

Andere Ergebnisse sind nicht unmittelbar erklärbar oder naheliegend und daher besonders für weitere Analysen interessant. Abb. 7.27 zeigt zum Beispiel die enge Verbindung der rheumatoiden Arthritis mit den Indikationsgruppen Lungenkrebs und Brustkrebs. Mit und ohne Gewichtung zeigen diese Indikationsgruppen eine enge Verbindung und stellen eine vielversprechende Kombination für die Entwicklung indikationsübergreifender Medikamente dar.

Auch die Analyse der Vernetzung der Indikationsgruppen der therapeutischen Gebiete Zentralnervensystem und Herz-Kreislauf-System lieferte sowohl naheliegende als auch unerwartete Ergebnisse (vergleiche Abb. 7.28). Zwei Verbindungen mit hoher Anzahl an gemeinsamen Signalwegen dominieren die ungewichtete Vernetzungsmatrix: die Verbindung von multiple Sklerose mit Thrombose und die Verbindung von Alzheimer und Demenz mit Thrombose. Die

gewichtete Vernetzungsmatrix konnte bestätigen, dass beide Verbindungspunkte nicht nur viele Signalwege teilen, sondern auch über eine hohe Anzahl relevanter Zielmoleküle verfügen. Während eine Verbindung zwischen multipler Sklerose und Thrombose aufgrund der involvierten Serotonin-Singalzusammenhänge (Linder u. a., 2007; Sheremata u. a., 2008) erklärbar scheint, überrascht die enge Verbindung der neurodegenerativen Erkrankungen mit Thrombose.

Zusammenfassend kann gesagt werden, dass der zweite Analyseschritt eine Identifikation derjenigen Indikationsgruppen erlaubt, die besonders eng miteinander verbunden sind. Die Analyse kann dabei für jede Kombination von therapeutischen Gebieten durchgeführt werden.

Einschränkungen

Ähnlich wie bei der Vernetzung der therapeutischen Gebiete spielen auch bei der Vernetzung der Indikationsgruppen spezifische Eigenschaften der Indikationsgruppen eine Rolle. Das bedeutet, dass zum Beispiel heterogene Indikationsgruppen mit vielen verschiedenen Indikationen eher zur Vernetzung neigen als Gruppen mit wenigen, nahe verwandten Indikationen. Diese spezifischen Kriterien müssen bei der Interpretation der Ergebnisse berücksichtigt werden.

8.1.3 Identifizierung und Priorisierung von gemeinsamen Signalwegen

Erreichen der Zielsetzung

Ziel des dritten Analyseschritts war es, diejenigen Signalwege zu identifizieren, die für eine Verbindung von Indikationsgruppen besonders wichtig sind. Dabei wurden diejenigen Signalwege priorisiert, die für eine bestimmte Verbindung die größte Verbindungsstärke haben (vergleiche Abschnitt 7.2.3 und Abschnitt 7.4). Im Ablauf der Analyseschritte ist die Signalwegpriorisierung der dritter Schritt nach der Fokussierung auf zwei Indikationsgruppen (vergleiche Abb. 8.1). Darüber hinaus ist es aber auch möglich, die Signalwege einer Verbindung von therapeutischen Gebieten aufzulisten und zu priorisieren. Aufgrund der Tatsache, dass in einem therapeutischen Gebiet jedoch mehrere unterschiedliche

Indikationsgruppen zusammengefasst sind, sollte die Signalwegpriorisierung möglichst stets auf der Detailebene der Indikationsgruppen erfolgen.

Interpretation der Ergebnisse

Die Rangliste von Signalwegen eines Verbindungspunkts von Indikationsgruppen schlüsselt die Verbindungsdichte, die im zweiten Analyseschritt gezeigt wurde, im Detail auf. Sie kann also als Hintergrundinformation zur Vernetzungsanalyse der Indikationsgruppen gesehen werden. Abb. 10.60 zeigt beispielsweise die Signalwegrangliste für die zuvor diskutierte Verbindung von Alzheimer und Demenz mit Thrombose und hilft dabei zu verstehen, welche Signalwege für die enge Verbindung verantwortlich sind.

Die Verbindung der Indikationsgruppen Alzheimer und Demenz mit Thrombose wird beispielsweise durch die Signalwege „Neuroactive ligand-receptor interaction“ und „Calcium signaling pathway“ dominiert (vergleiche Abb. 10.60). Sie machen zusammen beinahe ein Drittel der Gesamtverbindungsstärke aus. Beide Signalwege verfügen über viele indirekt verbindende Zielmoleküle, die bisher nur für eine der beiden Indikationsgruppen relevant sind. Darüber hinaus haben sie auch direkte Verbindungen durch indikationsübergreifende Zielmoleküle.

In der Rangliste werden die Signalwege entsprechend ihres jeweiligen Anteils an der Gesamtverbindungsstärke aufgelistet. Der Vergleich dieses Anteils bei Gewichtung mit einfacher und doppelter Durchschnittsbildung ermöglicht zu beurteilen, ob innerhalb des Signalwegs bestimmte Zielmoleküle vorhanden sind, für die es bereits besonders viele pharmazeutisch aktive Substanzen gibt.

Auch in der Signalwegrangliste der Verbindung von Alzheimer und Demenz mit Thrombose können Verbindungen, die durch besonders starkes kommerzielles Interesse und eine hohe Substanzanzahl pro Zielmolekül gekennzeichnet sind, identifiziert werden: Während der Signalweg „Neuroactive ligand-receptor interaction“ bei einfacher und doppelter Durchschnittsbildung gleich stark zur Verbindung beiträgt, gibt es bei anderen Signalwegen (z.B. „Complement and coagulation cascades“) deutliche Unterschiede in Abhängigkeit der Gewichtung (vergleiche Abb. 10.60). Offenbar gibt es zum Beispiel im „Complement and

coagulation cascades“ Signalweg einzelne Zielmoleküle, die für eine oder beide Indikationsgruppen stark kommerziell genutzt werden. Der Parameter des Anteils des Signalwegs, der an der Verbindung beteiligt ist, gibt einen Hinweis darauf, ob der Signalweg eine spezifische Funktion für die Verbindung hat. Sehr geringe Werte deuten darauf hin, dass lediglich ein kleiner Teil des Signalwegs für die Verbindung eine Rolle spielt. Dieses Phänomen kann vor allem bei sehr großen Signalwegen von Bedeutung sein.

Einschränkungen

Signalwege sind natürliche Einheiten unterschiedlicher Größe mit verschiedenen spezifischen Eigenschaften. Große Signalwege, die universelle Zellfunktionen beschreiben, haben *per se* eine höhere Chance auf den vorderen Rangplätzen zu erscheinen. Regressionsanalysen konnten zeigen, dass die Tendenz zur Vernetzung jedoch keinesfalls nur durch die Größe des Signalwegs bestimmt ist (vergleiche Abschnitt 7.2.2). Bei der Interpretation einer Signalwegrangliste sollten trotzdem auch spezifische Eigenschaften eines Signalwegs berücksichtigt werden. Es ist nicht möglich, den wichtigsten Signalweg einer Verbindung standardisiert aus nur einem Messparameter abzuleiten. Vielmehr müssen stets alle Parameter und die spezifischen Eigenschaften eines Signalwegs berücksichtigt werden, um zu einer finalen Priorisierung zu gelangen.

8.1.4 Identifizierung von möglichen Forschungskandidaten für Zielmoleküle mit indikationsübergreifender Funktion

Erreichen der Zielsetzung

Im vierten Analyseschritt sollten für eine bestimmte Verbindung von Indikationsgruppen durch einen Signalweg diejenigen Zielmoleküle ermittelt werden, die aufgrund ihrer Signalwegzugehörigkeit eine erhöhte Chance haben, in Zukunft indikationsübergreifend wichtig zu sein (vergleiche Abb. 8.1). Hintergrund dieser Zielsetzung war es, mit Hilfe der Methodik zu Ergebnissen zu kommen, die möglichst unmittelbar für pharmazeutische Forschung und Entwicklung genutzt werden können. Zu diesem Zweck wurden die Zielmoleküle eines verbindenden Signalwegs entsprechend ihrer bisherigen pharmazeutischen

Nutzung in drei Klassen eingeteilt (vergleiche Abschnitt 7.5). Für die vorliegende Arbeit waren vor allem die Zielmoleküle eines Signalwegs interessant, die bisher erst für eine der beiden Indikationsgruppen erforscht bzw. genutzt werden, aber aufgrund ihrer Rolle im Signalweg eine erhöhte Chance haben, in Zukunft für beide Indikationsgruppen übergreifend von Bedeutung zu sein. Zusätzlich dazu wurden alle bereits bekannten indikationsübergreifenden Zielmoleküle des Signalwegs, sowie Zielmoleküle, die für beide Indikationsgruppen neu sind, identifiziert.

Die Identifizierung der Zielmoleküle der einzelnen Klassen wurde in der Datenbank so umgesetzt, dass durch Makros die aufeinanderfolgenden Abfragen automatisch gesteuert und Ergebnistabellen erstellt werden können.

Interpretation der Ergebnisse

Die Zielmoleküle einer untersuchten Verbindung werden durch die Datenbank in Form einer Namensliste für jede der drei Molekülklassen ermittelt. Neben der Liste der Kandidaten für neue indikationsübergreifende Zielmoleküle, die im fünften Analyseschritt priorisiert werden, ist vor allem die Liste der bereits bestehenden indikationsübergreifenden Zielmoleküle von Interesse. Da durch die entwickelte Methodik automatisiert Forschungs- und Entwicklungsprojekte aller Firmen analysiert werden, ist es möglich, Verbindungen darzustellen, die aufgrund der unterschiedlichen Natur der Entwickler, gegebenenfalls zuvor nicht offensichtlich waren. Darüber hinaus kann diese Information helfen, das Vernetzungsbild aus den ersten Analyseschritten zu interpretieren.

So wird beispielsweise die enge Verbindung der Indikationsgruppen Alzheimer und Demenz mit Thrombose unter anderem durch fünf direkt indikationsübergreifende Zielmoleküle aus dem „Neuroactive ligand-receptor interaction“-Signalweg gestützt (vergleiche Abb. 7.28 und Abb. 10.60). Bei diesen Molekülen handelt es sich um Neurotransmitterrezeptoren für Glutamat, Serotonin und Cannabinoide.

Einschränkungen

Bei der Identifizierung von Kandidaten für zukünftig indikationsübergreifende Zielmoleküle wurde sich im Rahmen der Arbeit bewusst auf diejenigen Moleküle beschränkt, die bereits in einer der beiden Indikationsgruppen pharmazeutisch genutzt werden („Klasse 2“, vergleiche Abschnitt 7.6.1). Grundsätzlich sind auch die Moleküle des verbindenden Signalwegs für eine indikationsübergreifende Funktion von Interesse, die bisher noch für keine der beiden Indikationsgruppen relevant sind („Klasse 3“, vergleiche Abschnitt 7.6.1). Bei diesen Molekülen ist jedoch aufgrund der Tatsache, dass beide Indikationsgruppen neu erschlossen werden müssen, mit höheren Herausforderungen in der Medikamentenentwicklung zu rechnen. Im Gegenzug bieten diese Moleküle einen größeren Neuheitswert (vergleiche Abb. 7.35).

8.1.5 Priorisierung von Kandidaten für zukünftig indikationsübergreifende Zielmoleküle

Erreichen der Zielsetzung

Je nach Größe des untersuchten Signalwegs und den ausgewählten Indikationsgruppen schwankt die Anzahl der Kandidaten für zukünftig indikationsübergreifende Zielmoleküle. Um auch aus einer hohen Anzahl von gegebenenfalls mehreren dutzend Zielmolekülen diejenigen zu finden, die in der Praxis die höchsten Erfolgschancen haben, war es wichtig, mit Hilfe eines Bewertungssystem außerhalb der Datenbank eine Priorisierung vorzunehmen.

Zu diesem Zweck wurde eine Priorisierungslogik entlang von vier Bereichen definiert (vergleiche Abschnitt 7.6.2). Für die Bewertung eines einzelnen Zielmoleküls innerhalb eines Bereichs wurden genaue Kriterien abgrenzt, die es ermöglichten, jedes Molekül zunächst alleinstehend zu bewerten. Nach Abschluss der Einzelbewertungen konnte eine vergleichende Bewertung aller Kandidaten angeschlossen werden.

Bei der Priorisierung wurden zusätzlich zu den in der Datenbank enthaltenen Informationen weitere Parameter, wie zum Beispiel der Überblick über die aktuelle Forschungssituation eines Zielmoleküls, berücksichtigt. Zu diesem

Zweck wurde neben der Datenbank eine systematische PubMed- und Mesh-Browser Recherche (U.S. National Library of Medicine, 2008b; U.S. National Library of Medicine, 2008a) mit anschließender Überprüfung der einzelnen Veröffentlichungen und die Proteindatenbank UniProt (The UniProt, 2008) verwendet.

Interpretation der Ergebnisse

Die Priorisierung der Zielmolekülkandidaten wurde exemplarisch an fünf Zielmolekülen aus dem Apoptose-Signalweg für die Verbindung von Lungenkrebs mit rheumatoider Arthritis durchgeführt. Die exemplarische Bewertung zeigte, dass es möglich ist durch Kombination unterschiedlicher Bewertungsbereiche und Anwendung von standardisierten Kriterien, geeignete Zielmolkekülkandidaten von weniger geeigneten zu trennen.

Einschränkungen

Bei Definition der Priorisierungslogik wurden bewusst standardisierte Kriterien und externe Datenquellen für eine möglichst neutrale Bewertung der Zielmolekülkandidaten verwendet. Nichts desto trotz sind die Kriterien überwiegend qualitativ und subjektive Faktoren (zum Beispiel subjektive Bewertung der verfügbaren Literatur zur biologischen Evidenz für die indikationsübergreifende Funktion) gehen in die Priorisierung ein. Vor der finalen Entscheidung für oder gegen bestimmte Forschungskandidaten sollte die Bewertung daher kritisch diskutiert werden.

Die in dieser Arbeit vorgeschlagene Vorgehensweise ist nur eine Möglichkeit zur Priorisierung. Eine Priorisierung kann durch weitere Kriterien ergänzt bzw. abgewandelt werden.

8.2 Anwendung der Methodik zur Identifizierung von indikationsübergreifenden Zusammenhängen für pharmazeutische Forschung und Entwicklung

8.2.1 Besondere Aspekte der Methodik

Mit der in der vorliegenden Arbeit entwickelten Methodik können nicht nur Verbindungen zwischen therapeutischen Gebieten und Indikationsgruppen identifiziert werden, die ein besonders hohes Potential für indikationsübergreifende Medikamentenentwicklung bieten, sondern auch Signalwege und konkrete Zielmolekülkandidaten bestimmt werden, die in ausgewählten medizinischen Bereichen eine indikationsübergreifende Rolle spielen. Mit diesen Informationen ist es möglich, die Forschung und Entwicklung gezielt so auszurichten, dass die Wahrscheinlichkeit ein Medikament zu entwickeln, das in mehr als einer Indikation eine Rolle spielt, möglichst hoch ist.

Im Folgenden sollen sechs besondere Aspekte der Methodik diskutiert werden:

8.2.1.1 Berücksichtigung indirekter Verbindungen durch gemeinsame Signalwege

Indikationsübergreifende Zielmoleküle sind Moleküle, die von Medikamenten zur Behandlung von mehr als einer Krankheit angegriffen werden. Sie stellen eine direkte Verbindung derjenigen therapeutischen Gebieten bzw. der Indikationsgruppen dar, für die die Medikamente zugelassen sind. Diese direkten Verbindungen sind in der Regel bekannt und werden meist bereits durch eine Vielzahl von Forschungsprojekten adressiert. Die Komplexität des menschlichen Organismus sowie die Tatsache, dass Zielmoleküle im Körper keineswegs isoliert sondern stets in funktionellen Zusammenhängen agieren (Araujo u. a., 2007; Hellerstein, 2008b; Watterson u. a., 2008), lässt vermuten, dass es eine Vielzahl weiterer indikationsübergreifender Zielmoleküle gibt. Wenn ein Zielmolekül innerhalb eines Signalwegs eine direkte Verbindung in zwei therapeutische Gebiete hat, liegt die Hypothese nahe, dass ein anderes Molekül aus demselben Signalweg ebenfalls eine indikationsübergreifende Rolle spielen könnte.

Die in dieser Arbeit definierte Datenbank berücksichtigt neben direkten Verbindungen von therapeutischen Gebieten und Indikationsgruppen ebenfalls sogenannte indirekte Verbindungen. Damit sind Verbindungen gemeint, bei denen sich zwei Gebiete denselben Signalweg, nicht aber ein und dasselbe Zielmolekül teilen. Diese indirekten Verbindungen sind durch das Fehlen von gemeinsamen Zielmolekülen nicht offensichtlich. Sie bieten aber einen Hinweis auf mögliche zukünftige gemeinsame Zielmoleküle und sollten daher bei der Untersuchung der Vernetzung von therapeutischen Gebieten berücksichtigt werden. Im Gegensatz zur Verbindung durch ein direkt indikationsübergreifendes Molekül ist die indirekte Verbindung über einen Signalweg durch eine geringe Validität charakterisiert. Das bedeutet, dass ein gemeinsamer Signalweg zwar einen guten Anhaltspunkt aber keine Sicherheit dafür bietet, dass auch andere Moleküle des Signalwegs in Zukunft direkt indikationsübergreifend genutzt werden können. Diese unterschiedlichen Validitäten der direkten und indirekten Verbindungen werden in der Datenbank durch Gewichtung berücksichtigt.

8.2.1.2 Lernen von Anderen durch Berücksichtigung bekannter Forschungs- und Entwicklungsprojekte

Die Datenbank berücksichtigt durch Integration der Rohdaten der Pharmaprojects Datenbank (Informa UK Ltd, 2008) das öffentlich verfügbare Wissen über aktuelle und vergangene Projekte von Forschung und Entwicklung. Dies ermöglicht, die Erfahrungen anderer Unternehmen mit bestimmten Zielmolekülen systematisch und automatisiert bei einer neuen Entscheidung für ein Zielmolekül oder einen Signalweg als Forschungskandidaten zu berücksichtigen. Im Gegensatz zum klassischen Vorgehen, bei dem zunächst ein Zielmolekül ausgewählt wird, das dann erst im Anschluss unter Einbeziehung der Erfahrung anderer Unternehmen validiert wird (Brown und Superti-Furga, 2003; Butcher, 2005), werden bei der hier entwickelten Datenbank die Forschungsprojekte der Industrie bereits bei Auswahl eines Zielmolekülkandidaten berücksichtigt. Der Prozess erfolgt dabei automatisch und integriert alle Forschungsprojekte, die in den Rohdaten vorhanden sind. Damit wird sichergestellt, dass das Maximum an öffentlich verfügbaren Daten einbezogen wird. In Anbetracht der Tatsache, dass mehr als 90 % aller Projekte, die in die klinischen Phase I eintreten, im Verlauf

der klinischen Entwicklung scheitern (Brown und Superti-Furga, 2003), bietet dieses Lernen von Anderen eine Möglichkeit, von Anfang an alle verfügbaren Beobachtungen zu berücksichtigen. Im historischen Kontext der Medikamentenentwicklung wird die Berücksichtigung von Beobachtungen als zentraler Erfolgsfaktor gesehen (Hellerstein, 2008b). Auch bei neuen systembiologischen Ansätzen gilt die Fähigkeit zur Integration von Beobachtungen (sogenannte „Observability“) als wichtig (Hellerstein, 2008b).

8.2.1.3 Trennung von biologischen und marktseitigen Zusammenhängen

Die in der Datenbank verwendeten Gewichtungsmechanismen dienen grundsätzlich dazu, die einzelnen verbindenden Signalwegen entsprechend ihrer Verbindungsstärke zu differenzieren. Neben der unterschiedlichen Gewichtung von direkten und indirekten Verbindungen werden – je nach Art der Gewichtung – auch diejenigen Verbindungen stärker betont, die über eine Vielzahl von Forschungsprojekten verfügen. Durch Vergleich unterschiedlicher Gewichtungsarten (mit und ohne Berücksichtigung der Anzahl der Forschungsprojekte) können marktseitige Einflüsse identifiziert werden. Insbesondere unterschiedliche wirtschaftliche Attraktivität einzelner Gebiete oder eine kurze Forschungshistorie haben einen Einfluss auf die Anzahl der Forschungsprojekte im Gebiet. Eine Verbindung, die beispielsweise bei Gewichtung mit einfacher und doppelter Durchschnittsbildung sehr hohe Verbindungswerte erhält, ist nicht nur durch eine hohe Zahl von Zielmolekülen sondern auch durch viele Forschungsprojekte gekennzeichnet. Eine Verbindung, die jedoch bei Gewichtung mit einfacher Durchschnittsbildung niedrige Verbindungswerte enthält und nur bei doppelter Durchschnittbildung stark erscheint, hat zwar biologisch viele Zielmoleküle, ist jedoch nicht so stark durch Forschungsprojekte gestützt. Während im ersten Fall mit starkem Wettbewerbsdruck zu rechnen ist, sollte im zweiten Fall geprüft werden, wodurch die niedrigere Anzahl von Forschungsprojekten verursacht wird. Sollte es sich um ein junges Forschungsgebiet handeln, in dem erst seit kurzer Zeit Forschungsaktivitäten vorhanden sind, könnte das Gebiet ein attraktiver Schwerpunkt für indikationsübergreifende Forschung sein. Sollte die beschränkte Forschungsaktivität jedoch durch geringe wirtschaftliche Attraktivität des Gebiets

(zum Beispiel schwierige Kostenerstattungssituation) entstanden sein, müssen Vor- und Nachteile sorgfältig abgewogen werden.

8.2.1.4 Anwendung auf unterschiedlichen Detailebenen

Die meisten forschenden Unternehmen haben eine strategische Ausrichtung auf bestimmte therapeutische Gebiete oder Indikationsgruppen. Da mit einer solchen Spezialisierung in der Regel organisatorische Konsequenzen, wie zum Beispiel der Aufbau des Außendiensts zusammenhängen, ist eine Änderung des strategischen Schwerpunkts eines Unternehmens in der Praxis kaum denkbar. Im Gegenzug dazu gibt es biotechnologische „Start-up"-Unternehmen, die zum Beispiel basierend auf einer technischen Plattform (zum Beispiel Antikörpertechnologie, RNAi) beginnen und damit noch frei in der Auswahl von Schwerpunktgebieten sind. Um die Datenbank für alle möglichen Szenarien dieser Art verwendbar zu machen, wurden die Analysen in fünf Detailschritten implementiert. Zum einen können so für bereits bestehende Schwerpunkte von Indikationsgruppen und Signalwegen geeignete Kandidaten für Forschung und Entwicklung an indikationsübergreifenden Zielmolekülen identifiziert werden. Zum anderen ist es möglich, die Neuausrichtung strategischer Schwerpunkte von therapeutischen Gebieten auf diejenigen zu lenken, die die engste Verbindung untereinander zeigen und damit das größte indikationsübergreifende Potential bieten.

8.2.1.5 Automatisierung zur Vereinfachung der Analyse großer Datenmengen

Die im Rahmen dieser Arbeit aufgebaute Datenbank verwendet in Microsoft Access kombinierte Abfragefolgen und Makros, um die meisten Analysen automatisiert ablaufen zu lassen. Eingaben eines Benutzers (zum Beispiel Namen von Indikationsgruppen, die untersucht werden sollen) werden durch „Pop-up"-Fenster erzwungen. Diese Automatisierung ermöglicht eine verhältnismäßig einfache Nutzung der Datenbank und erlaubt die Analyse der großen Datenmengen in kurzer Zeit.

8.2.2 Herausforderungen und Grenzen der Methodik

Im Folgenden sollen besondere Herausforderungen bei der Nutzung der Methodik für Forschung und Entwicklung ebenso wie die Grenzen der Anwendung diskutiert werden.

8.2.2.1 Pflege der Datenbank

Wichtigste Herausforderung bei Aufbau und Nutzung einer Datenbank, die vergleichbar mit der hier beschriebenen ist, ist die Pflege der Datenbank. Um aktuelle und damit für Forschung und Entwicklung anwendbare Ergebnisse zu generieren, ist es notwendig, in regelmäßigen Abständen die Rohdatenbasis zu aktualisieren. Dadurch wird sichergestellt, dass nicht nur der aktuelle Stand von Forschungsprojekten, sondern auch der jüngste Wissensstand über Signalwege in die Analysen einfließt. Bei Nutzung durch ein Unternehmen ist es darüber hinaus empfehlenswert, diejenigen firmeninternen Forschungsprojekte zusätzlich in die Datenbank zu integrieren, die aufgrund von Vertraulichkeit noch nicht in der Pharmaprojects Rohdatenbasis enthalten sind.

8.2.2.2 Restriktionen der Rohdaten

Bei der Verwendung der Datenbank sowie der Interpretation der Rohdaten muss berücksichtigt werden, dass nicht alle theoretisch wünschenswerten Daten in die Datenbank integriert werden konnten.

Basierend auf den Pharmaprojects Rohdaten (Informa UK Ltd, 2008) konnten nur diejenigen Forschungsprojekte in die Datenbank integriert werden, die öffentlich bekannt sind und Pharmaprojects vorliegen. Das bedeutet, dass insbesondere präklinische Forschungsprojekte, die nicht über die Grenzen der jeweiligen Unternehmen hinaus bekannt sind, in der Analyse fehlen und damit gegebenenfalls Verbindungen nicht detektiert werden können. Es konnten außerdem nur diejenigen Projekte in die Datenbank aufgenommen werden, deren Zielmoleküle durch eine Gene ID-Nummer eindeutig gekennzeichnet sind. Damit fallen die Forschungsprojekte aus der Analyse heraus, deren Zielmolekül zum Beispiel noch unbekannt ist. Forschungsprojekte, die sich gegen chemische Intermediate richten, konnten ebenfalls nicht integriert werden, da eine eindeutige

Identifizierung der Intermediate vergleichbar zur Gene ID-Nummer nicht möglich war.

In diesem Zusammenhang muss weiterhin diskutiert werden, dass sich bei der Erstellung der Datenbank auf funktionelle Signalwege beschränkt wurde. Metabolische Signalwege wurden aufgrund der zahlreichen chemischen Intermediate, die in ihnen vorkommen, nicht integriert. Spezifische Krankheitssignalwege (zum Beispiel der Signalweg der Alzheimer Erkrankung) wurden ebenfalls nicht in die Datenbank aufgenommen. Grundsätzlich ist die Einbindung von weiteren Signalwegen durch Erweiterung der Inputtabellen technisch einfach möglich. Es muss jedoch berücksichtigt werden, dass eine zu hohe Anzahl von Signalwegen möglicherweise zu einer „alles mit allem"-Vernetzung führen könnte, deren Interpretation dann nur noch nach Gewichtung möglich ist.

8.2.2.3 Grundeigenschaften von therapeutischen Gebieten, Indikationsgruppen und Signalwegen

Therapeutische Gebiete, Indikationsgruppen und Signalwege sind keine künstlichen Einheiten mit einer von außen definierten Größe sondern natürlich gewachsene Komplexe, die sich in ihren Eigenschaften unterscheiden. Zum Beispiel gibt es große therapeutische Gebiete, die viele verschiedene Indikationsgruppen mit vielen Indikationen enthalten und kleine therapeutische Gebiete mit wenigen Indikationsgruppen und Indikationen. Dasselbe gilt für Signalwege. Aufgrund ihrer Größe haben bestimmte therapeutische Gebiete *a priori* eine höhere Chance sich mit anderen Gebieten zu verbinden. Genauso hat ein größerer Signalweg aufgrund seiner Anzahl an Zielmolekülen eine höhere Wahrscheinlichkeit vernetzend zu wirken als ein kleiner Signalweg. Um diese Effekte zu kontrollieren, wurden Regressionsanalysen durchgeführt. Darüber hinaus wurde bei Erstellung der Signalwegranglisten der Parameter des relevanten Anteils eingeführt. Nichts desto trotz fließen die spezifischen Eigenschaften der Grundgrößen in die Analysen ein und müssen bei der Interpretation der Analyseergebnisse berücksichtigt werden.

8.2.2.4 Abbildung von Signalwegen

Die in dieser Arbeit entwickelte Datenbank stellt einen Signalweg als funktionelle Gruppe von Zielmolekülen dar. In den Analysen der ersten vier Schritte, die innerhalb der relationalen Datenbank stattfinden, werden die Positionen der einzelnen Zielmoleküle innerhalb eines Signalwegs sowie ihr Verhältnis untereinander nicht berücksichtigt. Das bedeutet, dass es für die Analyse einer Verbindung zwischen verschiedenen Gebieten nicht von Bedeutung ist, ob sich ein Zielmolekül beispielsweise „upstream“ oder „downstream“ eines direkt indikationsübergreifenden Zielmoleküls befindet. Eine Berücksichtigung der Position im Signalweg erfolgt erst im Rahmen der Priorisierung von Kandidatenmolekülen.

Die Integration von Informationen über die Signalwegarchitektur, insbesondere die Position und die Ordnung der Zielmoleküle im Signalweg, ist ein wünschenswerter Aspekt, der bei einer Erweiterung der Datenbank berücksichtigt werden sollte. Je genauer ein Signalweg in der Datenbank abgebildet ist, desto besser wird das Ergebnis der Datenbank sein.

8.3 Chancen und Herausforderungen von indikationsübergreifenden Zusammenhängen für die pharmazeutische Industrie

Trotz der Zunahme des verfügbaren Wissens über molekulare Zusammenhänge und zelluläre Funktionen, sieht sich die pharmazeutische Forschung und Entwicklung einer schwierigen Situation gegenüber (Butcher, 2005; Araujo u. a., 2007; Hellerstein, 2008b). Steigende Kosten (Lynch u. a., 2004; Araujo u. a., 2007), Rückgang der Forschung- und Entwicklungsproduktivität (Couzin, 2002; Butcher, 2005) sowie Schwierigkeiten, Medikamente gegen neue Zielmoleküle zu entwickeln (Zambrowicz und Sands, 2003; Araujo u. a., 2007), belasten die Pharmaindustrie (vergleiche Abschnitt 3.2.2).

8.3.1 Indikationsübergreifende Zusammenhänge als Chance für Forschung und Entwicklung

Im Zusammenhang mit dieser schwierigen Situation ist es naheliegend, nach Ansätzen zu suchen, die helfen können, das verfügbare Potential an Zielmolekülen maximal auszuschöpfen. Aufgrund der zuvor beschriebenen hohen Unsicherheit von neuen Zielmolekülen und einer Scheiterquote von größer 90 % in der klinischen Entwicklung (Brown und Superti-Furga, 2003), rücken insbesondere Zielmoleküle in den Mittelpunkt der Aufmerksamkeit, die bereits in mindestens einer Indikation untersucht werden und damit eine höhere Erfolgssicherheit bieten als vollständig neue Zielmoleküle.

Wenn diese Moleküle für weitere Indikationen genutzt werden können, ist das nicht nur von medizinischem Nutzen für die Patienten sondern zahlt sich auch für das entwickelnde Unternehmen wirtschaftlich aus. Im Rahmen der Entwicklung einer zusätzlichen Indikation ist es möglich, die Entwicklungsphasen Zielmolekülidentifizierung („Target Discovery“), Zielmolekülvalidierung („Target Validation“) und die Identifizierung einer Leitsubstanz („Lead Discovery“) zunächst zu überspringen und existierende Leitsubstanzen aus der ersten Indikation ebenfalls in der neuen Indikation zu testen. Gegebenenfalls

müssen dann – je nach Wirkung der Leitsubstanz – einzelne Aspekte der Zielmolekülvalidierung ergänzt werden oder die Struktur der Leitsubstanz optimiert werden („Lead optimization“). Analog zum zellbasierten systembiologischen Forschungsansatz von Butcher würde ein solches Entwicklungsprogramm Millioneninvestitionen im dreistelligen Bereich einsparen (Butcher, 2005).

8.3.2 Indikationsübergreifende Zusammenhänge und Nebenwirkungsrisiko

Grundlegender Ansatz zur Identifizierung von indikationsübergreifenden Zusammenhängen mit der entwickelten Datenbank ist, dass Signalwegen, die Zielmoleküle enthalten, die in verschiedenen therapeutischen Gebieten eine Rolle spielen, eine erhöhte Chance für indikationsübergreifende Funktion zugeschrieben wird. Durch eben diese verbindenden Signalwege geht eine indikationsübergreifende Funktion *per se* auch immer mit einem erhöhten Risiko von unerwünschten Wirkungen (sogenannten Nebenwirkungen) einher. Im konkreten Beispiel bedeutet das, dass die Einnahme von Viagra® zur Behandlung von erektiler Dysfunktion automatisch auch zu einer Wirkung auf den Blutdruck führt (Pfizer Pharma GmbH, 2008a) – schließlich ist derselbe Wirkstoff auch als Revatio® für die Behandlung von pulmonalem Bluthochdruck zugelassen (Ghofrani u. a., 2004a; Ghofrani u. a., 2006; Pfizer Inc., 2008b; Pfizer Pharma GmbH, 2008b). Zum Teil lässt sich diese Wirkung auf andere physiologische Vorgänge durch Veränderung der chemischen Struktur der Substanz und, sofern möglich, durch lokale Applikation beheben. So kann das Anti-VEGF Fab-Fragment Ranibizumab (Lucentis®) beispielsweise außerhalb des Auges durch seine kurze Halbwertszeit kaum zu unerwünschten systemischen Nebenwirkungen führen (Ferrara u. a., 2006). Darüber hinaus kann durch eine lokale, intra-vitreale Injektion eine hohe Wirkstoffkonzentration im Auge erreicht werden, ohne dass diese den Rest des Organismus in vollem Maß belastet (Novartis International AG, 2008).

Ob der Effekt der indikationsübergreifenden Wirkungen im Einzelnen problematisch ist, muss für jeden Einzelfall separat bewertet werden.

Grundsätzlich lassen sich in der wissenschaftlichen Literatur Anzeichen für einen Paradigma-Wechsel erkennen: Wurde noch vor einigen Jahren große Hoffnung in die Bedeutung von Substanzen, die gezielt nur eine Krankheit behandeln können, gesetzt (sogenannte „Magic bullets“), so zeichnet sich nun die Erkenntnis ab, dass die Entwicklung dieser Medikamente aufgrund der Komplexität des Organismus und seiner vielfältigen Interdependenzen nur selten erfolgreich ist (Kitano, 2007). Darüber hinaus konnten sich einzelne Netzwerkknoten-Zielmoleküle, wie zum Beispiel Survivin, gerade aufgrund ihrer wichtigen Position in Signalwegen als geeignete Zielmoleküle ausweisen (Altieri, 2008).

Für die praktische Anwendung des in dieser Arbeit entwickelten Instruments ist es darüber hinaus denkbar, die Untersuchung der Vernetzungen aktiv zur Vorhersage von Nebenwirkungsrisiken zu nutzen. Eine solche Funktion ist bei einer Erweiterung des Instruments wünschenswert.

8.3.3 Spezifische Herausforderungen indikationsübergreifender Substanzen

Unternehmen, die indikationsübergreifende Substanzen entwickeln und diese dann – gegebenenfalls auch unter unterschiedlichem Medikamentennamen – auf den Markt bringen wollen, sehen sich spezifischen Herausforderungen gegenüber. Ein Beispiel für eine solche Situation kann bei den beiden anti-angiogenetischen Medikamenten Lucentis® und Avastin® beobachtet werden. Während der humanisierte monoklonale Antikörper Bevacizumab (Handelsname Avastin®) nur für die Behandlung onkologischer Indikationen zugelassen ist (Ferrara u. a., 2007; Genentech Inc., 2008c; Hoffmann-La Roche, 2008a; Hoffmann-La Roche, 2008b), hat das kleinere, stark verwandte Fab-Fragment Ranibizumab (Handelsname Lucentis®) die Zulassung zur Behandlung altersabhängiger Makuladegeneration (AMD) (Genentech Inc., 2008b; Novartis International AG, 2008). Eine intra-vitreale Injektion von Lucentis® kostet jedoch mit fast 2000 US $ derartig viel (Berenson, 2006), dass verstärkt die billigere Substanz Bevacizumab im Auge zur AMD-Behandlung verwendet wird (Fung u. a., 2006; Wu u. a., 2008). Mit Kosten von 30 US $ pro Injektion (Berenson, 2006) können damit 98 % der Substanzkosten gespart werden. Diese Nutzung ohne Zulassung in

der Indikation wird als „Off-Label“-Gebrauch bezeichnet und ist nicht nur für die entwickelnde Firma ein finanzielles Desaster. Insbesondere die Patienten, die aus Kostengründen Avastin® ins Auge injiziert bekommen, übernehmen ein hohes Risiko, da der Hersteller bei einem wissentlichen „Off-Label“-Gebrauch bei möglichen Problemen nicht haftbar gemacht werden kann (Rücker, 2007; Kuhrt, 2008). Im konkreten Beispiel wird diesem Problem nun durch groß angelegte klinische Studien entgegengewirkt, die einen Vergleich zwischen der Wirkung und der Verträglichkeit von Avastin® und Lucentis® bei AMD untersuchen (Klinikum Bremen-Mitte Institut für Pharmakologie, 2008; National Eye Institute, 2008). Sollte aus diesen Studien das Ergebnis hervorgehen, dass das günstigere Avastin® gleichermaßen im Auge verwendet werden kann, würde das für die Firmen Genentech und Novartis, die Lucentis® in den USA und in Europa vertreiben, deutliche finanzielle Einbußen bedeuten. Um dem entgegenzuwirken, versuchen die Firmen verschiedene Ansätze. In den USA hatte Genentech im Herbst 2007 versucht, die Verteilung von Avastin® an bestimmte Apotheken, die die Substanz für ophthalmologische Anwendungen portionieren können, zu unterbinden (Genentech Inc., 2007a). Dieser Vorstoß scheiterte am großen Widerstand der amerikanischen Akademie für Augenheilkunde (American Academy of Ophthalmology, 2007; Genentech Inc., 2007b). In Deutschland, wo Lucentis® von der Firma Novartis vertrieben wird, gab es Ende 2007 ebenfalls sehr umstrittene Bestrebungen, einen „Off-Label“-Gebrauch dauerhaft zu tolerieren (Rücker, 2007; Kuhrt, 2008). Im Juni 2008 wurde durch einen Vorstoß der Firma Novartis die Verwendung von Lucentis® zumindest vorerst abgesichert: Durch einen Rahmenvertrag mit der gesetzlichen Krankenkasse AOK verpflichtet sich Novartis, sich im dreistelligen Millionenbereich an den durch die Lucentis®-Behandlung entstehenden Kosten zu beteiligen (Seidlitz, 2008).

Um bei der Entwicklung und Markteinführung von indikationsübergreifenden Medikamenten Komplikationen durch „Off-Label“-Gebrauch und mögliche gesetzliche Preislimitationen zu verhindern, ist eine strategische Planung des klinischen Entwicklungsprozesses besonders wichtig. Integration von Produktdifferenzierung (zum Beispiel durch unterschiedliche Formulierungen, Konzentrationen und Applikationssysteme), sowie eine Fokussierung auf

wirtschaftlich besonders attraktive Indikationen in ersten Schritt, können mögliche Strategien sein. Neben dem Negativbeispiel von Avastin® und Lucentis® können andere Fallbeispiele zeigen, dass selbst ohne Produktdifferenzierung ein positiver wirtschaftlicher Effekt für die jeweilige Firma möglich ist. Die Markteinführung von Sildenafil als Revatio® zur Behandlung von pulmonaler Hypertonie beispielsweise brachte Pfizer im Jahr 2007 einen Zusatzumsatz von 200 Mio. US $ ein (Pfizer Inc., 2008a).

8.3.4 Kommunikation zwischen Forschungsgruppen als zentraler Erfolgsfaktor

Die Identifikation von geeigneten Zielmolekülkandidaten ist nur der erste Schritt zur erfolgreichen Entwicklung von indikationsübergreifenden Medikamenten. Um Forschung und Entwicklung so auszurichten, dass indikationsübergreifende Zusammenhänge optimal genutzt werden, sind eine intensive Zusammenarbeit und eine offene Kommunikation zwischen allen Arbeitsgruppen eines forschenden Unternehmens notwendig. Der Austausch über Arbeitsgruppen hinaus mag theoretisch selbstverständlich erscheinen, ist in der Praxis aber durch die historisch gewachsenen Einheiten der therapeutischen Gebiete erschwert. Viele Pharmaunternehmen haben sich entlang therapeutischer Gebiete aufgestellt, die jeweils eigene Forschungs- und Entwicklungsabteilungen haben. Eine Arbeitsgruppe, die jedoch nur für ein therapeutisches Gebiet arbeitet, kann allein keine indikationsübergreifende Entwicklung bewerkstelligen. Die aktive Förderung der Kommunikation und der Zusammenarbeit zwischen den einzelnen Bereichen ist daher ein zentraler Erfolgsfaktor für die Entwicklung indikationsübergreifender Medikamente. Einzelne Maßnahmen, wie zum Beispiel gemeinsame Workshops oder interne Konferenzen, sind dabei genauso wichtig wie die Bereitstellung angemessener Anreizsysteme. Dabei muss sichergestellt werden, dass es für eine Arbeitsgruppe besonders attraktiv ist, gemeinsam mit einer anderen einen indikationsübergreifenden Zusammenhang zu erforschen.

8.4 Fazit und Ausblick

Trotz großem Wissenszuwachs in den vergangenen Jahrzehnten muss die pharmazeutische Industrie im gleichen Zeitraum einem deutlichen Rückgang der Forschungs- und Entwicklungsproduktivität hinnehmen (Butcher, 2005; Fishman und Porter, 2005; Sams-Dodd, 2006; Aggarwal u. a., 2007; Araujo u. a., 2007). Auch „high-throughput screening"-Techniken und die Fokussierung auf gezielte Modulation einzelner Proteine zur spezifischen Behandlung von Krankheiten führten nur zu vereinzelten Erfolgsgeschichten dieses „genes to drugs"-Ansatzes (Butcher u. a., 2004; Araujo u. a., 2007; Hellerstein, 2008b; Hellerstein, 2008a). Die Entwicklung von Medikamenten gegen neue Zielmoleküle ist noch immer ausgesprochen schwierig und die meisten Forschungsprojekte scheitern im Verlauf der klinischen Entwicklung (Brown und Superti-Furga, 2003; Duyk, 2003).

Dies scheint Grund genug zu sein, Zweifel am aktuellen Forschungsparadigma der pharmazeutischen Industrie aufkommen zu lassen. In unterschiedlicher Radikalität werden Veränderungen des Forschungsablaufes diskutiert. Extreme Ansätze fordern eine Rückbesinnung auf Beobachtungen an *in vivo* Modellen (Hellerstein, 2008b; Hellerstein, 2008a). Andere postulieren die Verwendung von primären humanen Zellsystemen für Forschung und Entwicklung (Butcher u. a., 2004; Butcher, 2005; Sams-Dodd, 2006). Gemeinsame Grundlage dieser Diskussionen ist die Betonung zweier Aspekte des menschlichen Organismus: Zum einen sind funktionelle Zusammenhänge derart komplex und interdependent, dass es nur sehr schwer möglich ist, mit dem Eingriff an einem einzelnen Zielmolekül selektiv eine Krankheit zu behandeln (Apic u. a., 2005; Araujo u. a., 2007; Kitano, 2007; Hellerstein, 2008b). Zum anderen zeigen Signalwege die Fähigkeit, Störungen von außen durch verschiedene Regulationsmechanismen auszugleichen. Dieses als Robustheit bezeichnete Phänomen dient dazu sicherzustellen, dass komplexe funktionelle Zusammenhänge auch dann ablaufen, wenn eine einzelne Komponente des Signalwegs nicht funktionsfähig ist (Kitano, 2007; Hellerstein, 2008b). Im Umkehrschluss bedeutet dies für die pharmazeutische Forschung und Entwicklung, dass ein erfolgreicher Eingriff in

einen funktionellen Zusammenhang häufig nicht durch den Angriff an einem einzelnen Zielmolekül bewerkstelligt werden kann (Apic u. a., 2005; Araujo u. a., 2007; Hellerstein, 2008a; Jones, 2008; Schrattenholz und Soskic, 2008; Watterson u. a., 2008).

Als zentraler Erfolgsfaktor zur Gewährleistung erfolgreicher pharmazeutischer Forschung und Entwicklung gilt die Integration systembiologischer Informationen aus Genom, Proteom, metabolischen Zusammenhängen und funktionellen Signalwegen in die Medikamentenentwicklung (Bugrim u. a., 2004; Apic u. a., 2005; Butcher, 2005; Araujo u. a., 2007; Hellerstein, 2008b; Schrattenholz und Soskic, 2008; Watterson u. a., 2008). Mit Hilfe unterschiedlicher Werkzeuge kann so die Komplexität der Zusammenhänge modelliert werden und ein Eingriff durch eine pharmazeutisch aktive Substanz sicher, günstig und schnell simuliert werden (Watterson u. a., 2008). Im Rahmen der Medikamentenentwicklung kann auf diese Weise leichter eine Vorhersage getroffen werden, ob es sich zum Beispiel bei einem Zielmolekül um einen geeigneten Angriffspunkt für eine pharmazeutische Intervention handelt. Darüber hinaus hilft die Modellierung das Nebenwirkungsrisiko abzuschätzen und molekulare Ursachen einer möglichen Toxizität zu verstehen (Apic u. a., 2005).

Für die langfristige Entwicklung von pharmazeutischer Forschung und Entwicklung ist eine Verschiebung vom einzelnen Protein als Zielmolekül („protein drug target“) hin zum gesamten Signalweg als Ziel für pharmazeutische Intervention („pathway drug target“) zu erwarten (Aggarwal u. a., 2007; Araujo u. a., 2007; Jones, 2008). Dabei gilt, dass die konzertierte Modulation mehrerer Moleküle innerhalb eines Signalwegs durch eine Kombination von Substanzen gegenüber dem spezifischen Angriff an einzelnen Zielmolekülen an Bedeutung gewinnt (Aggarwal u. a., 2007; Araujo u. a., 2007; Hellerstein, 2008b; Hellerstein, 2008a; Watterson u. a., 2008). Die Hoffnung, mit Signalwegen als neuer „Grammatik der Medikamentenentwicklung“ (Fishman und Porter, 2005) zum Beispiel Krebserkrankungen besser behandeln zu können, ist groß (Aggarwal u. a., 2007; Jones, 2008).

Die in dieser Arbeit entwickelte Methodik ergänzt die bisher verfügbaren systembiologischen Ansätze, die sich in der Regel auf die Integration von experimentellen Daten in Signalwegmodelle beschränken, um die Fähigkeit, Beobachtungen aus dem Forschungs- und Entwicklungsalltag der Industrie mit biologischen Zusammenhängen zu kombinieren. Neben den unmittelbar nutzbaren Funktionen des Instruments für die Forschung und Entwicklung – wie der Ermittlung von Kandidaten für indikationsübergreifende Zielmoleküle und der Priorisierung von Signalwegen für bestimmte Forschungsschwerpunkte – ist dies ein weiterer Schritt zur Integration der Systembiologie in die Medikamentenentwicklung.

9 Literaturverzeichnis

ABOU-SHOUSHA, S. A., SALAH, E. und WAGDY, E.: Study of P53 in peripheral blood and synovial mononuclear cells of rheumatoid arthritis and osteoarthritis patients and its relation to the degree of disease activity. *Egypt J Immunol* (2005), **12** (2): 61-70

AGGARWAL, B. B., SETHI, G., BALADANDAYUTHAPANI, V., KRISHNAN, S. und SHISHODIA, S.: Targeting cell signaling pathways for drug discovery: an old lock needs a new key. *J Cell Biochem* (2007), **102** (3): 580-92

AIELLO, L. P., AVERY, R. L., ARRIGG, P. G., KEYT, B. A., JAMPEL, H. D., SHAH, S. T., PASQUALE, L. R., THIEME, H., IWAMOTO, M. A., PARK, J. E. und ET AL.: Vascular endothelial growth factor in ocular fluid of patients with diabetic retinopathy and other retinal disorders. *N Engl J Med* (1994), **331** (22): 1480-7

ALON, T., HEMO, I., ITIN, A., PE'ER, J., STONE, J. und KESHET, E.: Vascular endothelial growth factor acts as a survival factor for newly formed retinal vessels and has implications for retinopathy of prematurity. *Nat Med* (1995), **1** (10): 1024-8

ALP, S., SKRYGAN, M., SCHMIDT, W. E. und BASTIAN, A.: Sildenafil improves hemodynamic parameters in COPD--an investigation of six patients. *Pulm Pharmacol Ther* (2006), **19** (6): 386-90

ALTIERI, D. C.: Survivin, cancer networks and pathway-directed drug discovery. *Nat Rev Cancer* (2008), **8** (1): 61-70

AMERICAN ACADEMY OF OPHTHALMOLOGY, Ophthalmologists Concerned About AMD Patients' Access To Avastin, verfügbar unter http://www.aao.org./newsroom/release/20071012.cfm (2007)

APIC, G., IGNJATOVIC, T., BOYER, S. und RUSSELL, R. B.: Illuminating drug discovery with biological pathways. *FEBS Lett* (2005), **579** (8): 1872-7

ARAUJO, R. P., LIOTTA, L. A. und PETRICOIN, E. F.: Proteins, drug targets and the mechanisms they control: the simple truth about complex networks. *Nat Rev Drug Discov* (2007), **6** (11): 871-80

ARIADNE INC., Homepage, verfügbar unter http://www.ariadnegenomics.com/ (2008)

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. und SHERLOCK, G.: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* (2000), **25** (1): 25-9

AURON, P. E., WEBB, A. C., ROSENWASSER, L. J., MUCCI, S. F., RICH, A., WOLFF, S. M. und DINARELLO, C. A., (1984), Vol. 81, pp. 7907-7911

AZADZOI, K. M., KIM, N., BROWN, M. L., GOLDSTEIN, I., COHEN, R. A. und SAENZ DE TEJADA, I.: Endothelium-derived nitric oxide and cyclooxygenase

products modulate corpus cavernosum smooth muscle tone. *J Urol* (1992), **147** (1): 220-5

BADER, G. D., BETEL, D. und HOGUE, C. W.: BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res* (2003), **31** (1): 248-50

BAIER, A., MEINECKEL, I., GAY, S. und PAP, T.: Apoptosis in rheumatoid arthritis. *Curr Opin Rheumatol* (2003), **15** (3): 274-9

BARBERA, J. A., PEINADO, V. I. und SANTOS, S.: Pulmonary hypertension in chronic obstructive pulmonary disease. *Eur Respir J* (2003), **21** (5): 892-905

BERENSON, A.: A cancer drug shows promise, at a price that many can't pay. *NY Times (Print)* (2006), A1, C2

BOHLE, R. M., HARTMANN, E., KINFE, T., ERMERT, L., SEEGER, W. und FINK, L.: Cell type-specific mRNA quantitation in non-neoplastic tissues after laser-assisted cell picking. *Pathobiology* (2000), **68** (4-5): 191-5

BOOLELL, M., ALLEN, M. J., BALLARD, S. A., GEPI-ATTEE, S., MUIRHEAD, G. J., NAYLOR, A. M., OSTERLOH, I. H. und GINGELL, C.: Sildenafil: an orally active type 5 cyclic GMP-specific phosphodiesterase inhibitor for the treatment of penile erectile dysfunction. *Int J Impot Res* (1996a), **8** (2): 47-52

BOOLELL, M., GEPI-ATTEE, S., GINGELL, J. C. und ALLEN, M. J.: Sildenafil, a novel effective oral therapy for male erectile dysfunction. *Br J Urol* (1996b), **78** (2): 257-61

BRANER, D. A., FINEMAN, J. R., CHANG, R. und SOIFER, S. J.: M&B 22948, a cGMP phosphodiesterase inhibitor, is a pulmonary vasodilator in lambs. *Am J Physiol* (1993), **264** (1 Pt 2): H252-8

BRAY, T., PAOLI, J., SPERBERG-MCQUEEN, C. M., MALER, E. und YERGEAU, F., Extensible Markup Language (XML) 1.0 (Fourth Edition), World Wide Web Consortium recommendation 16 August 2006, edited in place 29 September 2006, verfügbar unter http://www.w3.org/TR/xml/ (2006)

BROWN, D. und SUPERTI-FURGA, G.: Rediscovering the sweet spot in drug discovery. *Drug Discov Today* (2003), **8** (23): 1067-77

BROWN, D. M., KAISER, P. K., MICHELS, M., SOUBRANE, G., HEIER, J. S., KIM, R. Y., SY, J. P. und SCHNEIDER, S.: Ranibizumab versus verteporfin for neovascular age-related macular degeneration. *N Engl J Med* (2006), **355** (14): 1432-44

BRUN, J.: Proteasome inhibition as a novel therapy in treating rheumatoid arthritis. *Med Hypotheses* (2008), **71** (1): 65-72

BUERK, D. G., ANCES, B. M., GREENBERG, J. H. und DETRE, J. A.: Temporal dynamics of brain tissue nitric oxide during functional forepaw stimulation in rats. *Neuroimage* (2003), **18** (1): 1-9

BUGRIM, A., NIKOLSKAYA, T. und NIKOLSKY, Y.: Early prediction of drug metabolism and toxicity: systems biology approach and modeling. *Drug Discov Today* (2004), **9** (3): 127-35

BUNT, S. K., YANG, L., SINHA, P., CLEMENTS, V. K., LEIPS, J. und OSTRAND-ROSENBERG, S.: Reduced inflammation in the tumor microenvironment

delays the accumulation of myeloid-derived suppressor cells and limits tumor progression. *Cancer Res* (2007), **67** (20): 10019-26

BUSH, P. A., ARONSON, W. J., BUGA, G. M., RAJFER, J. und IGNARRO, L. J.: Nitric oxide is a potent relaxant of human and rabbit corpus cavernosum. *J Urol* (1992), **147** (6): 1650-5

BUTCHER, E. C., BERG, E. L. und KUNKEL, E. J.: Systems biology in drug discovery. *Nat Biotechnol* (2004), **22** (10): 1253-9

BUTCHER, E. C.: Can cell systems biology rescue drug discovery? *Nat Rev Drug Discov* (2005), **4** (6): 461-7

CHEN, P. P.-S.: The entity-relationship model: toward a unified view of data. *ACM SIGIR Forum* (1975), **10** (3): 9-9

CHEN, Q., CASALI, B., PATTACINI, L., BOIARDI, L. und SALVARANI, C.: Tumor necrosis factor-alpha protects synovial cells from nitric oxide induced apoptosis through phosphoinositide 3-kinase Akt signal transduction. *J Rheumatol* (2006), **33** (6): 1061-8

CHEN, X., JI, Z. L. und CHEN, Y. Z.: TTD: Therapeutic Target Database. *Nucleic Acids Res* (2002), **30** (1): 412-5

CHOU, C. T., YANG, J. S. und LEE, M. R.: Apoptosis in rheumatoid arthritis--expression of Fas, Fas-L, p53, and Bcl-2 in rheumatoid synovial tissues. *J Pathol* (2001), **193** (1): 110-6

CODD, E. F.: A relational model of data for large shared data banks. *Communications of the ACM* (1970), **13** (6): 377-387

CODD, E. F.: Further normalization of the data base relational model. *Courant Computer Science Symposium 6: Data Base Systems* (1971), 65-98.

COHEN, A. H., HANSON, K., MORRIS, K., FOUTY, B., MCMURTY, I. F., CLARKE, W. und RODMAN, D. M.: Inhibition of cyclic 3'-5'-guanosine monophosphate-specific phosphodiesterase selectively vasodilates the pulmonary circulation in chronically hypoxic rats. *J Clin Invest* (1996), **97** (1): 172-9

CORBIN, J. D., BEASLEY, A., BLOUNT, M. A. und FRANCIS, S. H.: High lung PDE5: a strong basis for treating pulmonary hypertension with PDE5 inhibitors. *Biochem Biophys Res Commun* (2005), **334** (3): 930-8

COUZIN, J.: Cancer drugs. Smart weapons prove tough to design. *Science* (2002), **298** (5593): 522-5

CRAWFORD, L. V., PIM, D. C. und LAMB, P.: The cellular protein p53 in human tumours. *Mol Biol Med* (1984), **2** (4): 261-72

DE JONG, P. T.: Age-related macular degeneration. *N Engl J Med* (2006), **355** (14): 1474-85

DELOBEL, C.: Normalization and hierarchical dependencies in the relational data model. *ACM Transactions on Database Systems* (1978), **3** (3): 201-222

DICKSON, M. und GAGNON, J. P.: Key factors in the rising cost of new drug discovery and development. *Nat Rev Drug Discov* (2004), **3** (5): 417-29

DINARELLO, C. A.: The interleukin-1 family: 10 years of discovery. *Faseb J* (1994), **8** (15): 1314-25

DINSMORE, W. W., HODGES, M., HARGREAVES, C., OSTERLOH, I. H., SMITH, M. D. und ROSEN, R. C.: Sildenafil citrate (Viagra) in erectile dysfunction: near normalization in men with broad-spectrum erectile dysfunction compared with age-matched healthy control subjects. *Urology* (1999), **53** (4): 800-5

DRAGHICI, S., KHATRI, P., TARCA, A. L., AMIN, K., DONE, A., VOICHITA, C., GEORGESCU, C. und ROMERO, R.: A systems biology approach for pathway level analysis. *Genome Res* (2007), **17** (10): 1537-45

DREWS, J.: Genomic sciences and the medicine of tomorrow. *Nat Biotechnol* (1996), **14** (11): 1516-8

DREWS, J. und RYSER, S.: The role of innovation in drug development. *Nat Biotechnol* (1997), **15** (13): 1318-9

DREWS, J.: Drug discovery: a historical perspective. *Science* (2000), **287** (5460): 1960-4

DUYK, G.: Attrition and translation. *Science* (2003), **302** (5645): 603-5

EKINS, S., MESTRES, J. und TESTA, B.: In silico pharmacology for drug discovery: applications to targets and beyond. *Br J Pharmacol* (2007), **152** (1): 21-37

EL-SAYED, Z. A., FARAG, D. H. und EISSA, S.: Tumor suppressor protein p53 and anti-p53 autoantibodies in pediatric rheumatological diseases. *Pediatr Allergy Immunol* (2003), **14** (3): 229-33

ELARAJ, D. M., WEINREICH, D. M., VARGHESE, S., PUHLMANN, M., HEWITT, S. M., CARROLL, N. M., FELDMAN, E. D., TURNER, E. M. und ALEXANDER, H. R., (2006), Vol. 12, pp. 1088-1096

ELLIS, L. M. und HICKLIN, D. J.: VEGF-targeted therapy: mechanisms of anti-tumour activity. *Nat Rev Cancer* (2008), **8** (8): 579-91

ENGELS, E. A., WU, X., GU, J., DONG, Q., LIU, J. und SPITZ, M. R., (2007), Vol. 67, pp. 6520-6527

FAGENHOLZ, P. J., GUTMAN, J. A., MURRAY, A. F. und HARRIS, N. S.: Treatment of high altitude pulmonary edema at 4240 m in Nepal. *High Alt Med Biol* (2007), **8** (2): 139-46

FAGIN, R.: Multivalued dependencies and a new normal form for relational databases. *ACM Transactions on Database Systems* (1977), **2** (3): 262-278

FDA: Innovation and stagnation: challenge and opportunity on the critical path to new medical products (2004)

FERRARA, N. und HENZEL, W. J.: Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. *Biochem Biophys Res Commun* (1989), **161** (2): 851-8

FERRARA, N., HILLAN, K. J., GERBER, H. P. und NOVOTNY, W.: Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. *Nat Rev Drug Discov* (2004), **3** (5): 391-400

FERRARA, N., DAMICO, L., SHAMS, N., LOWMAN, H. und KIM, R.: Development of ranibizumab, an anti-vascular endothelial growth factor antigen binding fragment, as therapy for neovascular age-related macular degeneration. *Retina* (2006), **26** (8): 859-70

FERRARA, N., MASS, R. D., CAMPA, C. und KIM, R.: Targeting VEGF-A to treat cancer and age-related macular degeneration. *Annu Rev Med* (2007), **58** 491-504

FINZEL, B. C., CLANCY, L. L., HOLLAND, D. R., MUCHMORE, S. W., WATENPAUGH, K. D. und EINSPAHR, H. M.: Crystal structure of recombinant human interleukin-1 beta at 2.0 A resolution. *J Mol Biol* (1989), **209** (4): 779-91

FISHMAN, M. C. und PORTER, J. A.: Pharmaceuticals: a new grammar for drug discovery. *Nature* (2005), **437** (7058): 491-3

FOLKMAN, J.: Tumor angiogenesis: therapeutic implications. *N Engl J Med* (1971), **285** (21): 1182-6

FOLKMAN, J.: Angiogenesis: an organizing principle for drug discovery? *Nat Rev Drug Discov* (2007), **6** (4): 273-86

FRIES, R., SHARIAT, K., VON WILMOWSKY, H. und BOHM, M.: Sildenafil in the treatment of Raynaud's phenomenon resistant to vasodilatory therapy. *Circulation* (2005), **112** (19): 2980-5

FUNG, A. E., ROSENFELD, P. J. und REICHEL, E.: The International Intravitreal Bevacizumab Safety Survey: using the internet to assess drug safety worldwide. *Br J Ophthalmol* (2006), **90** (11): 1344-9

GALIE, N., GHOFRANI, H. A., TORBICKI, A., BARST, R. J., RUBIN, L. J., BADESCH, D., FLEMING, T., PARPIA, T., BURGESS, G., BRANZI, A., GRIMMINGER, F., KURZYNA, M. und SIMONNEAU, G.: Sildenafil citrate therapy for pulmonary arterial hypertension. *N Engl J Med* (2005), **353** (20): 2148-57

GENEGO BIOINFORMATICS SOFTWARE INC., Homepage, verfügbar unter http://www.genego.com/metacore.php (2008)

GENENTECH INC., Letter to Physicians, verfügbar unter http://www.gene.com/gene/features/avastin/letter-to-physicians.html (2007a)

GENENTECH INC., Open Letter, verfügbar unter http://www.gene.com/gene/features/avastin/open-letter.html (2007b)

GENENTECH INC., Genentech Homepage, Historical Product Sales, verfügbar unter http://www.gene.com/gene/ir/financials/historical/ (2008a)

GENENTECH INC., Genentech Homepage, Lucentis Produkt Informationen, verfügbar unter http://www.gene.com/gene/products/information/tgr/lucentis/ (2008b)

GENENTECH INC., Avastin Homepage USA, verfügbar unter http://www.avastin.com/avastin/index.jsp (2008c)

GERMAN, Z., CHAMBLISS, K. L., PACE, M. C., ARNET, U. A., LOWENSTEIN, C. J. und SHAUL, P. W.: Molecular basis of cell-specific endothelial nitric-oxide synthase expression in airway epithelium. *J Biol Chem* (2000), **275** (11): 8183-9

GHOFRANI, H. A., PEPKE-ZABA, J., BARBERA, J. A., CHANNICK, R., KEOGH, A. M., GOMEZ-SANCHEZ, M. A., KNEUSSL, M. und GRIMMINGER, F.: Nitric oxide pathway and phosphodiesterase inhibitors in pulmonary arterial hypertension. *J Am Coll Cardiol* (2004a), **43** (12 Suppl S): 68S-72S

GHOFRANI, H. A., REICHENBERGER, F., KOHSTALL, M. G., MROSEK, E. H., SEEGER, T., OLSCHEWSKI, H., SEEGER, W. und GRIMMINGER, F.: Sildenafil increased exercise capacity during hypoxia at low altitudes and at Mount Everest base camp: a randomized, double-blind, placebo-controlled crossover trial. *Ann Intern Med* (2004b), **141** (3): 169-77

GHOFRANI, H. A., OSTERLOH, I. H. und GRIMMINGER, F.: Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. *Nat Rev Drug Discov* (2006), **5** (8): 689-702

GOCKEL, H. R., LUGERING, A., HEIDEMANN, J., SCHMIDT, M., DOMSCHKE, W., KUCHARZIK, T. und LUGERING, N.: Thalidomide induces apoptosis in human monocytes by using a cytochrome c-dependent pathway. *J Immunol* (2004), **172** (8): 5103-9

GOLDSTEIN, I., LUE, T. F., PADMA-NATHAN, H., ROSEN, R. C., STEERS, W. D. und WICKER, P. A.: Oral sildenafil in the treatment of erectile dysfunction. Sildenafil Study Group. *N Engl J Med* (1998), **338** (20): 1397-404

GORDON, M. S., MARGOLIN, K., TALPAZ, M., SLEDGE, G. W., JR., HOLMGREN, E., BENJAMIN, R., STALTER, S., SHAK, S. und ADELMAN, D.: Phase I safety and pharmacokinetic study of recombinant human anti-vascular endothelial growth factor in patients with advanced cancer. *J Clin Oncol* (2001), **19** (3): 843-50

GORE, J. und SILVER, R.: Oral sildenafil for the treatment of Raynaud's phenomenon and digital ulcers secondary to systemic sclerosis. *Ann Rheum Dis* (2005), **64** (9): 1387

GRAGOUDAS, E. S., ADAMIS, A. P., CUNNINGHAM, E. T., JR., FEINSOD, M. und GUYER, D. R.: Pegaptanib for neovascular age-related macular degeneration. *N Engl J Med* (2004), **351** (27): 2805-16

GRIMMINGER, F., SPRIESTERSBACH, R., WEISSMANN, N., WALMRATH, D. und SEEGER, W.: Nitric oxide generation and hypoxic vasoconstriction in buffer-perfused rabbit lungs. *J Appl Physiol* (1995), **78** (4): 1509-15

HASSAN, M. A. und KETAT, A. F.: Sildenafil citrate increases myocardial cGMP content in rat heart, decreases its hypertrophic response to isoproterenol and decreases myocardial leak of creatine kinase and troponin T. *BMC Pharmacol* (2005), **5** 10

HAYNES, J., JR., KITHAS, P. A., TAYLOR, A. E. und STRADA, S. J.: Selective inhibition of cGMP-inhibitable cAMP phosphodiesterase decreases pulmonary vasoreactivity. *Am J Physiol* (1991), **261** (2 Pt 2): H487-92

HEIER, J. S., ANTOSZYK, A. N., PAVAN, P. R., LEFF, S. R., ROSENFELD, P. J., CIULLA, T. A., DREYER, R. F., GENTILE, R. C., SY, J. P., HANTSBARGER, G. und SHAMS, N.: Ranibizumab for treatment of neovascular age-related macular degeneration: a phase I/II multicenter, controlled, multidose study. *Ophthalmology* (2006), **113** (4): 633 e1-4

HELLERSTEIN, M. K.: A critique of the molecular target-based drug discovery paradigm based on principles of metabolic control: advantages of pathway-based discovery. *Metab Eng* (2008a), **10** (1): 1-9

HELLERSTEIN, M. K.: Exploiting complexity and the robustness of network architecture for drug discovery. *J Pharmacol Exp Ther* (2008b), **325** (1): 1-9

HELMCHEN, B., WECKAUF, H., EHEMANN, V., WITTMANN, I., MEYER-SCHOLTEN, C. und BERGER, I.: Expression pattern of cell cycle-related gene products in synovial stroma and synovial lining in active and quiescent stages of rheumatoid arthritis. *Histol Histopathol* (2005), **20** (2): 365-72

HOFFMANN-LA ROCHE: Roche Group Annual Report 2004 (2005)

HOFFMANN-LA ROCHE: Roche Group Annual Report 2005, Part 1 Business Report (2006)

HOFFMANN-LA ROCHE: Roche Group Annual Report 2006, Part 1 Business Report (2007)

HOFFMANN-LA ROCHE, Homepage Roche Pharma, verfügbar unter http://roche.de/pharma (2008a)

HOFFMANN-LA ROCHE, Avastin Homepage Deutschland, verfügbar unter http://www.avastin.de/ (2008b)

HOFFMANN-LA ROCHE: Roche Group Annual Report 2007, Part 1 Business Report (2008c)

HSU, A. R., BARNHOLT, K. E., GRUNDMANN, N. K., LIN, J. H., MCCALLUM, S. W. und FRIEDLANDER, A. L.: Sildenafil improves cardiac output and exercise performance during acute hypoxia, but not normoxia. *J Appl Physiol* (2006), **100** (6): 2031-40

HU, Z., SHAO, M., CHEN, Y., ZHOU, J., QIAN, J., XU, L., MA, H., WANG, X., XU, Y., LU, D. und SHEN, H.: Allele 2 of the interleukin-1 receptor antagonist gene (IL1RN*2) is associated with a decreased risk of primary lung cancer. *Cancer Lett* (2006), **236** (2): 269-75

HUANG, J., GAO, X., LI, S. und CAO, Z.: Recruitment of IRAK to the interleukin 1 receptor complex requires interleukin 1 receptor accessory protein. *Proc Natl Acad Sci U S A* (1997), **94** (24): 12829-32

HWANG, S. Y., KIM, J. Y., KIM, K. W., PARK, M. K., MOON, Y., KIM, W. U. und KIM, H. Y.: IL-17 induces production of IL-6 and IL-8 in rheumatoid arthritis synovial fibroblasts via NF-kappaB- and PI3-kinase/Akt-dependent pathways. *Arthritis Res Ther* (2004), **6** (2): R120-8

ICHIKAWA, K., LIU, W., FLECK, M., ZHANG, H., ZHAO, L., OHTSUKA, T., WANG, Z., LIU, D., MOUNTZ, J. D., OHTSUKI, M., KOOPMAN, W. J., KIMBERLY, R. und ZHOU, T.: TRAIL-R2 (DR5) mediates apoptosis of synovial fibroblasts in rheumatoid arthritis. *J Immunol* (2003), **171** (2): 1061-9

ICHINOSE, F., ADRIE, C., HURFORD, W. E. und ZAPOL, W. M.: Prolonged pulmonary vasodilator action of inhaled nitric oxide by Zaprinast in awake lambs. *J Appl Physiol* (1995), **78** (4): 1288-95

IDE, H., NAKANO, H., OGASA, T., OSANAI, S., KIKUCHI, K. und IWAMOTO, J.: Regulation of pulmonary circulation by alveolar oxygen tension via airway nitric oxide. *J Appl Physiol* (1999), **87** (5): 1629-36

IGNARRO, L. J., BUSH, P. A., BUGA, G. M., WOOD, K. S., FUKUTO, J. M. und RAJFER, J.: Nitric oxide and cyclic GMP formation upon electrical field

stimulation cause relaxation of corpus cavernosum smooth muscle. *Biochem Biophys Res Commun* (1990), **170** (2): 843-50

ILONIDIS, G., PARAPANISIOU, E., ANOGEIANAKI, A., GIAVAZIS, I., THEOFILOGIANNAKOS, E. K., TSEKOURA, P., KIDONOPOULOU, K., TRAKATELLI, C., POLIMENIDIS, Z., CONTI, P. und ANOGIANAKIS, G.: Interleukin -1beta (IL-1 beta), interleukin 6 (IL-6) and tumor necrosis factor (TNF) in plasma and pleural fluid of pneumonia, lung cancer and tuberculous pleuritis. *J Biol Regul Homeost Agents* (2006), **20** (1-2): 41-6

INAZUKA, M., TAHIRA, T., HORIUCHI, T., HARASHIMA, S., SAWABE, T., KONDO, M., MIYAHARA, H. und HAYASHI, K.: Analysis of p53 tumour suppressor gene somatic mutations in rheumatoid arthritis synovium. *Rheumatology (Oxford)* (2000), **39** (3): 262-6

INFORMA UK LTD, (2008)

INGENUITY SYSTEMS INC., Homepage, verfügbar unter http://www.ingenuity.com/ (2008)

INOHA, S., INAMURA, T., IKEZAKI, K., NAKAMIZO, A., AMANO, T. und FUKUI, M.: Type V phosphodiesterase expression in cerebral arteries with vasospasm after subarachnoid hemorrhage in a canine model. *Neurol Res* (2002), **24** (6): 607-12

ITOH, K., HASE, H., KOJIMA, H., SAOTOME, K., NISHIOKA, K. und KOBATA, T.: Central role of mitochondria and p53 in Fas-mediated apoptosis of rheumatoid synovial fibroblasts. *Rheumatology (Oxford)* (2004), **43** (3): 277-85

JAGER, R. D., MIELER, W. F. und MILLER, J. W.: Age-related macular degeneration. *N Engl J Med* (2008), **358** (24): 2606-17

JONES, D.: Pathways to cancer therapy. *Nat Rev Drug Discov* (2008), **7** (11): 875-6

JONES, P. F., JAKUBOWICZ, T., PITOSSI, F. J., MAURER, F. und HEMMINGS, B. A., (1991), Vol. 88, pp. 4171-4175

JOSHI-TOPE, G., GILLESPIE, M., VASTRIK, I., D'EUSTACHIO, P., SCHMIDT, E., DE BONO, B., JASSAL, B., GOPINATH, G. R., WU, G. R., MATTHEWS, L., LEWIS, S., BIRNEY, E. und STEIN, L.: Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* (2005), **33** (Database issue): D428-32

JUBILANT BIOSYS, Homepage, verfügbar unter http://www.jubilantbiosys.com/pathart.html (2008)

KABBINAVAR, F., HURWITZ, H. I., FEHRENBACHER, L., MEROPOL, N. J., NOVOTNY, W. F., LIEBERMAN, G., GRIFFING, S. und BERGSLAND, E.: Phase II, randomized trial comparing bevacizumab plus fluorouracil (FU)/leucovorin (LV) with FU/LV alone in patients with metastatic colorectal cancer. *J Clin Oncol* (2003), **21** (1): 60-5

KAISER, P. K., BROWN, D. M., ZHANG, K., HUDSON, H. L., HOLZ, F. G., SHAPIRO, H., SCHNEIDER, S. und ACHARYA, N. R.: Ranibizumab for predominantly classic neovascular age-related macular degeneration: subgroup analysis of first-year ANCHOR results. *Am J Ophthalmol* (2007), **144** (6): 850-857

KAMATA, Y., KAMIMURA, T., IWAMOTO, M. und MINOTA, S.: Comparable effects of sildenafil citrate and alprostadil on severe Raynaud's phenomenon in a patient with systemic sclerosis. *Clin Exp Dermatol* (2005), **30** (4): 451

KANEHISA LABORATORIES, KEGG: Kyoto Encyclopedia of Genes and Genomes, verfügbar unter http://www.genome.jp/kegg/ (2008)

KANEHISA, M. und GOTO, S.: KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* (2000), **28** (1): 27-30

KANEHISA, M., GOTO, S., HATTORI, M., AOKI-KINOSHITA, K. F., ITOH, M., KAWASHIMA, S., KATAYAMA, T., ARAKI, M. und HIRAKAWA, M.: From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* (2006), **34** (Database issue): D354-7

KANEHISA, M., ARAKI, M., GOTO, S., HATTORI, M., HIRAKAWA, M., ITOH, M., KATAYAMA, T., KAWASHIMA, S., OKUDA, S., TOKIMATSU, T. und YAMANISHI, Y.: KEGG for linking genomes to life and the environment. *Nucleic Acids Res* (2008), **36** (Database issue): D480-4

KATZ, S. D., BALIDEMAJ, K., HOMMA, S., WU, H., WANG, J. und MAYBAUM, S.: Acute type 5 phosphodiesterase inhibition with sildenafil enhances flow-mediated vasodilation in patients with chronic heart failure. *J Am Coll Cardiol* (2000), **36** (3): 845-51

KIM, G., JUN, J. B. und ELKON, K. B.: Necessary role of phosphatidylinositol 3-kinase in transforming growth factor beta-mediated activation of Akt in normal and rheumatoid arthritis synovial fibroblasts. *Arthritis Rheum* (2002), **46** (6): 1504-11

KIM, K. J., LI, B., WINER, J., ARMANINI, M., GILLETT, N., PHILLIPS, H. S. und FERRARA, N.: Inhibition of vascular endothelial growth factor-induced angiogenesis suppresses tumour growth in vivo. *Nature* (1993), **362** (6423): 841-4

KIM, K. W., CHO, M. L., PARK, M. K., YOON, C. H., PARK, S. H., LEE, S. H. und KIM, H. Y.: Increased interleukin-17 production via a phosphoinositide 3-kinase/Akt and nuclear factor kappaB-dependent pathway in patients with rheumatoid arthritis. *Arthritis Res Ther* (2005), **7** (1): R139-48

KIMBERLEY, F. C. und SCREATON, G. R.: Following a TRAIL: Update on a ligand and its five receptors. *Cell Res* (2004), **14** (5): 359-372

KITANO, H.: A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* (2007), **6** (3): 202-10

KLING, J.: From hypertension to angina to Viagra. *Modern Drug Discovery* (1998), **1** (2): 31-38

KLINGELHOFER, J., SENOLT, L., BASLUND, B., NIELSEN, G. H., SKIBSHOJ, I., PAVELKA, K., NEIDHART, M., GAY, S., AMBARTSUMIAN, N., HANSEN, B. S., PETERSEN, J., LUKANIDIN, E. und GRIGORIAN, M.: Up-regulation of metastasis-promoting S100A4 (Mts-1) in rheumatoid arthritis: putative involvement in the pathogenesis of rheumatoid arthritis. *Arthritis Rheum* (2007), **56** (3): 779-89

KLINIKUM BREMEN-MITTE INSTITUT FÜR PHARMAKOLOGIE, VIBERA Studie - Prävention des Sehverlustes bei Patienten mit altersabhängiger Makuladegeneration

durch intravitreale Injektion von Bevacizumab und Ranibizumab, verfügbar unter http://www.pharmakologie-bremen.de/index.php/klinische-forschung/vibera (2008)

KNOWLES, J. und GROMO, G.: A guide to drug discovery: Target selection in drug discovery. *Nat Rev Drug Discov* (2003), **2** (1): 63-9

KOTHARI, S. S. und DUGGAL, B.: Chronic oral sildenafil therapy in severe pulmonary artery hypertension. *Indian Heart J* (2002), **54** (4): 404-9

KRAMER, R. und COHEN, D.: Functional genomics to new drug targets. *Nat Rev Drug Discov* (2004), **3** (11): 965-72

KUHRT, N., in *Die Zeit* (2008), Vol. 20

LAMB, P. und CRAWFORD, L.: Characterization of the human p53 gene. *Mol Cell Biol* (1986), **6** (5): 1379-85

LANE, D. P.: Cancer. p53, guardian of the genome. *Nature* (1992), **358** (6381): 15-6

LAVI, G., VORONOV, E., DINARELLO, C. A., APTE, R. N. und COHEN, S.: Sustained delivery of IL-1Ra from biodegradable microspheres reduces the number of murine B16 melanoma lung metastases. *Journal of Controlled Release* (2007), **123** (2): 123-130

LEE, Y. H., KIM, Y. R., JI, J. D., SOHN, J. und SONG, G. G.: p53 codon 72 polymorphism and rheumatoid arthritis. *J Rheumatol* (2001), **28** (11): 2392-4

LEECH, M., LACEY, D., XUE, J. R., SANTOS, L., HUTCHINSON, P., WOLVETANG, E., DAVID, J. R., BUCALA, R. und MORAND, E. F.: Regulation of p53 by macrophage migration inhibitory factor in inflammatory arthritis. *Arthritis Rheum* (2003), **48** (7): 1881-9

LEECH, M., XUE, J. R., DACUMOS, A., HALL, P., SANTOS, L., YANG, Y., LI, M., KITCHING, A. R. und MORAND, E. F.: The tumour suppressor gene p53 modulates the severity of antigen-induced arthritis and the systemic immune response. *Clin Exp Immunol* (2008), **152** (2): 345-53

LEWIS, A., VARGHESE, S., XU, H. und ALEXANDER, H. R., (2006), Vol. 4, pp. 48

LICHTENSTEIN, J. R.: Use of sildenafil citrate in Raynaud's phenomenon: comment on the article by Thompson et al. *Arthritis Rheum* (2003), **48** (1): 282-3; author reply 283

LIND, H., ZIENOLDDINY, S., RYBERG, D., SKAUG, V., PHILLIPS, D. H. und HAUGEN, A.: Interleukin 1 receptor antagonist gene polymorphism and risk of lung cancer: A possible interaction with polymorphisms in the interleukin 1 beta gene. *Lung Cancer* (2005), **50** (3): 285-290

LINDER, A. E., NI, W., DIAZ, J. L., SZASZ, T., BURNETT, R. und WATTS, S. W.: Serotonin (5-HT) in veins: not all in vain. *J Pharmacol Exp Ther* (2007), **323** (2): 415-21

LIU, H. und POPE, R. M.: The role of apoptosis in rheumatoid arthritis. *Curr Opin Pharmacol* (2003), **3** (3): 317-22

LIU, Z., XU, X., HSU, H. C., TOUSSON, A., YANG, P. A., WU, Q., LIU, C., YU, S., ZHANG, H. G. und MOUNTZ, J. D.: CII-DC-AdTRAIL cell gene therapy inhibits infiltration of CII-reactive T cells and CII-induced arthritis. *J Clin Invest* (2003), **112** (9): 1332-41

LUDIKHUIZE, J., DE LAUNAY, D., GROOT, D., SMEETS, T. J., VINKENOOG, M., SANDERS, M. E., TAS, S. W., TAK, P. P. und REEDQUIST, K. A.: Inhibition of forkhead box class O family member transcription factors in rheumatoid synovial tissue. *Arthritis Rheum* (2007), **56** (7): 2180-91

LYNCH, T. J., BELL, D. W., SORDELLA, R., GURUBHAGAVATULA, S., OKIMOTO, R. A., BRANNIGAN, B. W., HARRIS, P. L., HASERLAT, S. M., SUPKO, J. G., HALUSKA, F. G., LOUIS, D. N., CHRISTIANI, D. C., SETTLEMAN, J. und HABER, D. A.: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* (2004), **350** (21): 2129-39

MAAS, K., WESTFALL, M., PIETENPOL, J., OLSEN, N. J. und AUNE, T.: Reduced p53 in peripheral blood mononuclear cells from patients with rheumatoid arthritis is associated with loss of radiation-induced apoptosis. *Arthritis Rheum* (2005), **52** (4): 1047-57

MACCHIONI, P., NICOLI, D., CASALI, B., CATANOSO, M., FARNETTI, E., BOIARDI, L. und SALVARANI, C.: The codon 72 polymorphic variants of p53 in Italian rheumatoid arthritis patients. *Clin Exp Rheumatol* (2007), **25** (3): 416-21

MAGLOTT, D., OSTELL, J., PRUITT, K. D. und TATUSOVA, T.: Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* (2005), **33** (Database issue): D54-8

MAGLOTT, D. R., KATZ, K. S., SICOTTE, H. und PRUITT, K. D.: NCBI's LocusLink and RefSeq. *Nucleic Acids Res* (2000), **28** (1): 126-8

MALECAZE, F., CLAMENS, S., SIMORRE-PINATEL, V., MATHIS, A., CHOLLET, P., FAVARD, C., BAYARD, F. und PLOUET, J.: Detection of vascular endothelial growth factor messenger RNA and vascular endothelial growth factor-like activity in proliferative diabetic retinopathy. *Arch Ophthalmol* (1994), **112** (11): 1476-82

MANNING, B. D. und CANTLEY, L. C.: AKT/PKB signaling: navigating downstream. *Cell* (2007), **129** (7): 1261-74

MARGOLIN, K., GORDON, M. S., HOLMGREN, E., GAUDREAULT, J., NOVOTNY, W., FYFE, G., ADELMAN, D., STALTER, S. und BREED, J.: Phase Ib trial of intravenous recombinant humanized monoclonal antibody to vascular endothelial growth factor in combination with chemotherapy in patients with advanced cancer: pharmacologic and long-term safety data. *J Clin Oncol* (2001), **19** (3): 851-6

MARONE, R., CMILJANOVIC, V., GIESE, B. und WYMANN, M. P.: Targeting phosphoinositide 3-kinase: moving towards therapy. *Biochim Biophys Acta* (2008), **1784** (1): 159-85

MAURICE, D. H., PALMER, D., TILLEY, D. G., DUNKERLEY, H. A., NETHERTON, S. J., RAYMOND, D. R., ELBATARNY, H. S. und JIMMO, S. L.: Cyclic

nucleotide phosphodiesterase activity, expression, and targeting in cells of the cardiovascular system. *Mol Pharmacol* (2003), **64** (3): 533-46

MCVARY, K. T.: Clinical practice. Erectile dysfunction. *N Engl J Med* (2007), **357** (24): 2472-81

MICHAEL, V. V. und ALISA, K. E.: Cell cycle implications in the pathogenesis of rheumatoid arthritis. *Front Biosci* (2000), **5** D594-601

MIGITA, K., HONDA, S., YAMASAKI, S., HIRAI, Y., FUKUDA, T., AOYAGI, T., KITA, M., IDA, H., TSUKADA, T., KAWAKAMI, A., KAWABE, Y. und EGUCHI, K.: Regulation of rheumatoid synovial cell growth by ceramide. *Biochem Biophys Res Commun* (2000), **269** (1): 70-5

MIGITA, K., TANAKA, F., YAMASAKI, S., SHIBATOMI, K., IDA, H., KAWAKAMI, A., AOYAGI, T., KAWABE, Y. und EGUCHI, K.: Regulation of rheumatoid synoviocyte proliferation by endogenous p53 induction. *Clin Exp Immunol* (2001), **126** (2): 334-8

MILLER, J. W., ADAMIS, A. P., SHIMA, D. T., D'AMORE, P. A., MOULTON, R. S., O'REILLY, M. S., FOLKMAN, J., DVORAK, H. F., BROWN, L. F., BERSE, B. und ET AL.: Vascular endothelial growth factor/vascular permeability factor is temporally and spatially correlated with ocular angiogenesis in a primate model. *Am J Pathol* (1994), **145** (3): 574-84

MIRANDA-CARUS, M. E., BALSA, A., BENITO-MIGUEL, M., DE AYALA, C. P. und MARTIN-MOLA, E.: Rheumatoid arthritis synovial fluid fibroblasts express TRAIL-R2 (DR5) that is functionally active. *Arthritis Rheum* (2004), **50** (9): 2786-93

MIYASHITA, T., KAWAKAMI, A., TAMAI, M., IZUMI, Y., MINGGUO, H., TANAKA, F., ABIRU, S., NAKASHIMA, K., IWANAGA, N., ARATAKE, K., KAMACHI, M., ARIMA, K., IDA, H., MIGITA, K., ORIGUCHI, T., TAGASHIRA, S., NISHIKAKU, F. und EGUCHI, K.: Akt is an endogenous inhibitor toward tumor necrosis factor-related apoptosis inducing ligand-mediated apoptosis in rheumatoid synovial cells. *Biochem Biophys Res Commun* (2003), **312** (2): 397-404

MIYASHITA, T., KAWAKAMI, A., NAKASHIMA, T., YAMASAKI, S., TAMAI, M., TANAKA, F., KAMACHI, M., IDA, H., MIGITA, K., ORIGUCHI, T., NAKAO, K. und EGUCHI, K.: Osteoprotegerin (OPG) acts as an endogenous decoy receptor in tumour necrosis factor-related apoptosis-inducing ligand (TRAIL)-mediated apoptosis of fibroblast-like synovial cells. *Clin Exp Immunol* (2004), **137** (2): 430-6

MONTORSI, F., MCDERMOTT, T. E., MORGAN, R., OLSSON, A., SCHULTZ, A., KIRKEBY, H. J. und OSTERLOH, I. H.: Efficacy and safety of fixed-dose oral sildenafil in the treatment of erectile dysfunction of various etiologies. *Urology* (1999), **53** (5): 1011-8

MOREL, J., AUDO, R., HAHNE, M. und COMBE, B.: Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) induces rheumatoid arthritis synovial fibroblast proliferation through mitogen-activated protein kinases and phosphatidylinositol 3-kinase/Akt. *J Biol Chem* (2005), **280** (16): 15709-18

NAEIJE, R.: Pulmonary hypertension and right heart failure in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* (2005), **2** (1): 20-2

NAKAJIMA, T., ARATANI, S., NAKAZAWA, M., HIROSE, T., FUJITA, H. und NISHIOKA, K.: Implications of transcriptional coactivator CREB binding protein complexes in rheumatoid arthritis. *Mod Rheumatol* (2004), **14** (1): 6-11

NAKAO, S., KUWANO, T., TSUTSUMI-MIYAHARA, C., UEDA, S., KIMURA, Y. N., HAMANO, S., SONODA, K. H., SAIJO, Y., NUKIWA, T., STRIETER, R. M., ISHIBASHI, T., KUWANO, M. und ONO, M.: Infiltration of COX-2-expressing macrophages is a prerequisite for IL-1 beta-induced neovascularization and tumor growth. *J Clin Invest* (2005), **115** (11): 2979-91

NAKAZAWA, M., ARATANI, S., HATTA, M., ARAYA, N., DAITOKU, H., KAWAHARA, K., WATANABE, S., NAKAMURA, H., YOSHINO, S., FUJII, R., FUJITA, H., FUKAMIZU, A., NISHIOKA, K. und NAKAJIMA, T.: TNFalpha induces acetylation of p53 but attenuates its transcriptional activation in rheumatoid synoviocytes. *Int J Mol Med* (2002), **10** (3): 269-75

NATIONAL EYE INSTITUTE, NEI Press release - Comparison of AMD Treatments Trials (CATT): Lucentis - Avastin Trial, verfügbar unter http://www.nei.nih.gov/news/pressreleases/022208.asp (2008)

NICHOLSON, J. K. und WILSON, I. D.: Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov* (2003), **2** (8): 668-76

NIKOLSKY, Y., NIKOLSKAYA, T. und BUGRIM, A.: Biological networks and analysis of experimental data in drug discovery. *Drug Discov Today* (2005), **10** (9): 653-62

NOVARTIS INTERNATIONAL AG, Lucentis Homepage Deutschland, verfügbar unter http://www.lucentis.de (2008)

OMORI, K. und KOTERA, J.: Overview of PDEs and their regulation. *Circ Res* (2007), **100** (3): 309-27

OVERINGTON, J. P., AL-LAZIKANI, B. und HOPKINS, A. L.: How many drug targets are there? *Nat Rev Drug Discov* (2006), **5** (12): 993-6

PAP, T., AUPPERLE, K. R., GAY, S., FIRESTEIN, G. S. und GAY, R. E.: Invasiveness of synovial fibroblasts is regulated by p53 in the SCID mouse in vivo model of cartilage invasion. *Arthritis Rheum* (2001), **44** (3): 676-81

PARK, Y. W., JI, J. D., LEE, J. S., RYANG, D. W. und YOO, D. H.: Actinomycin D renders cultured synovial fibroblasts susceptible to tumour necrosis factor related apoptosis-inducing ligand (TRAIL)-induced apoptosis. *Scand J Rheumatol* (2003), **32** (6): 356-63

PATEL, M. D. und KATZ, S. D.: Phosphodiesterase 5 inhibition in chronic heart failure and pulmonary hypertension. *Am J Cardiol* (2005), **96** (12B): 47M-51M

PERI, S., NAVARRO, J. D., KRISTIANSEN, T. Z., AMANCHY, R., SURENDRANATH, V., MUTHUSAMY, B., GANDHI, T. K., CHANDRIKA, K. N., DESHPANDE, N., SURESH, S., RASHMI, B. P., SHANKER, K., PADMA, N., NIRANJAN, V.,

HARSHA, H. C., TALREJA, N., VRUSHABENDRA, B. M., RAMYA, M. A., YATISH, A. J., JOY, M., SHIVASHANKAR, H. N., KAVITHA, M. P., MENEZES, M., CHOUDHURY, D. R., GHOSH, N., SARAVANA, R., CHANDRAN, S., MOHAN, S., JONNALAGADDA, C. K., PRASAD, C. K., KUMAR-SINHA, C., DESHPANDE, K. S. und PANDEY, A.: Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Res* (2004), **32** (Database issue): D497-501

PERIMENIS, P.: Sildenafil for the treatment of altitude-induced hypoxaemia. *Expert Opin Pharmacother* (2005), **6** (5): 835-7

PERLMAN, H., NGUYEN, N., LIU, H., ESLICK, J., ESSER, S., WALSH, K., MOORE, T. L. und POPE, R. M.: Rheumatoid arthritis synovial fluid macrophages express decreased tumor necrosis factor-related apoptosis-inducing ligand R2 and increased decoy receptor tumor necrosis factor-related apoptosis-inducing ligand R3. *Arthritis Rheum* (2003), **48** (11): 3096-101

PFIZER INC.: 2002 Annual Report (2003)

PFIZER INC.: 2003 Annual Review, Financial Review (2004)

PFIZER INC.: 2007 Financial Report (2008a)

PFIZER INC., Revatio Homepage, verfügbar unter http://www.revatio.com/ (2008b)

PFIZER PHARMA GMBH, Mann-Info, Fachinformationen zu Viagra, verfügbar unter http://www.mann-info.de/fach/wissen_viagra/index.htm (2008a)

PFIZER PHARMA GMBH, Herz-Info, Fachinformationen zu Revatio, verfügbar unter http://www.herz-info.de/fach/produkte/revatio/index.htm (2008b)

PRESTA, L. G., CHEN, H., O'CONNOR, S. J., CHISHOLM, V., MENG, Y. G., KRUMMEN, L., WINKLER, M. und FERRARA, N.: Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders. *Cancer Res* (1997), **57** (20): 4593-9

PRIESTLE, J. P., SCHAR, H. P. und GRUTTER, M. G.: Crystal structure of the cytokine interleukin-1 beta. *Embo J* (1988), **7** (2): 339-43

PRIESTLE, J. P., SCHAR, H. P. und GRUTTER, M. G.: Crystallographic refinement of interleukin 1 beta at 2.0 A resolution. *Proc Natl Acad Sci U S A* (1989), **86** (24): 9667-71

PROTEIN LOUNGE, Homepage, verfügbar unter www.proteinlounge.com (2008)

RICHALET, J. P., GRATADOUR, P., ROBACH, P., PHAM, I., DECHAUX, M., JONCQUIERT-LATARJET, A., MOLLARD, P., BRUGNIAUX, J. und CORNOLO, J.: Sildenafil inhibits altitude-induced hypoxemia and pulmonary hypertension. *Am J Respir Crit Care Med* (2005), **171** (3): 275-81

ROSENFELD, P. J., SCHWARTZ, S. D., BLUMENKRANZ, M. S., MILLER, J. W., HALLER, J. A., REIMANN, J. D., GREENE, W. L. und SHAMS, N.: Maximum tolerated dose of a humanized anti-vascular endothelial growth factor antibody fragment for treating neovascular age-related macular degeneration. *Ophthalmology* (2005), **112** (6): 1048-53

ROSENFELD, P. J., BROWN, D. M., HEIER, J. S., BOYER, D. S., KAISER, P. K., CHUNG, C. Y. und KIM, R. Y.: Ranibizumab for neovascular age-related macular degeneration. *N Engl J Med* (2006a), **355** (14): 1419-31

ROSENFELD, P. J., HEIER, J. S., HANTSBARGER, G. und SHAMS, N.: Tolerability and efficacy of multiple escalating doses of ranibizumab (Lucentis) for neovascular age-related macular degeneration. *Ophthalmology* (2006b), **113** (4): 623 e1

ROSENGARTEN, B., SCHERMULY, R. T., VOSWINCKEL, R., KOHSTALL, M. G., OLSCHEWSKI, H., WEISSMANN, N., SEEGER, W., KAPS, M., GRIMMINGER, F. und GHOFRANI, H. A.: Sildenafil improves dynamic vascular function in the brain: studies in patients with pulmonary hypertension. *Cerebrovasc Dis* (2006), **21** (3): 194-200

ROTH, J. A. und GRAMMER, S. F.: Gene replacement therapy for non-small cell lung cancer: a review. *Hematol Oncol Clin North Am* (2004), **18** (1): 215-29

RÜCKER, D., in *Pharmazeutische Zeitung online* (2007), Vol. 45/2007

RUSS, A. P. und LAMPEL, S.: The druggable genome: an update. *Drug Discov Today* (2005), **10** (23-24): 1607-10

SALVADOR, G., SANMARTI, R., GARCIA-PEIRO, A., RODRIGUEZ-CROS, J. R., MUNOZ-GOMEZ, J. und CANETE, J. D.: p53 expression in rheumatoid and psoriatic arthritis synovial tissue and association with joint damage. *Ann Rheum Dis* (2005), **64** (2): 183-7

SALWINSKI, L., MILLER, C. S., SMITH, A. J., PETTIT, F. K., BOWIE, J. U. und EISENBERG, D.: The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res* (2004), **32** (Database issue): D449-51

SAMS-DODD, F.: Drug discovery: selecting the optimal approach. *Drug Discov Today* (2006), **11** (9-10): 465-72

SASTRY, B. K., NARASIMHAN, C., REDDY, N. K., ANAND, B., PRAKASH, G. S., RAJU, P. R. und KUMAR, D. N.: A study of clinical efficacy of sildenafil in patients with primary pulmonary hypertension. *Indian Heart J* (2002), **54** (4): 410-4

SASTRY, B. K., NARASIMHAN, C., REDDY, N. K. und RAJU, B. S.: Clinical efficacy of sildenafil in primary pulmonary hypertension: a randomized, placebo-controlled, double-blind, crossover study. *J Am Coll Cardiol* (2004), **43** (7): 1149-53

SCHNEIDER, P., BODMER, J. L., THOME, M., HOFMANN, K., HOLLER, N. und TSCHOPP, J.: Characterization of two receptors for TRAIL. *FEBS Lett* (1997a), **416** (3): 329-34

SCHNEIDER, P., THOME, M., BURNS, K., BODMER, J.-L., HOFMANN, K., KATAOKA, T., HOLLER, N. und TSCHOPP, J.: TRAIL Receptors 1 (DR4) and 2 (DR5) Signal FADD-Dependent Apoptosis and Activate NF-[kappa]B. *Immunity* (1997b), **7** (6): 831-836

SCHRATTENHOLZ, A. und SOSKIC, V.: What does systems biology mean for drug development? *Curr Med Chem* (2008), **15** (15): 1520-8

SEEMAYER, C. A., KUCHEN, S., NEIDHART, M., KUENZLER, P., RIHOSKOVA, V., NEUMANN, E., PRUSCHY, M., AICHER, W. K., MULLER-LADNER, U., GAY, R. E., MICHEL, B. A., FIRESTEIN, G. S. und GAY, S.: p53 in rheumatoid

arthritis synovial fibroblasts at sites of invasion. *Ann Rheum Dis* (2003), **62** (12): 1139-44

SEIDLITZ, F., in *Welt Online* (2008)

SENGER, D. R., GALLI, S. J., DVORAK, A. M., PERRUZZI, C. A., HARVEY, V. S. und DVORAK, H. F.: Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science* (1983), **219** (4587): 983-5

SHEREMATA, W. A., JY, W., HORSTMAN, L. L., AHN, Y. S., ALEXANDER, J. S. und MINAGAR, A.: Evidence of platelet activation in multiple sclerosis. *J Neuroinflammation* (2008), **5** 27

SIMELYTE, E., ROSENGREN, S., BOYLE, D. L., CORR, M., GREEN, D. R. und FIRESTEIN, G. S.: Regulation of arthritis by p53: critical role of adaptive immunity. *Arthritis Rheum* (2005), **52** (6): 1876-84

SIMS, J. E., ACRES, R. B., GRUBIN, C. E., MCMAHAN, C. J., WIGNALL, J. M., MARCH, C. J. und DOWER, S. K., (1989), Vol. 86, pp. 8946-8950

SMITH, J. M. und SMITH, D. C. P.: Database abstractions: aggregation and generalization. (1977), **2** (2): 105-133

SONG, K., CHEN, Y., GOKE, R., WILMEN, A., SEIDEL, C., GOKE, A., HILLIARD, B. und CHEN, Y.: Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) is an inhibitor of autoimmune inflammation and cell cycle progression. *J Exp Med* (2000), **191** (7): 1095-104

SONG, X., VORONOV, E., DVORKIN, T., FIMA, E., CAGNANO, E., BENHARROCH, D., SHENDLER, Y., BJORKDAHL, O., SEGAL, S., DINARELLO, C. A. und APTE, R. N., (2003), Vol. 171, pp. 6448-6456

STOREY, V. C.: Comparing Relationships in Conceptual Modeling: Mapping to Semantic Classifications. (2005), **17** (11): 1478-1489

SUN, Y. und CHEUNG, H. S.: p53, proto-oncogene and rheumatoid arthritis. *Semin Arthritis Rheum* (2002), **31** (5): 299-310

TAK, P. P., ZVAIFLER, N. J., GREEN, D. R. und FIRESTEIN, G. S.: Rheumatoid arthritis and p53: how oxidative stress might alter the course of inflammatory diseases. *Immunol Today* (2000), **21** (2): 78-82

TAKIMOTO, E., CHAMPION, H. C., LI, M., BELARDI, D., REN, S., RODRIGUEZ, E. R., BEDJA, D., GABRIELSON, K. L., WANG, Y. und KASS, D. A.: Chronic inhibition of cyclic GMP phosphodiesterase 5A prevents and reverses cardiac hypertrophy. *Nat Med* (2005), **11** (2): 214-22

TARANTO, E., XUE, J. R., LACEY, D., HUTCHINSON, P., SMITH, M., MORAND, E. F. und LEECH, M.: Detection of the p53 regulator murine double-minute protein 2 in rheumatoid arthritis. *J Rheumatol* (2005), **32** (3): 424-9

TERZIOGLU, E., BISGIN, A., SANLIOGLU, A. D., ULKER, M., YAZISIZ, V., TUZUNER, S. und SANLIOGLU, S.: Concurrent gene therapy strategies effectively destroy synoviocytes of patients with rheumatoid arthritis. *Rheumatology (Oxford)* (2007), **46** (5): 783-9

THE UNIPROT, C.: The universal protein resource (UniProt). *Nucleic Acids Res* (2008), **36** (Database issue): D190-5

THUSU, K. G., MORIN, F. C., 3RD, RUSSELL, J. A. und STEINHORN, R. H.: The cGMP phosphodiesterase inhibitor zaprinast enhances the effect of nitric oxide. *Am J Respir Crit Care Med* (1995), **152** (5 Pt 1): 1605-10

TSOKOS, G. C. und TSOKOS, M.: The TRAIL to arthritis. *J Clin Invest* (2003), **112** (9): 1315-7

U.S. NATIONAL LIBRARY OF MEDICINE: Medical Subject Headings (MeSH) Browser. (2008a),

U.S. NATIONAL LIBRARY OF MEDICINE, Entrez Pubmed, verfügbar unter http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed (2008b)

UEMURA, Y., NAKATA, H., KOBAYASHI, M., HARADA, R., ASAHI, Y. und TAGUCHI, H.: Regulation of granulocyte colony-stimulating factor and parathyroid hormone-related protein production in lung carcinoma cell line OKa-C-1. *Jpn J Cancer Res* (2000), **91** (9): 911-7

VIGERS, G. P., ANDERSON, L. J., CAFFES, P. und BRANDHUBER, B. J.: Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1beta. *Nature* (1997), **386** (6621): 190-4

VIGERS, G. P., DRIPPS, D. J., EDWARDS, C. K., 3RD und BRANDHUBER, B. J.: X-ray crystal structure of a small antagonist peptide bound to interleukin-1 receptor type 1. *J Biol Chem* (2000), **275** (47): 36927-33

VOELKEL, N. F. und COOL, C. D.: Pulmonary vascular involvement in chronic obstructive pulmonary disease. *Eur Respir J Suppl* (2003), **46** 28s-32s

VORONOV, E., SHOUVAL, D. S., KRELIN, Y., CAGNANO, E., BENHARROCH, D., IWAKURA, Y., DINARELLO, C. A. und APTE, R. N., (2003), Vol. 100, pp. 2645-2650

WALKER, D. K., ACKLAND, M. J., JAMES, G. C., MUIRHEAD, G. J., RANCE, D. J., WASTALL, P. und WRIGHT, P. A.: Pharmacokinetics and metabolism of sildenafil in mouse, rat, rabbit, dog and man. *Xenobiotica* (1999), **29** (3): 297-310

WANG, L., ZHANG, B., WOLFINGER, R. D. und CHEN, X.: An integrated approach for the analysis of biological pathways using mixed models. *PLoS Genet* (2008), **4** (7): e1000115

WANG, S., SIM, T. B., KIM, Y. S. und CHANG, Y. T.: Tools for target identification and validation. *Curr Opin Chem Biol* (2004), **8** (4): 371-7

WATTERSON, S., MARSHALL, S. und GHAZAL, P.: Logic models of pathway biology. *Drug Discov Today* (2008), **13** (9-10): 447-56

WEBB, D. J., FREESTONE, S., ALLEN, M. J. und MUIRHEAD, G. J.: Sildenafil citrate and blood-pressure-lowering drugs: results of drug interaction studies with an organic nitrate and a calcium antagonist. *Am J Cardiol* (1999), **83** (5A): 21C-28C

WEINREICH, D. M., ELARAJ, D. M., PUHLMANN, M., HEWITT, S. M., CARROLL, N. M., FELDMAN, E. D., TURNER, E. M., SPIESS, P. J. und ALEXANDER, H. R.: Effect of interleukin 1 receptor antagonist gene transduction on human melanoma xenografts in nude mice. *Cancer Res* (2003), **63** (18): 5957-61

WEISS, B. und HAIT, W. N.: Selective cyclic nucleotide phosphodiesterase inhibitors as potential therapeutic agents. *Annu Rev Pharmacol Toxicol* (1977), **17** 441-77

WHO COLLABORATING CENTRE FOR DRUG STATISTICS METHODOLOGY, Homepage, verfügbar unter http://www.whocc.no/atcddd/ (2008)

WILKINS, M. R., PAUL, G. A., STRANGE, J. W., TUNARIU, N., GIN-SING, W., BANYA, W. A., WESTWOOD, M. A., STEFANIDIS, A., NG, L. L., PENNELL, D. J., MOHIADDIN, R. H., NIHOYANNOPOULOS, P. und GIBBS, J. S.: Sildenafil versus Endothelin Receptor Antagonist for Pulmonary Hypertension (SERAPH) study. *Am J Respir Crit Care Med* (2005), **171** (11): 1292-7

WILSON, D. R.: Viral-mediated gene transfer for cancer treatment. *Curr Pharm Biotechnol* (2002), **3** (2): 151-64

WU, L., MARTINEZ-CASTELLANOS, M. A., QUIROZ-MERCADO, H., AREVALO, J. F., BERROCAL, M. H., FARAH, M. E., MAIA, M., ROCA, J. A. und RODRIGUEZ, F. J.: Twelve-month safety of intravitreal injections of bevacizumab (Avastin(R)): results of the Pan-American Collaborative Retina Study Group (PACORES). *Graefes Arch Clin Exp Ophthalmol* (2008), **246** (1): 81-7

XU, H., XU, H., LIN, M., WANG, W., LI, Z., HUANG, J., CHEN, Y. und CHEN, X.: Learning the drug target-likeness of a protein. *Proteomics* (2007), **7** (23): 4255-63

YAMANISHI, Y. und FIRESTEIN, G. S.: Pathogenesis of rheumatoid arthritis: the role of synoviocytes. *Rheum Dis Clin North Am* (2001), **27** (2): 355-71

YAMANISHI, Y., BOYLE, D. L., PINKOSKI, M. J., MAHBOUBI, A., LIN, T., HAN, Z., ZVAIFLER, N. J., GREEN, D. R. und FIRESTEIN, G. S.: Regulation of joint destruction and inflammation by p53 in collagen-induced arthritis. *Am J Pathol* (2002a), **160** (1): 123-30

YAMANISHI, Y., BOYLE, D. L., ROSENGREN, S., GREEN, D. R., ZVAIFLER, N. J. und FIRESTEIN, G. S.: Regional analysis of p53 mutations in rheumatoid arthritis synovium. *Proc Natl Acad Sci U S A* (2002b), **99** (15): 10025-30

YAMANISHI, Y., BOYLE, D. L., GREEN, D. R., KEYSTONE, E. C., CONNOR, A., ZOLLMAN, S. und FIRESTEIN, G. S.: p53 tumor suppressor gene mutations in fibroblast-like synoviocytes from erosion synovium and non-erosion synovium in rheumatoid arthritis. *Arthritis Res Ther* (2005), **7** (1): R12-8

YAMASAKI, S., YAGISHITA, N., NISHIOKA, K. und NAKAJIMA, T.: The roles of synoviolin in crosstalk between endoplasmic reticulum stress-induced apoptosis and p53 pathway. *Cell Cycle* (2007), **6** (11): 1319-23

YANG, J. C., HAWORTH, L., SHERRY, R. M., HWU, P., SCHWARTZENTRUBER, D. J., TOPALIAN, S. L., STEINBERG, S. M., CHEN, H. X. und ROSENBERG, S. A.: A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer. *N Engl J Med* (2003), **349** (5): 427-34

YANG, X., WANG, J., LIU, C., GRIZZLE, W. E., YU, S., ZHANG, S., BARNES, S., KOOPMAN, W. J., MOUNTZ, J. D., KIMBERLY, R. P. und ZHANG, H. G.: Cleavage of p53-vimentin complex enhances tumor necrosis factor-related

apoptosis-inducing ligand-mediated apoptosis of rheumatoid arthritis synovial fibroblasts. *Am J Pathol* (2005), **167** (3): 705-19

YANG, Z. Z., TSCHOPP, O., BAUDRY, A., DUMMLER, B., HYNX, D. und HEMMINGS, B. A.: Physiological functions of protein kinase B/Akt. *Biochem Soc Trans* (2004), **32** (Pt 2): 350-4

YANO, S., NOKIHARA, H., YAMAMOTO, A., GOTO, H., OGAWA, H., KANEMATSU, T., MIKI, T., UEHARA, H., SAIJO, Y., NUKIWA, T. und SONE, S.: Multifunctional interleukin-1beta promotes metastasis of human lung cancer cells in SCID mice via enhanced expression of adhesion-, invasion- and angiogenesis-related molecules. *Cancer Sci* (2003), **94** (3): 244-52

YAO, Q., WANG, S., GLORIOSO, J. C., EVANS, C. H., ROBBINS, P. D., GHIVIZZANI, S. C. und OLIGINO, T. J.: Gene transfer of p53 to arthritic joints stimulates synovial apoptosis and inhibits inflammation. *Mol Ther* (2001), **3** (6): 901-10

YAO, Q., WANG, S., GAMBOTTO, A., GLORIOSO, J. C., EVANS, C. H., ROBBINS, P. D., GHIVIZZANI, S. C. und OLIGINO, T. J.: Intra-articular adenoviral-mediated gene transfer of trail induces apoptosis of arthritic rabbit synovium. *Gene Ther* (2003), **10** (12): 1055-60

YATSUGI, N., TSUKAZAKI, T., OSAKI, M., KOJI, T., YAMASHITA, S. und SHINDO, H.: Apoptosis of articular chondrocytes in rheumatoid arthritis and osteoarthritis: correlation of apoptosis with degree of cartilage destruction and expression of apoptosis-related proteins of p53 and c-myc. *J Orthop Sci* (2000), **5** (2): 150-6

YAZDANI-BIUKI, B., BRICKMANN, K., LANGSENLEHNER, U., RENNER, W., TRUSCHNIG, M., KRIPPL, P., FURST, F., GRANINGER, W. B. und BREZINSCHEK, H. P.: The Glu228Ala polymorphism in the ligand-binding domain of death receptor 4 is not associated with rheumatoid arthritis. *Ann Rheum Dis* (2008), **67** (7): 1053-4

ZAMBROWICZ, B. P. und SANDS, A. T.: Knockouts model the 100 best-selling drugs--will they model the next 100? *Nat Rev Drug Discov* (2003), **2** (1): 38-51

ZARBIN, M. A.: Current concepts in the pathogenesis of age-related macular degeneration. *Arch Ophthalmol* (2004), **122** (4): 598-614

ZHANG, H. G., WANG, Y., XIE, J. F., LIANG, X., LIU, D., YANG, P., HSU, H. C., RAY, R. B. und MOUNTZ, J. D.: Regulation of tumor necrosis factor alpha-mediated apoptosis of rheumatoid arthritis synovial fibroblasts by the protein kinase Akt. *Arthritis Rheum* (2001), **44** (7): 1555-67

ZHANG, R., WANG, Y., ZHANG, L., ZHANG, Z., TSANG, W., LU, M., ZHANG, L. und CHOPP, M.: Sildenafil (Viagra) induces neurogenesis and promotes functional recovery after stroke in rats. *Stroke* (2002), **33** (11): 2675-80

ZHANG, R., WANG, L., ZHANG, L., CHEN, J., ZHU, Z., ZHANG, Z. und CHOPP, M.: Nitric oxide enhances angiogenesis via the synthesis of vascular endothelial growth factor and cGMP after stroke in the rat. *Circ Res* (2003), **92** (3): 308-13

ZHENG, C. J., HAN, L. Y., YAP, C. W., JI, Z. L., CAO, Z. W. und CHEN, Y. Z.: Therapeutic targets: progress of their exploration and investigation of their characteristics. *Pharmacol Rev* (2006), **58** (2): 259-79

ZHU, Y. M., BAGSTAFF, S. M. und WOLL, P. J.: Production and upregulation of granulocyte chemotactic protein-2/CXCL6 by IL-1beta and hypoxia in small cell lung cancer. *Br J Cancer* (2006), **94** (12): 1936-41

ZIENOLDDINY, S., RYBERG, D., MAGGINI, V., SKAUG, V., CANZIAN, F. und HAUGEN, A.: Polymorphisms of the interleukin-1 beta gene are associated with increased risk of non-small cell lung cancer. *Int J Cancer* (2004), **109** (3): 353-6

10 Anhang

10.1 Detaillierte graphische Darstellung der Microsoft Access Abfragen

10.1.1 Abfragen zur Vernetzungsmatrix der therapeutischen Gebiete

T100_TAs and Pathways
TA name
Pathway

T100_TAs and Pathways_1
TA name
Pathway

Field:	TA name1: TA name	TA name2: TA name	Pathway1: Pathway	Pathway2: Pathway
Table:	T100_TAs and Pathways Pathways	T100_TAs and Pathways_1	T100_TAs and Pathways	T100_TAs and Pathways_1
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name1	TA name2	Pathway1	Pathway2
Criteria:				

Field:	Expr1: IIf([T100_TAs and Pathways]![TA name]=[T100_TAs and Pathways_1]![TA name];0;1)	Expr2: IIf([T100_TAs and Pathways]!Pathway=[T100_TAs and Pathways_1]!Pathway; 1;0)
Table:		
Total:		
Sort:		
Append to:	Expr1	Expr2
Criteria:	1	1

Abb. 10.1: **„Append"-Abfrage erstellte ein kartesisches Produkt zur Identifizierung der verbindenden Signalwege zwischen zwei therapeutischen Gebieten**
Für diese Abfrage war die Intermediärtabelle der vorherigen Abfrage (vergleiche Abb. 6.8) als Inputtabelle nötig. Durch zweimaliges Auflisten der Felder „TA name" und „Pathway" wurde zunächst ein kartesisches Produkt erstellt. Durch Verwenden von logischen Bedingungen („Expr1" und „Expr2") wurden in der durch die Abfrage befüllten Intermediärtabelle nur diejenigen Zeilen angezeigt, bei denen die therapeutischen Gebiete verschieden („Expr1") der Signalweg jedoch übereinstimmend („Expr2") ist.

T101_TA Matrix
TA name1
TA name2
Pathway1
Pathway2
Expr1
Expr2

Field:	TA name1	TA name2	Pathway1
Table:	T101_TA_Matrix	T101_TA_Matrix	T101_TA_Matrix
Total:	Group by	Group by	Count
Sort:			
Append to:	TA name1		CountOf Pathway1
Criteria:			

Abb. 10.2: „Append"-Abfrage ermittelte die Anzahl der gemeinsamen Signalwege zwischen zwei therapeutischen Gebieten

Die Intermediärtabelle der vorherigen Abfrage wurde als Inputtabelle für diese Abfrage verwendet. Die Felder „TA name1" und „TA name2" benannten das jeweilige Paar von therapeutischen Gebieten. Mit Hilfe einer Zählfunktion „Count" wurden die Namen der Signalwege aus dem Feld „Pathway1" gezählt und in das Feld „CountOf Pathway1" der neuen Intermediärtabelle geschrieben. Damit ordnete diese Abfrage jedem Paar von therapeutischen Gebieten eine Zahl verbindender Signalwege zu.

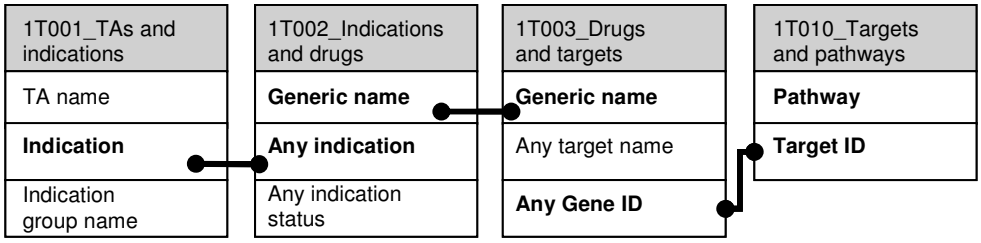
T102_TA Matrix_Count
TA name1
TA name2
CountOfPathway1

Field:	TA name1	TA name2	CountOfPathway1
Table:	T102_TA_Matrix_ Count	T102_TA_Matrix_ Count	T102_TA_Matrix_ Count
Total:	Group by	Group by	Sum
Crosstab:	Row Heading	Column Heading	Value

Abb. 10.3: „Crosstab"-Abfrage erstellte die Vernetzungsmatrix zwischen therapeutischen Gebieten in Form einer Kreuztabelle

Diese Abfrage nutzte die Intermediärtabelle der vorherigen Abfrage (Abb. 10.2), um eine Kreuztabelle zu erstellen. Dabei wurde das Feld „TA name1" auf der y-Achse der Matrix als Zeilenbeschriftung („Row Headings") und das Feld „TA name2" auf der x-Achse als Spaltenbeschriftung („Column Headings") verwendet. In die mittleren Felder der Matrix wurde jeweils die Anzahl der verbindenden Signalwege eingetragen.

10.1.2 Abfragen zur Vernetzungsmatrix der Indikationsgruppen



Field:	TA name	Sub TA name: Indication group name	Pathway
Table:	1T001_TAs and indications	1T001_TAs and indications	1T001_TAs and indications
Total:	Group by	Group by	Group by
Sort:			
Append to:	TA name	Sub TA name	Pathway
Criteria:			

Abb. 10.4: **„Append"-Abfrage ordnete Signalwege zu Indikationsgruppen zu**

Diese Abfrage verwendete die vier Inputtabellen „1T001_TAs and indications", „1T002_Indications and drugs", „1T003_Drugs and targets" und „1T010_Targets and pathways" zur Zuordnung von therapeutischen Gebieten („TA name"), Indikationsgrppen („Sub TA name"/„Indication group name") und Signalwegen („Pathway"). Der aus den Rohdaten hervorgehende längere Begriff „Indication group name" wurde im Rahmen dieser Abfrage in das synonyme kürzere „Sub TA name" umgewandelt. Die Verbindung zwischen Indikationsgruppe und Signalweg wurde durch die Beziehungen gleichartiger Felder zwischen den Inputtabellen gewährleistet.

Die neue Intermediärtabelle enthielt die Spalten „TA name", „SubTA name" und „Pathway" und gruppierte die Informationen so, dass jede Kombination nur einmal angezeigt wurde.

T110_SubTAs and Pathways
TA name
Sub TA name
Pathway

T110_SubTAs and Pathways_1
TA name
Sub TA name
Pathway

Field:	Sub TA name1: TA name	Sub TA name2: TA name	TA name1: TA name	TA name2: TA name
Table:	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1
Total:				
Sort:	Ascending	Ascending		
Append to:	Sub TA name1	Sub TA name2	TA name1	TA name2
Criteria:				

Field:	Pathway1: Pathway	Pathway2: Pathway	Expr1: IIf([T110_SubTAs and Pathways]![Sub TA name]=[T110_SubTAs and Pathways_1]![Sub TA name];0;1)	Expr2: IIf([T110_SubTAs and Pathways]!Pathway=[T110_SubTAs and Pathways_1]!Pathway; 1;0)
Table:	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1		
Total:				
Sort:	Ascending	Ascending		
Append to:	Pathway1	Pathway2	Expr1	Expr2
Criteria:			1	1

Abb. 10.5: **„Append“-Abfrage erstellte ein kartesisches Produkt und identifizierte die verbindenden Signalwege zwischen zwei Indikationsgruppen**
Inputtabelle für diese Abfrage war die Intermediärtabelle, die die Signalwege den Indikationsgruppen zugeordnet hatte. Durch zweimaliges Auflisten der Felder „TA name“, „Sub TA name“ und „Pathway“ wurde zunächst ein kartesisches Produkt erstellt. Durch Verwenden von logischen Bedingungen („Expr1“ und „Expr2“) wurden in der durch die Abfrage befüllten Intermediärtabelle nur diejenigen Zeilen angezeigt, bei denen die Indikationsgruppen verschieden („Expr1“) der Signalweg jedoch übereinstimmend („Expr2“) ist.

T111_SubTA Matrix
Sub TA name1
Sub TA name2
TA name1
TA name2
Pathway1
Pathway2
Expr1
Expr2

Field:	Sub TA name1	Sub TA name2	TA name1	TA name2
Table:	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix
Total:				
Sort:	Ascending	Ascending		
Append to:	Sub TA name1	Sub TA name2	TA name1	TA name2
Criteria:				

Field:	Pathway1
Table:	T111_SubTA_Matrix
Total:	Count
Sort:	
Append to:	CountOf Pathway1
Criteria:	

Abb. 10.6: : „Append"-Abfrage ermittelt die Anzahl gemeinsamer Signalwege zwischen zwei Indikationsgruppen
Die Intermediärtabelle der vorherigen Abfrage („T111_SubTA Matrix") wurde als Inputtabelle für diese Abfrage verwendet. Die Felder „Sub TA name1" und „Sub TA name2" benannten das jeweilige Paar von Indikationsgruppen. Zur weiteren Information wurden auch die entsprechenden therapeutischen Gebiete mit in die Tabelle aufgenommen. Mit Hilfe einer Zählfunktion „Count" wurden die Namen der Signalwege aus dem Feld „Pathway1" gezählt und in das Feld „CountOf Pathway1" der neuen Intermediärtabelle geschrieben. Damit ordnete diese Abfrage jedem Paar von Indikationsgruppen eine Zahl verbindender Signalwege zu.

T112_SubTA_Matrix_Pway_Count
Sub TA name1
Sub TA name2
TA name1
TA name2
CountOfPathway1

Field:	Sub TA name1	Sub TA name2	SumOfCountOfPath-way: CountOfPathway1
Table:	T112_SubTA_Matrix_ Pway_Count	T112_SubTA_Matrix_ Pway_Count	T112_SubTA_Matrix_ Pway_Count
Total:	Group by	Group by	Sum
Crosstab:	Row Heading	Column Heading	Value

Abb. 10.7: „Crosstab"-Abfrage erstellte die Vernetzungsmatrix zwischen den Indikationsgruppen in Form einer Kreuztabelle

Diese Abfrage nutzte die Intermediärtabelle der vorherigen Abfrage (Abb. 10.6, „T112_SubTA Matrix_Pway_Count"), um eine Kreuztabelle zu erstellen. Dabei wurde das Feld „Sub TA name1" auf der y-Achse der Matrix als Zeilenbeschriftung („Row Headings") und das Feld „Sub TA name2" auf der x-Achse als Spaltenbeschriftung („Column Headings") aufgetragen. In die mittleren Felder der Matrix wurde jeweils die Anzahl der verbindenden Signalwege eingetragen („Value").

10.1.3 Abfragen zur Vernetzungsanalyse ausgewählter therapeutischer Gebiete auf Ebene von Indikationsgruppen

T110_SubTAs and Pathways
TA name
Sub TA name
Pathway

T110_SubTAs and Pathways_1
TA name
Sub TA name
Pathway

Field:	Sub TA name1: TA name	Sub TA name2: TA name	TA name1: TA name	TA name2: TA name
Table:	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1
Total:				
Sort:	Ascending	Ascending		
Append to:	Sub TA name1	Sub TA name2	TA name1	TA name2
Criteria:			[Please enter TA name1]	[Please enter TA name2]

Field:	Pathway1: Pathway	Pathway2: Pathway	Expr1: IIf([T110_SubTAs and Pathways]![Sub TA name]=[T110_SubTAs and Pathways_1]![Sub TA name];0;1)	Expr2: IIf([T110_SubTAs and Pathways]!Pathway=[T110_SubTAs and Pathways_1]!Pathway; 1;0)
Table:	T110_SubTAs and Pathways	T110_SubTAs and Pathways_1		
Total:				
Sort:	Ascending	Ascending		
Append to:	Pathway1	Pathway2	Expr1	Expr2
Criteria:			1	1

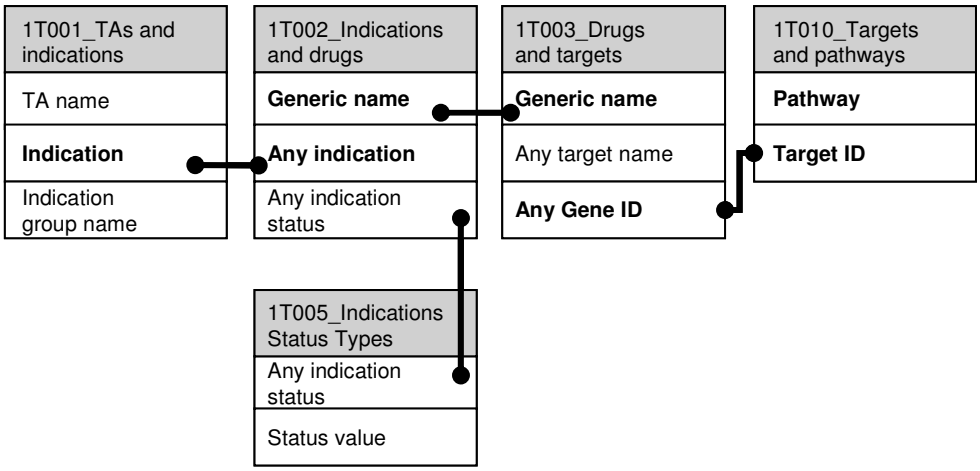
Abb. 10.8: „Append"-Abfrage erstellte ein kartesisches Produkt für bestimmte ausgewählte therapeutische Gebiete

Die oben dargestellte Abfrage gleicht von Aufbau der Abfrage aus Abb. 10.5. Inputtabelle war die Intermediärtabelle, die die Signalwege Indikationsgruppen zugeordnet hatte. Durch zweimaliges Auflisten der Felder „TA name", „Sub TA name" und „Pathway" wurde ein kartesisches Produkt erstellt. Durch Verwenden von logischen Bedingungen („Expr1" und „Expr2") wurden in der durch die Abfrage befüllten Intermediärtabelle nur diejenigen Zeilen angezeigt, bei denen die Indikationsgruppen verschieden („Expr1") der Signalweg jedoch übereinstimmend („Expr2") ist.

In der „Criteria"-Zeile wurden statt festen Auswahlkriterien die Parameter [Please enter TA name1 bzw. 2] eingegeben. Diese sogenannte „Parameter"-Abfrage führte zum Öffnen von Eingabefenstern. In die Eingabefenster konnten diejenigen therapeutischen Gebiete eingegeben werden, die im Detail untersucht werden sollten. Das kartesische Produkt wurde dann nur für die ausgewählten therapeutischen Gebiete berechnet, was die Rechenzeit deutlich reduzierte.

10.1.4 Abfragen zur Gewichtung der Vernetzung von therapeutischen Gebieten

10.1.4.1 Abfragen zur Gewichtung ohne Durchschnittsbildung



Field:	TA name	Pathway	Target ID	Generic Name
Table:	1T001_TAs and indications	1T010_Targets and pathways	1T010_Targets and pathways	1T002_Indications and drugs
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name	Pathway	Target ID	Generic Name
Criteria:				

Field:	IndCount: Any Indication	Any Indication Status	Status Value
Table:	1T002_Indications and drugs	1T002_Indications and drugs	1T305_Indication Status Types
Total:	Count	Group By	Group By
Sort:		Ascending	
Append to:	IndCount: Any Indication	Any Indication Status	Status Value
Criteria:			

Abb. 10.9: „Append"-Abfrage „Weighting Base" vernetzte alle Rohdaten und ordnete den Gewichtungsfaktor zu

Die „Weighting Base" Abfrage war die erste Abfrage der Gewichtung und benötigte die Rohdatentabellen 1T001, 1T002, 1T003, 1T005 und 1T010, die jeweils über gleichartige Felder miteinander in Beziehung standen.

Die Abfrage ordnete pro therapeutischen Gebiet („TA name") und Signalweg („Pathway") diejenigen Zielmoleküle zu („Target ID"), die eine Indikation im therapeutischen Gebiet haben. Diese Zuordnung erfolgte durch Verknüpfung eines Zielmolekül mit seinen Substanzen („Generic name") und den für die Substanzen vorhandenen Forschungsprojekte („Any indication"). Dabei wurden pro therapeutischem Gebiet, Signalweg und Zielmolekül alle relevanten Substanzen aufgelistet. Für jede Substanz wurde mit Hilfe einer „Count" Funktion die Anzahl aller Indikationen mit demselben Forschungsstatus ermittelt. Jeder Gruppe von Indikationen mit demselben Status wurde aus der Rodatentabelle 1T010 der entsprechende Gewichtungsfaktor zugeordnet.

T200J_Weighting_ Base
TA name
Pathway
Target ID
Generic name
IndCount
Any Indication Status
Status Value

Field:	TA name	Pathway	Target ID	Generic Name
Table:	T200J_Weighting_ Base	T200J_Weighting_ Base	T200J_Weighting_ Base	T200J_Weighting_ Base
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Pathway	Target ID	Generic Name
Criteria:				

Field:	WeigtedInd: [IndCount]*[Status Value]	IndCount	Status Value	Any Indication Status
Table:		T200J_Weighting_ Base	T200J_Weighting_ Base	T200J_Weighting_ Base
Total:				
Sort:				
Append to:	WeigtedInd	IndCount	Status Value	Any Indication Status
Criteria:			1	1

Abb. 10.10: „Append"-Abfrage „Weighting Base 1" berechnete ein Gesamtgewicht pro Entwicklungsstatus und Substanz

Inputtabelle für die zweite Gewichtungsabfrage war die Intermediärtabelle der Abfrage „Weighting Base".

Mit Hilfe einer Multiplikationsformel (WeigtedInd: [IndCount]*[Status Value]) wurde die Anzahl der Indikationen einer Substanz pro Entwicklungsstatus („IndCount") mit dem Gewichtungsfaktor („Status Value") multipliziert und das Gewicht wurde in einem neuen Feld als „WeigtedInd" angezeigt. Die Berechnung erfolgte dabei wie bei der Abfrage „Weighting Base" jeweils speziell in jedem therapeutischen Gebiet und Signalweg. Die letzten drei Felder der neuen Intermediärtabelle listeten zu Informationszwecken noch einmal die in die Berechnung eingegangenen Rohdaten (Anzahl der Indikationen = „IndCount", Gewichtungsfaktor = „Status Value" und den zum Gewichtungsfaktor gehörenden Status = „Any indication Status").

T201J_Weighting_ Base1
TA name
Pathway
Target ID
Generic name
WeigtedInd
IndCount
Any Indication Status
Status Value

Field:	TA name	Pathway	Target ID	Generic Name
Table:	T201J_Weighting_ Base1	T201J_Weighting_ Base1	T201J_Weighting_ Base1	T201J_Weighting_ Base1
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Pathway	Target ID	Generic Name
Criteria:				

Field:	IndCount	WeigtedInd
Table:	T201J_Weighting_ Base1	T201J_Weighting_ Base1
Total:	Sum	Sum
Sort:		
Append to:	SumOfIndCount	SumOfWeigtedInd
Criteria:		

Abb. 10.11: „Append"-Abfrage „Weighting Base 2" summierte die Gewichte der Entwicklungsstatus pro Substanz

Inputtabelle für die dritte Gewichtungsabfrage war die Intermediärtabelle der Abfrage „Weighting Base 1". Genauso wie bei den vorherigen Abfragen wurden zunächst Name des therapeutischen Gebiets („TA name"), des Signalwegs („Pathway"), des Zielmoleküls („Target ID") und der Substanz („Generic name") in eine neue Intermediärtabelle gezogen.

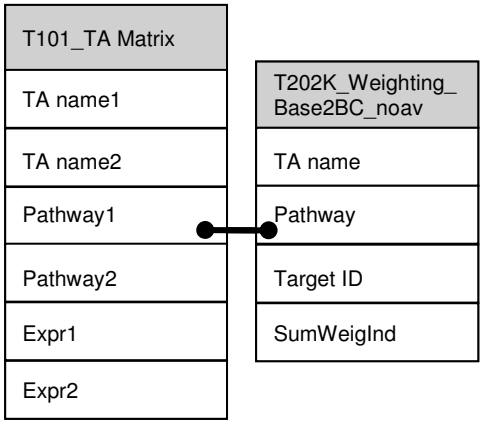
Eine Summenfunktion „Sum" addierte die Anzahl der Indikationen pro Entwicklungsstatus einer Substanz („IndCount") und ermittelte so die Gesamtzahl der Indikationen dieser Substanz innerhalb des therapeutischen Gebiets („SumOfIndCount"). Ebenfalls über eine Summenfunktion wurden die Gesamtgewichte pro Entwicklungsstatus einer Substanz („WeigtedInd") zu einem Gesamtgewicht für die Substanz berechnet („SumOfWeigtedInd").

T202J_Weighting_ Base2
TA name
Pathway
Target ID
Generic name
SumOfIndCount
SumOfWeigtedInd

Field:	TA name	Pathway	Target ID	SumWeigInd: SumOfWeigtedInd
Table:	T202J_Weighting_ Base2	T202J_Weighting_ Base2	T202J_Weighting_ Base2	T202J_Weighting_ Base2
Total:	Group By	Group By	Group By	Sum
Sort:	Ascending	Ascending		
Append to:	TA name	Pathway	Target ID	SumWeigInd
Criteria:				

Abb. 10.12: „Append"-Abfrage „Weighting Base 2BC noav" berechnete ein Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und eines therapeutischen Gebiets („TA name"). Die Berechnung des Gewichts pro Zielmolekül erfolgte über eine Summenfunktion „Sum", die pro Zielmolekül alle gewichteten Substanzen („SumOfWeigtedInd") addierte. Das Gewicht pro Zielmolekül wurde als „SumWeigInd" bezeichnet. Wie an der tabellarischen Darstellung erkannt werden kann, war jetzt das Zielmolekül („Target ID") und nicht mehr die Substanz („Generic name") die detaillierteste Ebene der Tabelle (vergleiche Abb. 10.10, Abb. 10.11, Abb. 10.12).



Field:	TA name1	TA name2	Pathway1	SumWeigInd
Table:	T101_TA_Matrix	T101_TA_Matrix	T101_TA_Matrix	T202K_Weighting_ Base2BC_noav
Total:	Group By	Group By	Group By	Sum
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name1	TA name2	Pathway1	SumOfSumWeigInd
Criteria:				

Field:	Expr3: IIf(T101_TA_Matrix![TA name1]=T202K_Weighting_Base2BC_noav![TA name] Or T101_TA_Matrix![TA name2]=T202K_Weighting_Base2BC_noav![TA name];1;0)
Table:	
Total:	Group By
Sort:	
Append to:	Expr3
Criteria:	1

Abb. 10.13: „Append"-Abfrage „Weighting Base 3 noav" berechnete ein Gesamtgewicht pro Signalweg

Inputtabellen für die Abfrage war die Intermediärtabelle der vorhergehenden Abfrage, die das Gesamtgewicht pro Zielmolekül berechnet hatte („SumWeigInd") sowie die Tabelle der Vernetzungsanalyse der therapeutischen Gebiete. Die Felder, die den Namen des Signalwegs enthalten („Pathway1" und „Pathway") waren durch eine Beziehung miteinander verbunden. Diese Beziehung stellte sicher, dass bei der Abfrage nur diejenigen Signalwege berücksichtigt wurden, die bereits in der Tabelle „T101_TA Matrix" als verbindend identifiziert worden waren.

Die Abfrage zog die Namen eines Paars von therapeutischen Gebieten („TA name1" und „TA name2") sowie den Namen eines verbindenden Signalwegs („Pathway") aus den Inputtabellen. Pro Signalweg wurden über eine Summenfunktion „Sum" die Gewichte der Zielmoleküle zu einem Gesamtgewicht des Signalwegs („SumOfSumWeigInd") addiert. Die logische Bedingung „Expr3" stellte sicher, dass dabei nur die Zielmoleküle berücksichtigt wurden, für die es Substanzen mit Indikationen in den ausgewählten therapeutischen Gebieten gibt.

T203K_Weighting_ Base3_noav
TA name1
TA name2
Pathway1
CountOfTargetID
SumOfSumWeigInd
Expr3

Field:	TA name1	TA name2	Pathway Score: SumOfSumWeigInd	CountOfPathway11: Pathway1
Table:	T203K_Weighting_ Base3_noav	T203K_Weighting_ Base3_noav	T203K_Weighting_ Base3_noav	T203K_Weighting_ Base3_noav
Total:	Group By	Group By	Sum	Count
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name1	TA name2	PathwayScore	PathwayCount
Criteria:				

Abb. 10.14: **„Append"-Abfrage „Weighting Base 3B noav" berechnete die gewichtete Verbindungsstärke zwischen zwei therapeutischen Gebieten**
Inputtabelle für den letzten Schritt der Gewichtungsabfragen war die Inputtabelle der Abfrage „Weighting Base 3 noav".
Über eine Summenfunktion „Sum" wurden die gewichteten Verbindungsstärken der einzelnen verbindenden Signalwege („SumOfSumWeigInd") addiert. Es resultierte ein gewichteter Gesamtwert für die Verbindungsstärke zwischen zwei therapeutischen Gebieten („Pathway Score"). Durch eine Zählfunktion „Count" wurde außerdem die Anzahl der verbindenden Signalwege berechnet („PathwayCount").

10.1.4.2 Abfragen zur Gewichtung mit einfacher Durchschnittsbildung

T202J_Weighting_ Base2
TA name
Pathway
Target ID
Generic name
SumOfIndCount
SumOfWeigtedInd

Field:	TA name	Pathway	Target ID	Generic name
Table:	T202J_Weighting_ Base2	T202J_Weighting_ Base2	T202J_Weighting_ Base2	T202J_Weighting_ Base2
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Pathway	Target ID	Generic name
Criteria:				

Field:	AvgInd: [SumOfWeigtedInd]/ [SumOfIndCount]
Table:	
Total:	
Sort:	Ascending
Append to:	AvgInd
Criteria:	

Abb. 10.15: „Append"-Abfrage „Weighting Base 2B" berechnete das Durchschnittsgewicht pro Substanz bei der Gewichtung mit einfacher Durchschnittsbildung

Inputtabelle für diese Abfrage war die Intermediärtablelle der Abfrage „Weighting Base 2" (vergleiche Abb. 10.11), die für jede Substanz („Generic name") pro therapeutischem Gebiet („TA name") und Signalweg („Pathway") die Summe aller gewichteten Indikationen („SumOfWeigtedInd") sowie die absolute Anzahl der Indikationen („SumOfIndCount") berechnet hatte.

Die Abfrage „Weighting Base 2B" berechnete nun das durchschnittliche Gewicht einer Substanz indem die Summe der gewichteten Indikationen durch die absolute Anzahl geteilt wurde ([SumOfWeigtedInd]/[SumOfIndCount]). Dieses durchschnittliche Gewicht (AvgInd) wurde in eine neue Intermediärtabelle geschrieben.

T202J_Weighting_ Base2B
TA name
Pathway
Target ID
Generic name
AvgInd

Field:	TA name	Pathway	Target ID	AvgInd
Table:	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B
Total:	Group By	Group By	Group By	Sum
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name	Pathway	Target ID	SumAvgInd
Criteria:				

Abb. 10.16: „Append"-Abfrage „Weighting Base 2C" berechnete das Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und eines therapeutischen Gebiets („TA name"). Die Berechnung des Gesamtgewichts erfolgte über eine Summenfunktion „Sum", die pro Zielmolekül alle Durchschnitts-Gewichte der Substanzen („AvgInd") addierte. Das Gewicht pro Zielmolekül wurde als „SumAvgInd" bezeichnet.

10.1.4.3 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung

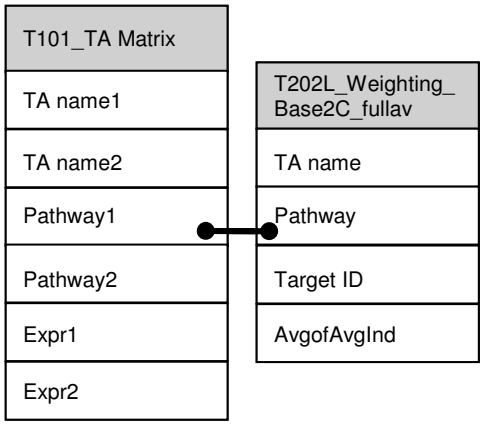
T202J_Weighting_ Base2B
TA name
Pathway
Target ID
Generic name
AvgInd

Field:	TA name	Pathway	Target ID	AvgInd
Table:	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B	T202J_Weighting_ Base2B
Total:	Group By	Group By	Group By	Avg
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name	Pathway	Target ID	AvgofAvgInd
Criteria:				

Abb. 10.17: „Append"-Abfrage „Weighting Base 2C fullav" berechnete das Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und eines therapeutischen Gebiets („TA name"). Mit Hilfe einer Durchschnittsfunktion „Avg" wurde dazu pro Zielmolekül der Durchschnitt alle Durchschnitts-Gewichte der Substanzen („AvgInd)" gebildet. Dieses neue Durchschnitts-Gewicht pro Zielmolekül wurde als „AvgofAvgInd" bezeichnet.

10.1.4.4 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung und Multiplikator



Field:	TA name1	TA name2	Pathway1	Target ID
Table:	T101_TA_Matrix	T101_TA_Matrix	T101_TA_Matrix	T202L_Weighting_ Base2C_fullav
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Show:	✓	✓	✓	✓
Criteria:				

Field:	AvgofAvgInd	Expr3: IIf(T101_TA_Matrix![TA name1]=T202L_Weighting_Base2C_fullav![TA name] Or T101_TA_Matrix![TA name2]=T202L_Weighting_Base2C_fullav![TA name];1;0)
Table:	T202L_Weighting_ Base2C_fullav	
Total:		
Sort:	Ascending	
Show:	✓	✓
Criteria:		1

Abb. 10.18: „Select"-Abfrage „Weighting Base 3A fullav" listete alle Zielmoleküle und ihre Gewichte für die Gewichtung mit doppelter Durchschnittbildung und Multiplikator

Die hier dargestellte Abfrage war eine sogenannte „Select query". Im Gegensatz zur „Append"-Abfrage wurde keine Intermediärtabelle befüllt, sondern nur eine temporäre Tabelle generiert. Felder, die in der Zeile „Show" der tabellarischen Ansicht einen Haken haben, wurden in der temporären Tabelle gezeigt. Eine Inputtabelle für diese Abfrage war die Intermediärtabelle der Abfrage „Weighting Base 2C fullav", die bei Gewichtung mit doppelter Durchschnittsbildung das Durchschnitts-Gewicht der Zielmoleküle berechnete („AvgofAvgInd"). Diese wurde über eine Beziehung mit der Intermediärtabelle der Vernetzungsanalyse therapeutischer Gebiete verbunden, in der die Signalwege gelistet waren, die zwei therapeutische Gebiete verbinden. Mit Hilfe der „Select"-Abfrage wurde eine temporäre Liste erstellt, in der für jedes Paar von therapeutischen Gebieten und jeden verbindenden Signalweg alle Zielmoleküle („Target ID") mit ihren Durchschnitts-Gewichten („AvgofAvgInd") aufgelistet wurden, die entweder für eines der beiden therapeutischen Gebiete relevant sind („Expr3"). Dabei galt, dass Zielmoleküle, die für beide Gebiete von Bedeutung sind zweifach aufgezählt wurden.

Q203X_Weighting_ Base3A_fullav
TA name1
TA name2
Pathway1
Target ID
AvgofAvgInd
Expr3

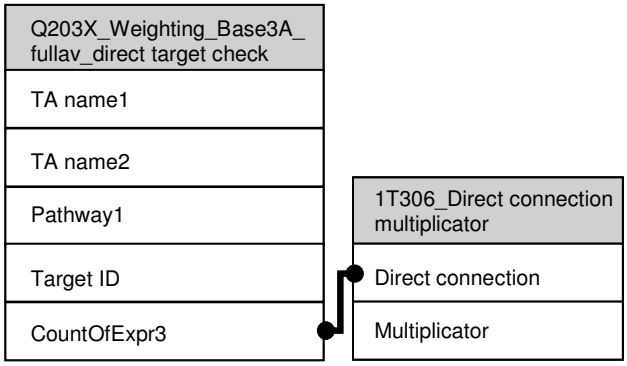
Field:	TA name1	TA name2	Pathway1	Target ID
Table:	Q203X_Weighting_ Base3A_fullav	Q203X_Weighting_ Base3A_fullav	Q203X_Weighting_ Base3A_fullav	Q203X_Weighting_ Base3A_fullav
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Expr3
Table:	Q203X_Weighting_ Base3A_fullav
Total:	Count
Sort:	
Show:	✓
Criteria:	

Abb. 10.19: „Select"-Abfrage „Weighting Base 3A fullav direct target check" identifizierte alle direkt indikationsübergreifenden Zielmoleküle

Diese „Select"-Abfrage verwendete das Ergebnis der vorhergehenden Abfrage, die für jede Kombination von therapeutischen Gebieten und verbindendem Signalweg die relevanten Zielmoleküle mit ihrem Durchschnittsgewicht aufgelistet hatte, als Input.

Durch eine Zählfunktion „Count" wurde gezählt, wie oft ein Zielmolekül (Target ID) für eine solche Kombination in der Liste erwähnt wurde. Dazu wurden die Zielmoleküleinträge gruppiert („Group by" von „Target ID") und die Anzahl der Zeilen durch Zählen des Feldes „Expr3" ermittelt. Dabei wurde jedem Zielmolekül, das für beide therapeutischen Gebiete von Bedeutung ist, die Ziffer 2 zugeordnet und jedem Zielmolekül, das nur in einem der beiden Gebiete aktiv ist, die Ziffer 1.



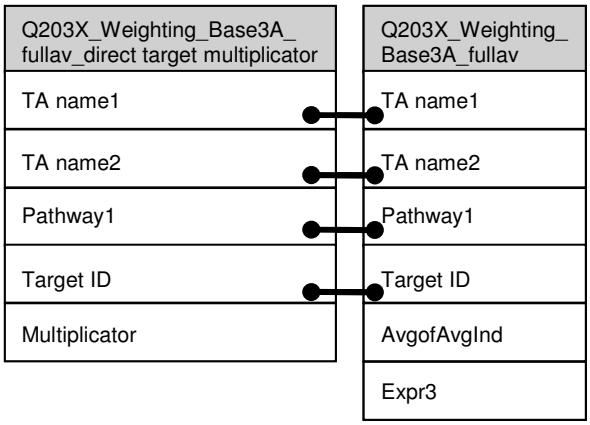
Field:	TA name1	TA name2	Pathway1	Target ID
Table:	Q203X_Weighting_Base3A_fullav_direct target check	Q203X_Weighting_Base3A_fullav_direct target check	Q203X_Weighting_Base3A_fullav_direct target check	Q203X_Weighting_Base3A_fullav_direct target check
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Multiplicator
Table:	1T306_Direct connection multiplicator
Total:	Group By
Sort:	
Show:	✓
Criteria:	

Abb. 10.20: “Select”-Abfrage “Weighting Base 3A fullav direct target multiplicator” ordnete den Zielmolekülen den Multiplikator zu

Inputtabellen für diese Abfrage waren das Ergebnis der vorhergehenden „Select“-Abfrage, die durch Zuordnung einer 1 bzw. 2 im Feld „CountOfExpr3“ zwischen direkten und indirekten Verbindungen unterschieden hatte, sowie die Inputtabelle 1T306, die den Wert des Multiplikators gespeichert hatte.

Die Abfrage ordnete jedem Zielmolekül, das an einer Signalwegverbindung von zwei therapeutischen Gebieten teilhat, entsprechend seines „CountOfExpr3“-Werts einen Multiplikator zu („Multiplicator“).



Field:	TA name1	TA name2	Pathway1	Target ID
Table:	Q203X_Weighting_Base3A_fullav	Q203X_Weighting_Base3A_fullav	Q203X_Weighting_Base3A_fullav	Q203X_Weighting_Base3A_fullav
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Show:	✓	✓	✓	✓
Criteria:				

Field:	MulAvgofAvgInd: [Multiplicator]*[AvgofAvgInd]
Table:	
Total:	
Sort:	
Show:	✓
Criteria:	

Abb. 10.21: „Select"-Abfrage „Weighting Base 3A2 fullav" multiplizierte die durchschnittlichen Zielmolekülgewichte mit dem Multiplikator

Inputtabellen für diese Abfrage waren das Ergebnis der ersten Gewichtungsabfrage mit Multiplikator (siehe Abb. 10.18) sowie das Ergebnis des ersten Schritts der Multiplikator-Anwendung (siehe Abb. 10.20). Die gleichartigen Felder der beiden Tabellen wurden durch Beziehungen verbunden.

Die Abfrage berechnete durch eine algebraische Operation das Produkt aus dem durchschnittlichen Gewicht des Zielmoleküls („AvgofAvgInd") und des Multiplikators („Multiplicator"). Im Endprodukt entstand ein neues Gewicht, das mit „MulAvsofAvgInd" bezeichnet wurde.

Q203X_Weighting_ Base3A2_fullav
TA name1
TA name2
Pathway1
Target ID
MulAvgofAvgInd

Field:	TA name1	TA name2	Pathway1	SumofPWay: MulAvgofAvgInd
Table:	Q203X_Weighting_ Base3A2_fullav	Q203X_Weighting_ Base3A2_fullav	Q203X_Weighting_ Base3A2_fullav	Q203X_Weighting_ Base3A2_fullav
Total:	Group By	Group By	Group By	Sum
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

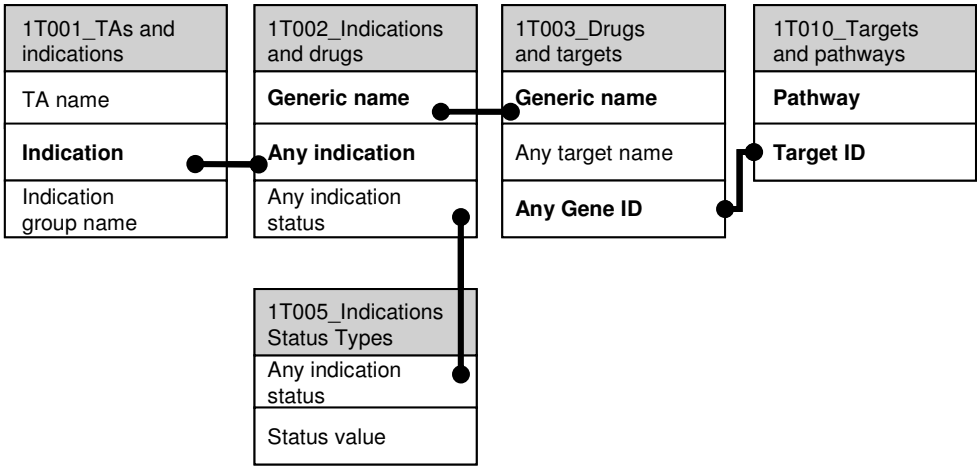
Abb. 10.22: „Select"-Abfrage „Weighting Base 3 fullav" berechnete ein Gesamtgewicht pro Signalweg

Die Abfrage verwendete das Ergebnis der vorhergehenden Abfrage, die ein Gesamtgewicht pro Zielmolekül unter Berücksichtigung des Multiplikators berechnet hatte.

Durch eine Summenfunktion „Sum" wurden die einzelnen Gewichte der Zielmoleküle addiert und zu einem Gesamtgewicht eines Signalwegs („SumofPWay") in der Verbindung von zwei therapeutischen Gebieten („TA name1" und „TA name2") zusammengefügt.

10.1.5 Abfragen zur Gewichtung der Vernetzung von Indikationsgruppen

10.1.5.1 Abfragen zur Gewichtung ohne Durchschnittsbildung



Field:	TA name	Sub TA name: Indication group name	Pathway	Target ID
Table:	T001_TAs and indications	T001_TAs and indications	1T010_Targets and pathways	1T010_Targets and pathways
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	Generic Name	IndCount: Any Indication	Any Indication Status	Status Value
Table:	T002_Indications and drugs	1T002_Indications and drugs	1T002_Indications and drugs	1T305_Indication Status Types
Total:	Group By	Count	Group By	Group By
Sort:			Ascending	
Append to:	Generic Name	IndCount	Any Indication Status	Status Value
Criteria:				

Abb. 10.23: „Append"-Abfrage „SubTA Weighting Base" für Indikationsgruppen vernetzte alle Rohdaten und ordnete den Gewichtungsfaktor zu

Die „Weighting Base" Abfrage war die erste Abfrage der Gewichtung der Indikationsgruppen und benötigte die Rohdatentabellen 1T001, 1T002, 1T003, 1T005 und 1T010, die jeweils über gleichartige Felder miteinander in Beziehung standen.

Die Abfrage ordnete pro Indikationsgruppe („SubTA name") und Signalweg („Pathway") diejenigen Zielmoleküle zu („Target ID"), die eine Indikation in der entsprechenden Indikationsgruppe haben. Diese Zuordnung erfolgte durch Verknüpfung des Zielmolekül („Target ID") mit seinen Substanzen („Generic name") und den für die Substanzen vorhandenen Forschungsprojekten („Any indication"). Dabei wurden pro Indikationsgruppe, Signalweg und Zielmolekül alle relevanten Substanzen aufgelistet und für jede Substanz mit Hilfe einer „Count" Funktion die Anzahl aller Indikationen mit demselben Forschungsstatus gezählt. Pro Gruppe von Indikationen mit demselben Status wurde aus der Rohdatentabelle 1T010 der dazugehörige Gewichtungsfaktor „Status Value" zugeordnet.

T220J_SubTAs_ Weighting_Base
TA name
Sub TA name
Pathway
Target ID
Generic Name
IndCount
Any Indication Status
Status Value

Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T220J_SubTAs_ Weighting_Base	T220J_SubTAs_ Weighting_Base	T220J_SubTAs_ Weighting_Base	T220J_SubTAs_ Weighting_Base
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	Generic Name	WeigtedInd: [IndCount]*[Status Value]	IndCount	Status Value
Table:	T220J_SubTAs_ Weighting_Base		T220J_SubTAs_ Weighting_Base	T220J_SubTAs_ Weighting_Base
Total:				
Sort:	Ascending			
Append to:	Generic Name	WeigtedInd	IndCount	Status Value
Criteria:				

Field:	Any Indication Status
Table:	T220J_SubTAs_ Weighting_Base
Total:	
Sort:	
Append to:	Any Indication Status
Criteria:	

Abb. 10.24: **„Append"-Abfrage „SubTA Weighting Base 1" berechnete ein Gesamtgewicht pro Entwicklungsstatus und Substanz**
Die in der Abbildung gezeigte Abfrage führte die eigentliche Multiplikation mit dem Gewichtungsfaktor durch. Mit Hilfe einer Multiplikationsformel (WeigtedInd: [IndCount]*[Status Value]) wurden die Anzahl der Indikationen einer Substanz pro Entwicklungsstatus („IndCount") und der Gewichtungsfaktor („Status Value") multipliziert. Das Gesamtgewicht wurde in einem neuen Feld als „WeigtedInd" angezeigt. Die Berechnung erfolgte dabei wie bei der Abfrage „Weighting Base" jeweils speziell in jedem therapeutischen Gebiet und Signalweg. Die letzten drei Felder der neuen Intermediärtabelle listeten zu Informationszwecken noch einmal die in die Berechnung eingegangenen Rohdaten (Anzahl der Indikationen „IndCount", Gewicht „Status Value") und den zum Gewicht gehörenden Status („Any indication Status") auf.

T221J_SubTAs_ Weighting_Base1
TA name
Sub TA name
Pathway
Target ID
Generic Name
WeigtedInd
IndCount
Status Value
Any Indication Status

Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T221J_SubTAs_ Weighting_Base1	T221J_SubTAs_ Weighting_Base1	T221J_SubTAs_ Weighting_Base1	T221J_SubTAs_ Weighting_Base1
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	Generic Name	SumOfIndCount1: IndCount	SumOfWeigtedInd1: WeigtedInd
Table:	T221J_SubTAs_ Weighting_Base1	T221J_SubTAs_ Weighting_Base1	T221J_SubTAs_ Weighting_Base1
Total:	Group By	Sum	Sum
Sort:	Ascending		
Append to:	Generic Name	SumOfIndCount	SumOfWeigtedInd
Criteria:			

Abb. 10.25: „Append"-Abfrage „SubTA Weighting Base 2" summierte die Gewichte der Entwicklungsstatus pro Substanz

Inputtabelle für die dritten Gewichtungsabfrage war die Intermediärtabelle der Abfrage „SubTA Weighting Base 1". Genauso wie bei den vorherigen Abfragen wurden zunächst Name des therapeutischen Gebiets („TA name"), der Indikationsgruppe („Sub TA"), des Signalwegs („Pathway"), des Zielmoleküls („Target ID") und der Substanz („Generic name") in eine neue Intermediärtabelle gezogen.
Eine Summenfunktion „Sum" addierte die Anzahl der Indikationen pro Entwicklungsstatus einer Substanz („IndCount") und ermittelte so die Gesamtanzahl der Indikationen dieser Substanz innerhalb des therapeutischen Gebiets („SumOfIndCount"). Genauso wurden mit Hilfe einer Summenfunktion die Gewichte pro Entwicklungsstatus einer Substanz („WeigtedInd") zu einem Gesamtgewicht für die Substanz berechnet („SumOfWeigtedInd").

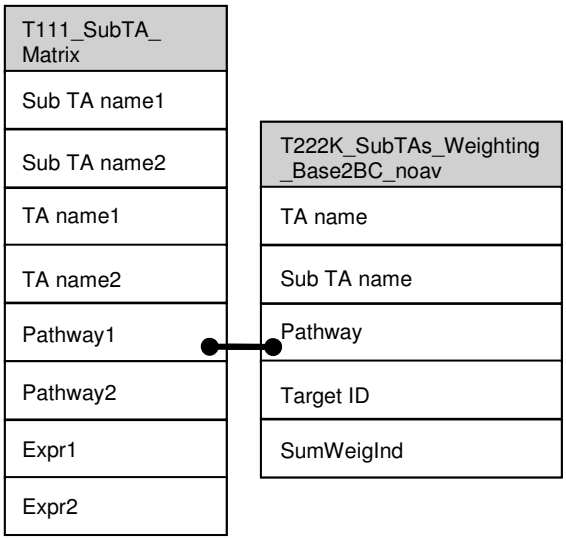
T222J_SubTAs_ Weighting_Base2
TA name
Sub TA name
Pathway
Target ID
Generic Name
SumOfIndCount
SumOfWeigtedInd

Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	SumWeigInd: SumOfWeigtedInd
Table:	T222J_SubTAs_ Weighting_Base2
Total:	Sum
Sort:	
Append to:	SumWeigInd
Criteria:	

Abb. 10.26: „Append"-Abfrage „SubTA Weighting Base 2BC noav" berechnete ein Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und einer Indikationsgruppe („Sub TA name"). Die Berechnung des Gesamtgewichts erfolgte über eine Summenfunktion „Sum", die pro Zielmolekül alle gewichteten Substanzen („SumOfWeigtedInd") addierte. Das Gewicht pro Zielmolekül wurde als „SumWeigInd" bezeichnet. Wie an der tabellarischen Darstellung erkannt werden kann, war jetzt das Zielmolekül („Target ID") und nicht mehr die Substanz („Generic name") die detaillierteste Ebene der Tabelle (vergleiche Abb. 10.23, Abb. 10.24 und Abb. 10.25). Der Name des therapeutischen Gebiets („TA name"), in das die Indikationsgruppe gehört, wurde zu Informationszwecken in die Tabelle aufgenommen.



Field:	Sub TA name1	Sub TA name2	TA name1	TA name2
Table:	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending		
Append to:	Sub TA name1	Sub TA name2	TA name1	TA name2
Criteria:				

Field:	Pathway1	SumWeigInd	Expr3: IIf(T111_SubTA_Matrix ![Sub TA name1]= T222K_SubTAs_Weighting_Base2BC_noav![Sub TA name] Or T111_SubTA_Matrix![Sub TA name2]= T222K_SubTAs_Weighting_Base2BC_noav![Sub TA name];1;0)	Expr4: IIf(T111_SubTA_Matrix ![TA name1] <>T111_SubTA_Matrix ![TA name2];1;0)
Table:	T111_SubTA_Matrix	T222K_SubTAs_Weighting_Base2BC_noav		
Total:	Group By	Sum	Group By	Group By
Sort:	Ascending			
Append to:	Pathway1	SumOfSumWeigInd	Expr3	Expr4
Criteria:			1	1

Abb. 10.27: „Append"-Abfrage „SubTA Weighting Base 3 noav" berechnete ein Gesamtgewicht pro Signalweg

Inputtabellen für die Abfrage waren die Intermediärtabelle der vorhergehenden Abfrage, die das Gesamtgewicht pro Zielmolekül berechnet hatte („SumWeigInd") und die Tabelle der Vernetzungsanalyse der Indikationsgruppen („T111_SubTA Matrix"). Die Felder, die den Namen des Signalwegs enthalten („Pathway1" und „Pathway") wurden durch eine Beziehung miteinander verbunden. Diese Beziehung stellte sicher, dass bei der Abfrage nur diejenigen Signalwege berücksichtigt wurden, die bereits in der Tabelle „T111_SubTA Matrix" als Verbindung zwischen zwei Indikationsgruppen identifiziert worden waren.

Die Abfrage zog die Namen eines Paars von Indikationsgruppen („Sub TA name1" und „Sub TA name2") sowie den Namen eines verbindenden Signalwegs („Pathway") aus den Inputtabellen.

(Fortsetzung der Bildunterschrift auf nächster Seite)

*(Fortsetzung der Bildunterschrift von **Abb. 10.27**.)*
Pro Signalweg wurden über eine Summenfunktion „Sum“ die Gewichte der Zielmoleküle zu einem Gesamtgewicht des Signalwegs („SumOfSumWeigInd“) addiert. Die logische Bedingung „Expr3“ stellte sicher, dass dabei nur die Zielmoleküle berücksichtigt wurden, für die es Substanzen mit Indikationen in den ausgewählten therapeutischen Gebieten gibt. Die logische Bedingung „Expr4“ sorgte dafür, dass ein und derselbe Signalweg in der neu befüllten Intermediärtabelle nicht doppelt eingetragen wurde.

T223K_SubTAs_ Weighting_Base3_noav
Sub TA name1
Sub TA name2
TA name1
TA name2
Pathway1
SumOfSumWeigInd
Expr3
Expr4

Field:	TA name1	TA name2	Sub TA name1	Sub TA name2
Table:	T223K_SubTAs_Weighting_Base3_noav	T223K_SubTAs_Weighting_Base3_noav	T223K_SubTAs_Weighting_Base3_noav	T223K_SubTAs_Weighting_Base3_noav
Total:	Group By	Group By	Group By	Group By
Sort:				
Append to:	TA name1	TA name2	Sub TA name1	Sub TA name2
Criteria:				

Field:	Pathway Score: SumOfSumWeigInd	Pathway1
Table:	T223K_SubTAs_Weighting_Base3_noav	T223K_SubTAs_Weighting_Base3_noav
Total:	Sum	Count
Sort:		
Append to:	PathwayScore	PathwayCount
Criteria:		

Abb. 10.28: „Append“-Abfrage „SubTA Weighting Base 3B noav“ berechnete die gewichtete Verbindungsstärke zwischen zwei Indikationsgruppen
Inputtabelle für den letzten Schritt der Gewichtungsabfragen war die Inputtabelle der Abfrage „SubTA Weighting Base 3 noav“.
Über eine Summenfunktion „Sum“ wurden die gewichteten Verbindungsstärken der einzelnen verbindenden Signalwege („SumOfSumWeigInd“) addiert. Es resultierte ein gewichteter Gesamtwert für die Verbindungsstärke zwischen Indikationsgruppen („Pathway Score“). Durch eine Zählfunktion „Count“ wurde zusätzlich die Anzahl der verbindenden Signalwege berechnet („PathwayCount“). Die Namen der therapeutischen Gebiete („TA name1“ und „TA name2“), zu denen die Indikationsgruppen gehören, dienten nur der Information und waren für die eigentliche Berechnung unerheblich.

10.1.5.2 Abfragen zur Gewichtung mit einfacher Durchschnittsbildung

T222J_SubTAs_ Weighting_Base2
TA name
Sub TA name
Pathway
Target ID
Generic Name
SumOfIndCount
SumOfWeigtedInd

Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2	T222J_SubTAs_ Weighting_Base2
Total:				
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	Generic Name	AvgInd: [SumOfWeigtedInd]/ [SumOfIndCount]
Table:	T222J_SubTAs_ Weighting_Base2	
Total:		
Sort:		
Append to:	PathwayScore	AvgInd
Criteria:		

Abb. 10.29: „Append"-Abfrage „SubTA Weighting Base 2B" berechnete das Durchschnittsgewicht pro Substanz bei der Gewichtung mit einfacher Durchschnittsbildung

Inputtabelle für diese Abfrage war die Intermediärtablelle der Abfrage „Sub TA Weighting Base 2" (vergleiche Abb. 10.25), die für jede Substanz („Generic name") pro Indikationsgruppe („Sub TA name") und Signalweg („Pathway") die Summe aller gewichteten Indikationen („SumOfWeigtedInd") sowie die absolute Anzahl der Indikationen („SumOfIndCount") berechnet hatte.

Die oben dargestellte Abfrage „SubTA Weighting Base 2B" berechnete das durchschnittliche Gewicht einer Substanz, indem die Summe der gewichteten Indikationen durch die absolute Anzahl geteilt wurde ([SumOfWeigtedInd]/[SumOfIndCount]). Dieses durchschnittliche Gewicht („AvgInd") wurde in eine neue Intermediärtabelle geschrieben.

T222J_SubTAs_ Weighting_Base2B
TA name
Sub TA name
Pathway
Target ID
Generic Name
AvgInd

Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	SumOfAvgInd1: AvgInd
Table:	T222J_SubTAs_ Weighting_Base2B
Total:	Sum
Sort:	
Append to:	SumAvgInd
Criteria:	

Abb. 10.30: „Append"-Abfrage „SubTA Weighting Base 2C" berechnete das Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und einer Indikationsgruppe („Sub TA name"). Die Berechnung des Gesamtgewichts erfolgte über eine Summenfunktion „Sum", die pro Zielmolekül alle Durchschnitts-Gewichte der Substanzen („AvgInd") addierte. Das Gewicht pro Zielmolekül wurde als „SumAvgInd" bezeichnet. Der Name des therapeutischen Gebiets („TA name"), zu dem die Indikationsgruppe gehört, wurde nur zu Informationszwecken in die Tabelle integriert und war nicht für die Berechnung notwendig.

10.1.5.3 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung

T222J_SubTAs_ Weighting_Base2B
TA name
Sub TA name
Pathway
Target ID
Generic Name
AvgInd

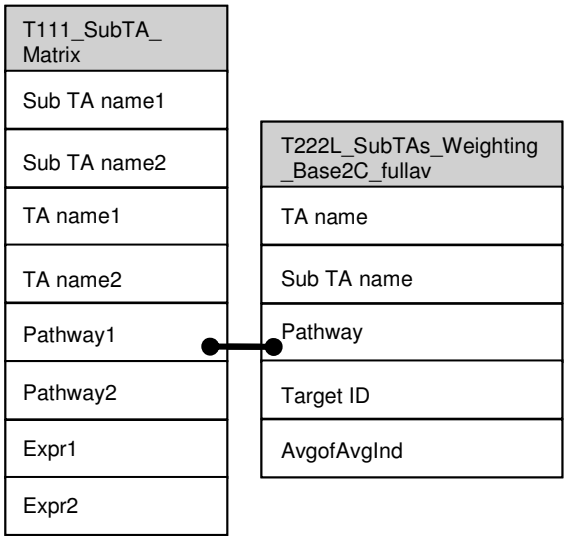
Field:	TA name	Sub TA name	Pathway	Target ID
Table:	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B	T222J_SubTAs_ Weighting_Base2B
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending	Ascending	Ascending	Ascending
Append to:	TA name	Sub TA name	Pathway	Target ID
Criteria:				

Field:	AvgInd
Table:	T222J_SubTAs_ Weighting_Base2B
Total:	Avg
Sort:	
Append to:	AvgofAvgInd
Criteria:	

Abb. 10.31: „Append"-Abfrage „SubTA Weighting Base 2C fullav" berechnete das Gesamtgewicht pro Zielmolekül

Aufbauend auf der Intermediärtabelle der vorhergehenden Abfrage berechnete diese Abfrage ein Gesamtgewicht für ein Zielmolekül („Target ID") innerhalb eines Signalwegs („Pathway") und einer Indikationsgruppe („Sub TA name").
Mit Hilfe einer Durchschnittsfunktion „Avg" wurde dazu pro Zielmolekül der Durchschnitt aller Gewichte der Substanzen („AvgInd") gebildet. Dieses Durchschnitts-Gewicht pro Zielmolekül wurde als „AvgofAvgInd" bezeichnet.

10.1.5.4 Abfragen zur Gewichtung mit doppelter Durchschnittsbildung und Mulitplikator



Field:	Sub TA name1	Sub TA name2	Pathway1	Target ID
Table:	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix	T222L_SubTAs_Weighting_Base2C_fullav
Total:				
Sort:	Ascending	Ascending	Ascending	
Show:	✓	✓	✓	✓
Criteria:				

Field:	AvgofAvgInd	Expr3: IIf(T111_SubTA_Matrix![Sub TA name1] =T222L_SubTAs_Weighting_Base2C_fullav![Sub TA name] Or T111_SubTA_Matrix![Sub TA name2] =T222L_SubTAs_Weighting_Base2C_fullav![Sub TA name];1;0)	Expr4: IIf(T111_SubTA_Matrix![TA name1] <>T111_SubTA_Matrix![TA name2];1;0)	TA name1
Table:	T222L_SubTAs_Weighting_Base2C_fullav			T111_SubTA_Matrix
Total:				
Sort:				
Show:	✓	✓	✓	✓
Criteria:		1	1	[Enter TA name1 here]

Field:	TA name2
Table:	T111_SubTA_Matrix
Total:	
Sort:	
Show:	✓
Criteria:	[Enter TA name2 here]

Abb. 10.32: Bildunterschrift auf der nächsten Seite

Abb. 10.32: „Select“-Abfrage „SubTA Weighting Base 3A fullav“ listete alle Zielmoleküle und ihre Gewichte für die Gewichtung mit doppelter Durchschnittbildung und Multiplikator

Diese „Select“-Abfrage verwendete als Inputtabelle die Intermediärtabelle der Abfrage „SubTA Weighting Base 2C fullav“, die bei Gewichtung mit doppelter Durchschnittsbildung das Durchschnitts-Gewicht der Zielmoleküle berechnet hatte („AvgofAvgInd“). Diese wurde über eine Beziehung mit der Vernetzungsmatrix der Indikationsgruppen („T111_SubTA_Matrix“) verbunden. Die Vernetzungsmatrix listet die Signalwege, die zwei Indikationsgruppen verbinden.

Mit Hilfe der Abfrage wurde nun eine temporäre Liste erstellt, in der für jedes Paar von Indikationsgruppen und jeden verbindenden Signalweg alle Zielmoleküle („Target ID“) mit ihren Durchschnitts-Gewichten („AvgofAvgInd“) aufgelistet wurden, die für eines der beiden therapeutischen Gebiete relevant sind („Expr3“). „Expr4“ verhinderte, dass gleiche Zeilen mehrfach angezeigt werden. In den Felder „TA name1“ und „TA name2“ wurde in der „Criteria“-Zeile eine Parameterfunktion ([Enter TA name1/2 here]) eingetragen. Diese Funktion führte zum Öffnen eines Fensters, in das die Namen der zu untersuchenden therapeutischen Gebiete eingetragen werden mussten. Erst nach Eingabe dieser Einschränkung konnte die Berechnung der Abfrage stattfinden.

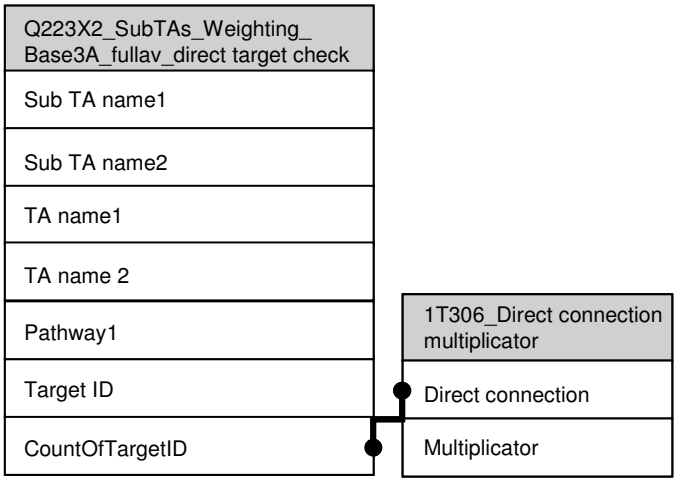
Q223X1_SubTAs_Weighting _Base3A_fullav_multiply
Sub TA name1
Sb TA name2
Pathway1
Target ID
AvgofAvgInd
Expr3
Expr4
TA name1
TA name2

Field:	Sub TA name1	Sub TA name2	TA name1	TA name2
Table:	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Pathway1	Target ID	Target ID
Table:	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply
Total:	Group By	Group By	Count
Sort:			
Show:	✓	✓	✓
Criteria:			

Abb. 10.33: „Select"-Abfrage „SubTA Weighting Base 3A fullav direct target check" identifizierte alle direkt indikationsübergreifenden Zielmoleküle

Diese „Select"-Abfrage verwendete das Ergebnis der vorhergehenden Abfrage, die für jede Kombination von Indikationsgruppen und verbindendem Signalweg die relevanten Zielmoleküle mit ihrem Durchschnittsgewicht aufgelistet hatte, als Input.
Durch eine Zählfunktion „Count" wurde gezählt, wie oft ein Zielmolekül („Target ID") für eine solche Kombination in der Liste erwähnt wurde. Dazu wurden die Zielmoleküleinträge gruppiert („Group by" von „Target ID") und die Anzahl der Zeilen durch Zählen des Feldes „Target ID" ermittelt. Dabei wurde jedem Zielmolekül, das für beide Indikationsgruppen von Bedeutung ist, die Zahl 2 zugeordnet und jedem Zielmolekül, das nur in einem der beiden Indikationsgruppen aktiv ist eine 1.



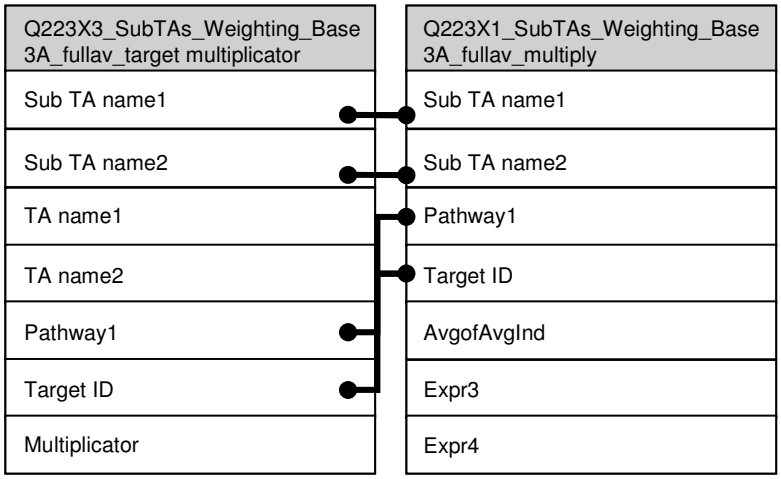
Field:	Sub TA name1	Sub TA name2	TA name1	TA name2
Table:	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Pathway1	Target ID	Target ID
Table:	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check	1T306_Direct connection multiplicator
Total:	Group By	Group By	Group By
Sort:			
Show:	✓	✓	✓
Criteria:			

Abb. 10.34: „Select"-Abfrage „SubTAs Weighting Base 3A fullav direct target multiplicator ordnete den Zielmolekülen den Multiplikator zu

Inputtabellen für diese Abfrage waren das Ergebnis der vorhergehenden „Select"-Abfrage, die durch Zuordnung einer 1 bzw. 2 im Feld „CountofTargetID" zwischen direkten und indirekten Verbindungen unterschieden hatte sowie die Inputtabelle 1T306, in der der Werte des Multiplikators gespeichert ist.

Die oben gezeigte Abfrage ordnete jedem Zielmolekül, das an einer Signalwegverbindung von zwei therapeutischen Gebieten teilhat, entsprechend seines „CountOfTargetID"-Werts einen Multiplikator zu („Multiplicator").



Field:	Sub TA name1	Sub TA name2	TA name1	TA name2
Table:	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X3_SubTAs_Weighting_Base3A_fullav_target multiplicator	Q223X3_SubTAs_Weighting_Base3A_fullav_target multiplicator
Total:				
Sort:	Ascending	Ascending		
Show:	✓	✓	✓	✓
Criteria:				

Field:	Pathway1	Target ID	MulAvgofAvgInd: [Multiplicator]*[AvgofAvgInd]
Table:	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	Q223X1_SubTAs_Weighting_Base3A_fullav_multiply	
Total:			
Sort:	Ascending		
Show:	✓	✓	✓
Criteria:			

Abb. 10.35: „Select"-Abfrage „SubTAs Weighting Base 3A2 fullav" multiplizierte die durchschnittlichen Zielmolekülgewichte mit dem Multiplikator

Inputtabellen für diese Abfrage war das Ergebnis der ersten Gewichtungsabfrage mit Multiplikator (vergleiche Abb. 10.32) sowie das Ergebnis des ersten Schritts der Multiplikator-Anwendung (vergleiche Abb. 10.34). Die gleichartigen Felder der beiden Tabellen waren durch Beziehungen verbunden.

Die Abfrage berechnete das Produkt aus dem durchschnittlichen Gewicht des Zielmoleküls („AvgofAvgInd") und des Multiplikators („Multiplicator"). Das Endprodukt des multiplizierten Gewichts wurde als „MulAvgofAvgInd" bezeichnet.

Q223X4_SubTAs_Weighting_Base 3A_fullav_multiply
Sub TA name1
Sub TA name2
TA name1
TA name2
Pathway1
Target ID
MulAvgofAvgInd

Field:	Sub TA name1	Sub TA name2	Pathway1	SumofPWay: MulAvgofAvgInd
Table:	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply
Total:	Group By	Group By	Group By	Sum
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	TA name1	TA name2
Table:	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply	Q223X4_SubTAs_Weighting_Base3A2_fullav_multiply
Total:	Group By	Group By
Sort:		
Show:	✓	✓
Criteria:		

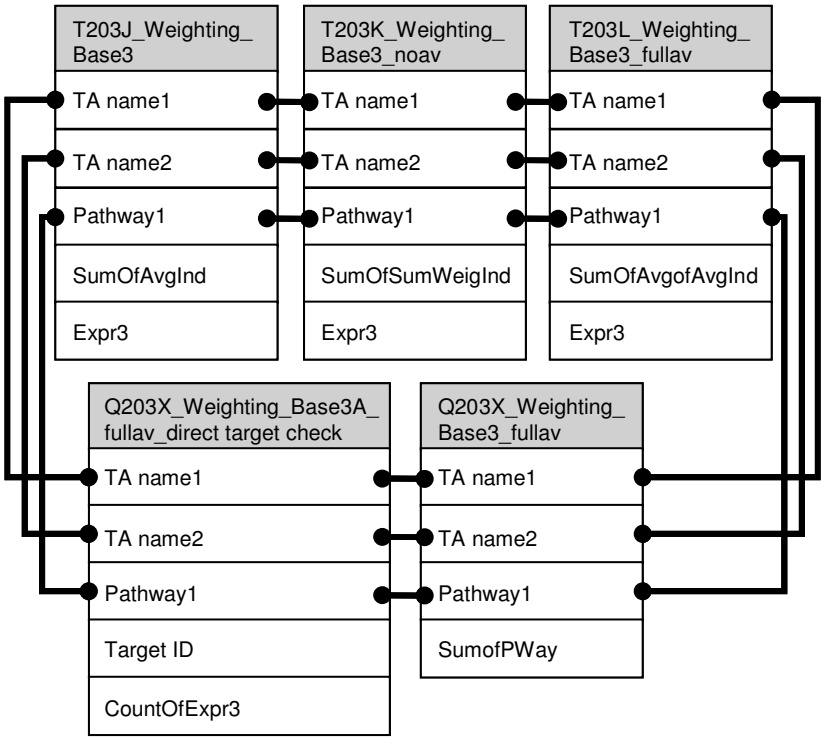
Abb. 10.36: „Select"-Abfrage „SubTAs Weighting Base 3 fullav" berechnete ein Gesamtgewicht pro Signalweg

Die Anfrage verwendete das Ergebnis der vorhergehenden Abfrage, die ein Gesamtgewicht pro Zielmolekül unter Berücksichtigung des Multiplikators berechnet hatte.

Durch eine Summenfunktion „Sum" wurden die einzelnen Gewichte der Zielmoleküle addiert und zu einem Gesamtgewicht eines Signalwegs („SumofPWay") in der Verbindung von zwei Indikationsgruppen („Sub TA name1" und „Sub TA name2") zusammengefügt. Die Felder mit den Namen der therapeutischen Gebiete („TA name1" und „TA name2"), zu denen die Indikationsgruppen gehören, dienten lediglich der Information.

10.1.6 Abfragen zu Signalwegranglisten

10.1.6.1 Abfrage für die Signalwegranglisten für Verbindungen zwischen therapeutischen Gebieten



Field:	TA name1	TA name2	Pathway1	SPway Score weighted: SumOfAvgInd
Table:	T203J_Weighting_ Base3	T203J_Weighting_ Base3	T203J_Weighting_ Base3	T203J_Weighting_ Base3
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:	[Enter TA name1 here]	[Enter TA name2 here]		

Field:	SPway Score fullav: SumOfAvgofAvgInd	SPway Score noav: SumOfSumWeigInd	SPway Score fullav mutliplicator: SumofPWay	Count of targets: CountOfExpr3
Table:	T203L_Weighting_ Base3_fullav	T203K_Weighting_ Base3_noav	Q203X_Weighting_ Base3_fullav	Q203X_Weighting_ Base3A_fullav_direct target check
Total:	Group By	Group By	Group By	Count
Sort:			Descending	
Show:	✓	✓	✓	✓
Criteria:				

Fortsetzung der Tabelle und Bildunterschrift auf der nächsten Seite

Field:	Indirect targets: Sum(IIf([CountOfExpr3] =1;1;0))	Direct targets: Sum(IIf([CountOfExpr3] =2;1;0))
Table:		
Total:	Expression	Expression
Sort:		
Show:	✓	✓
Criteria:		

Abb. 10.37: „Select"-Abfrage erstellte eine Signalwegrangliste für die Verbindung von zwei therapeutischen Gebieten

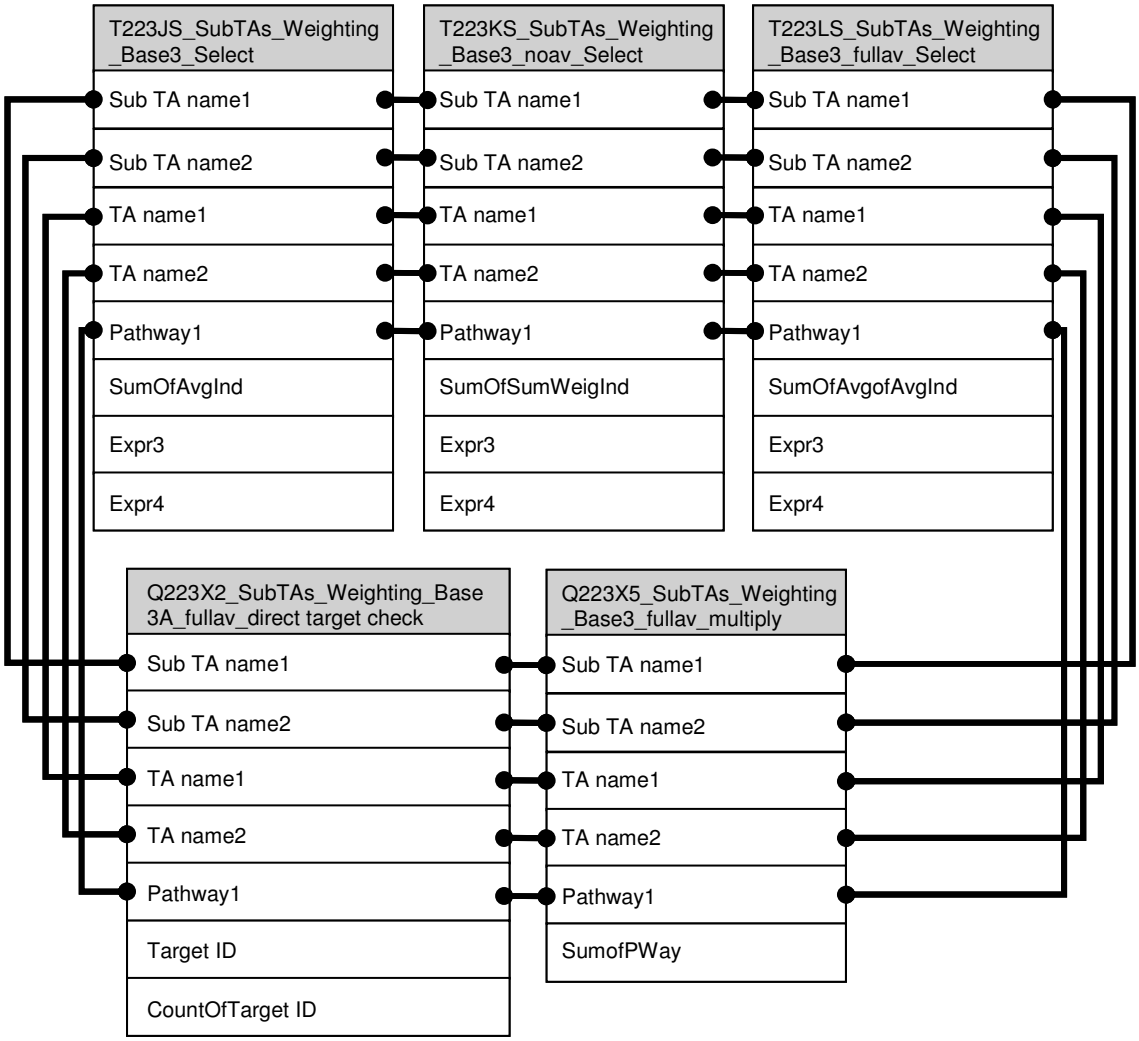
Die oben gezeigte Abfrage verwendete als Inputtabellen die Intermediärtabellen „Weighting Base 3" aller Gewichtungstypen (ohne Durchschnitt, mit einfachem Durchschnitt, mit doppelten Durchschnitt und mit doppelten Durchschnitt und Multiplikator). In diesen Tabellen war jeweils die Verbindungsstärke eines einzelnen Signalwegs berechnet worden, die nun für die Rangliste verwendet wurde. Zusätzlich zu diesen vier Tabellen wurde die Tabelle „Q203X_Weighting_Base3A_fullav_direct target check" benötigt. Die gleichartigen Felder „TA name1", „TA name2" und „Pathway 1" waren durch Beziehungen untereinander verbunden.

Die tabellarische Übersicht zeigt die Struktur der Rangliste. Für die Verbindung von zwei therapeutischen Gebieten wurden in der Spalte „Pathway1" die verbindenden Signalwege aufgelistet.

In den folgenden Spalten folgten die Gewichte eines jeden Signalwegs entsprechend der Art der Gewichtung („SPway Score weighted" = Gewicht mit einfacher Durchschnittsbildung, „SPway Score fullav" = Gewicht mit doppelter Durchschnittsbildung, „SPway Score noav" = Gewicht ohne Durchschnitt, „ SPway Score fullav mutliplicator" = Gewicht mit doppelten Durchschnitt und Multiplikator).

In der Spalte „Count of targets" wurde durch eine Zählfunktion die Gesamtanzahl der verbindenden Zielmoleküle gezeigt; die Spalten „Indirect targets" und „Direct target" zählten die indirekten und direkten Verbindungsmoleküle. Die Zuordnung zu diesen Kategorien erfolgte durch selektive Summenbildung. Das bedeutet, dass die Summe der Zielmoleküle, die in der Inputtabelle Q203X eine bestimmte Ausprägung des Feldes „CountofExpr3" (indirekt: „CountofExpr3" = 1, direkt: „CountofExpr3" = 2) haben, gebildet wurde.

10.1.6.2 Abfrage für die Signalwegranglisten für Verbindungen zwischen Indikationsgruppen



Field:	Sub TA name1	Sub TA name2	Pathway1	SPwayScore weighted: SumOfAvgInd
Table:	T223JS_SubTAs_Weighting_Base3_Select	T223JS_SubTAs_Weighting_Base3_Select	T223JS_SubTAs_Weighting_Base3_Select	T223JS_SubTAs_Weighting_Base3_Select
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:	[Enter SubTA name1, refer to TA1]	[Enter SubTA name2, refer to TA2]		

Fortsetzung der Tabelle und Bildunterschrift auf der nächsten Seite

Field:	SPway Score fullav: SumOfAvgofAvgInd	SPway Score noav: SumOfSumWeigInd	SPway Score fullav multiply: SumofPWay	Count of Targets: CountOfTarget ID
Table:	T223LS_SubTAs_Weighting_Base3_fullav_Select	T223KS_SubTAs_Weighting_Base3_noav_Select	Q223X5_SubTAs_Weighting_Base3_fullav_multiply	Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check
Total:	Group By	Group By	Group By	Count
Sort:			Descending	
Show:	✓	✓	✓	✓
Criteria:				

Field:	Indirect targets: Sum(IIf([CountOfTarget ID]=1;1;0))	Direct targets: Sum(IIf([CountOfTarget ID]=2;1;0))
Table:		
Total:	Expression	Expression
Sort:		
Show:	✓	✓
Criteria:		

Abb. 10.38: „Select"-Abfrage erstellte eine Signalwegrangliste für die Verbindung von zwei Indikationsgruppen

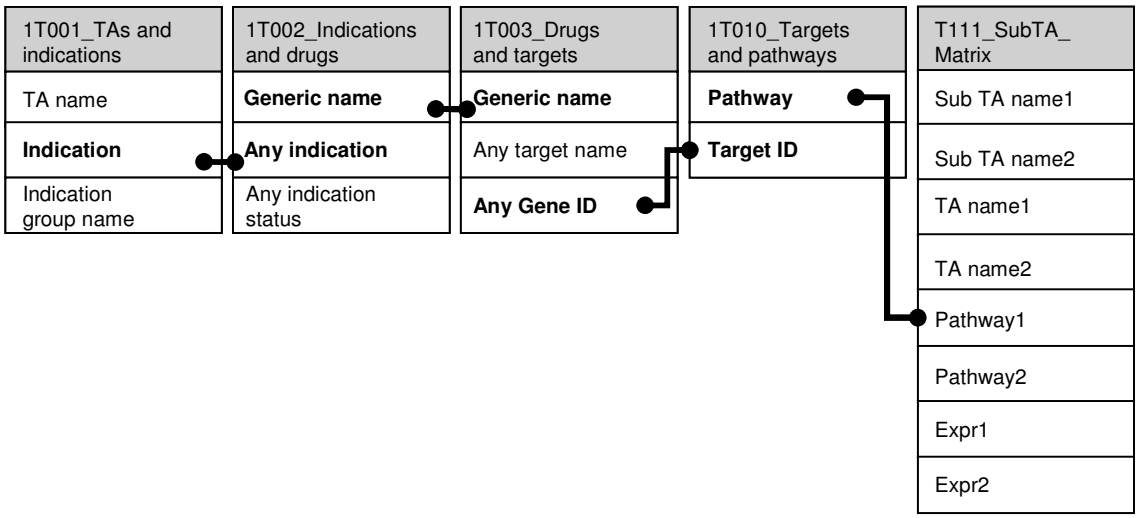
Die oben gezeigte Abfrage verwendete als Inputtabellen die Intermediärtabellen „SubTAs Weighting Base 3" aller Gewichtungstypen (ohne Durchschnitt, mit einfachem Durchschnitt, mit doppelten Durchschnitt und mit doppelten Durchschnitt und Multiplikator). In diesen Tabellen war jeweils die Verbindungsstärke eines einzelnen Signalwegs berechnet worden, die jetzt für die Rangliste verwendet wurde. Zusätzlich zu diesen vier Tabellen wurde die Tabelle „Q223X2_SubTAs_Weighting_Base3A_fullav_direct target check" benötigt. Die gleichartigen Felder „TA name1", „TA name2", „SubTA name1", „SubTA name2" und „Pathway 1" waren durch Beziehungen untereinander verbunden.

Die tabellarische Übersicht zeigt die Struktur der Rangliste. Für die Verbindung von Indikationsgruppen wurden in der Spalte „Pathway1" die verbindenden Signalwege aufgelistet.

In den folgenden Spalten folgten die Gewichte eines jeden Signalwegs entsprechend der Art der Gewichtung („SPwayScore weighted" = Gewicht mit einfacher Durchschnittsbildung, „SPway Score fullav" = Gewicht mit doppelter Durchschnittsbildung, „SPway Score noav" = Gewicht ohne Durchschnitt, „SPway Score fullav multiply" = Gewicht mit doppelten Durchschnitt und Multiplikator).

In der Spalte „Count of targets" wurde durch eine Zählfunktion die Gesamtanzahl der verbindenden Zielmoleküle gezeigt; die Spalten „Indirect targets" und „Direct target" zählten die indirekten und direkten Verbindungsmoleküle. Die Zuordnung zu diesen Kategorien erfolgte durch selektive Summenbildung. Das bedeutet, dass die Summe der Zielmoleküle, die in der Inputtabelle Q223X eine bestimmte Ausprägung des Feldes „CountofTargetID" (indirekt: „ CountofTargetID „ = 1, direkt: „ CountofTargetID „ = 2) hatten, gebildet wurden.

10.1.7 Abfragen zur Identifikation von gemeinsamen Zielmolekülen



Field:	Sub TA name1	Sub TA name2	Pathway1	Target ID
Table:	T111_SubTA_Matrix	T111_SubTA_Matrix	T111_SubTA_Matrix	1T010_Targets and pathways
Total:	Group By	Group By	Group By	Group By
Sort:				
Append to:	Sub TA name1	Sub TA name2	Pathway1	Target ID
Criteria:	[Enter Sub TA name1 here]	[Enter Sub TA name2 here]	[Enter pathway here]	

Field:	Any Target Name	Expr3: IIf([1T001_TAs and indications]![Indication group name]=T111_SubTA_Matrix![Sub TA name1] Or [1T001_TAs and indications]![Indication group name]=T111_SubTA_Matrix![Sub TA name2];1;0)	Indication group name
Table:	1T003_Drugs and targets		1T001_TAs and indications
Total:	Group By		Group By
Sort:			
Append to:	Any Target Name	Expr3	Indication group name
Criteria:		1	

Abb. 10.39: Bildunterschrift auf der nächsten Seite

Abb. 10.39: „Append"-Abfrage listete aller Zielmoleküle innerhalb eines ausgewählten Signalwegs, die für mindestens eine der beiden ausgewählten Indikationsgruppen bereits pharmazeutisch genutzt werden

Die oben dargestellte Abfrage nutzte die Rohdatentabellen „TAs and indications", „Indications and drugs", „Drugs and targets" und „Targets and pathways" und die Vernetzungsmatrix der Indikationsgruppen zur Erstellung einer Liste von allen Zielmolekülen, die Teil des untersuchten Signalwegs sind und in mindestens einer der beiden ausgewählten Indikationsgruppen pharmazeutisch aktiv sind. Gleichartige Felder waren durch Beziehungen miteinander verbunden.

Mit Hilfe einer Parameterfunktion in der „Criteria"-Zeile der Spalten „Sub TA name1", „Sub TA name2" und „Pathway" wurden in der oben darstellten Abfrage vom Nutzer der Datenbank zwei Indikationsgruppen sowie ein verbindender Signalweg für die Analyse ausgewählt. Die Abfrage erstellte eine Liste der Zielmoleküle („Target ID" und „Any target name"), die von einer Substanz adressiert werden, die über eine Indikation in einer der beiden Indikationsgruppen verfügt. Diese Auswahl wurde durch die logische Bedingung „Expr3" gesteuert. War ein Zielmolekül über Substanzen für beide Indikationsgruppen relevant, wird es zweifach aufgeführt.

T401_CT_Step1
Sub TA name1
Sub TA name2
Pathway1
Target ID
Any Target Name
Expr3
Indication group name

Field:	Sub TA name1	Sub TA name2	Pathway1	Target ID
Table:	T401_CT_Step1	T401_CT_Step1	T401_CT_Step1	T401_CT_Step1
Total:	Group By	Group By	Group By	Group By
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Any Target Name	Indication group name
Table:	T401_CT_Step1	T401_CT_Step1
Total:	Group By	Count
Sort:		
Show:	✓	✓
Criteria:		

Abb. 10.40: „Select"-Abfrage trennte Moleküle, die bereits heute für beide Indikationsgruppen relevant sind von Zielmolekülen, die erst in einer Indikationsgruppe genutzt werden

Die Abfrage erstellte eine Liste derjenigen Moleküle, die für eine bestimmte Signalwegverbindung von Indikationsgruppen wichtig sind.

Die Abfrage zählte mit einer „Count"-Funktion die Einträge im „Indication group name"-Feld jedes Zielmoleküls („Target ID"/"Any target name"). Ist ein Zielmolekül in beiden ausgewählten Indikationsgruppen aktiv, dann verfügte es in der Inputtabelle über zwei Einträge und wurde mit „2" gekennzeichnet. Ist das Zielmolekül bisher nur für eine Indikation pharmazeutisch genutzt, dann hatte es nur einen Eintrag und wurde mit „1" markiert.

Q402_CT_Step2
Sub TA name1
Sub TA name2
Pathway1
Target ID
Any Target Name
CountOfIndication group name

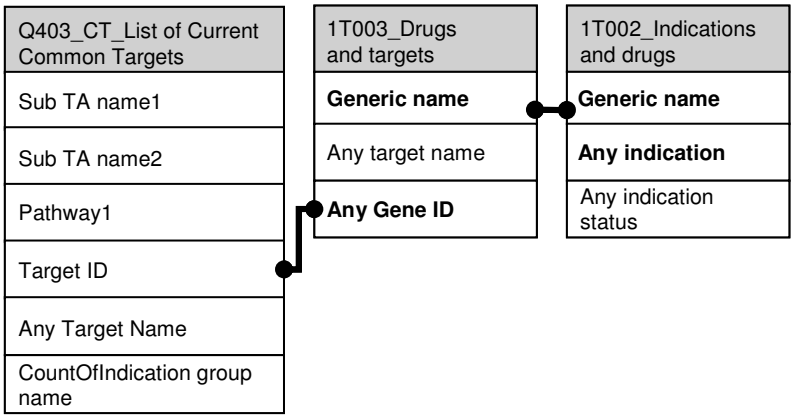
Field:	Sub TA name1	Sub TA name2	Pathway1	Target ID
Table:	Q402_CT_Step2	Q402_CT_Step2	Q402_CT_Step2	Q402_CT_Step2
Total:				
Sort:				
Show:	✓	✓	✓	✓
Criteria:				

Field:	Any Target Name	CountOfIndication group name
Table:	Q402_CT_Step2	Q402_CT_Step2
Total:		
Sort:		
Show:	✓	✓
Criteria:		2

Abb. 10.41: „Select"-Abfrage erstellte Liste von Kandidaten für zukünftig indikationsübergreifende Zielmoleküle

Die Abfrage verwendet das Ergebnis der vorhergehenden Abfrage und erstellt eine Liste, die nur diejenigen Zielmoleküle enthält, die für die Verbindung von zwei ausgewählten Indikationsgruppen durch einen bestimmten Signalweg bereits heute indikationsübergreifend sind.

In den ersten drei Spalten werden die untersuchten Indikationsgruppen („Sub TA name1/2") sowie der verbindende Signalweg („Pathway") gezeigt. In den nächsten Spalten werden dann Gene ID und Name der Zielmoleküle („Target ID"/"Any Target Name") gelistet. Die Eingabe des Werts „2" in das „Criteria"-Feld der Spalte „CountOfIndication group name" stellt dabei sicher, dass nur diejenigen Zielmoleküle genannt werden, die in der vorhergehenden Abfrage (vergleiche Abb. 10.40) durch eine „2" als indikationsübergreifende Zielmoleküle identifiziert wurden.



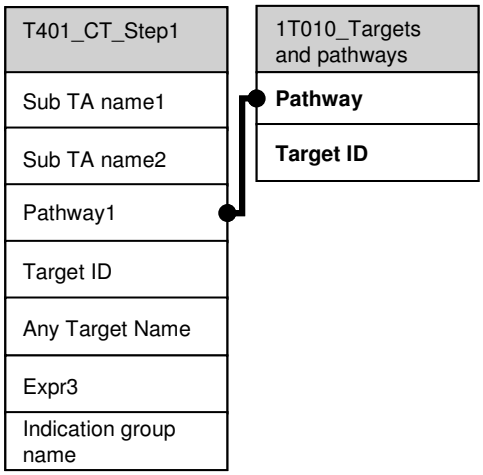
Field:	Target ID	Any Target Name	Generic name	Any Indication
Table:	Q403_CT_List of Current Common Targets	Q403_CT_List of Current Common Targets	1T002_Indications and drugs	1T002_Indications and drugs
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending		Ascending	
Show:	✓	✓	✓	✓
Criteria:				

Field:	Any Indication Status
Table:	1T002_Indications and drugs
Total:	Group By
Sort:	
Show:	✓
Criteria:	

Abb. 10.42: „Select"-Abfrage „Detailed info 1" stellte Informationen über Substanzen zusammen

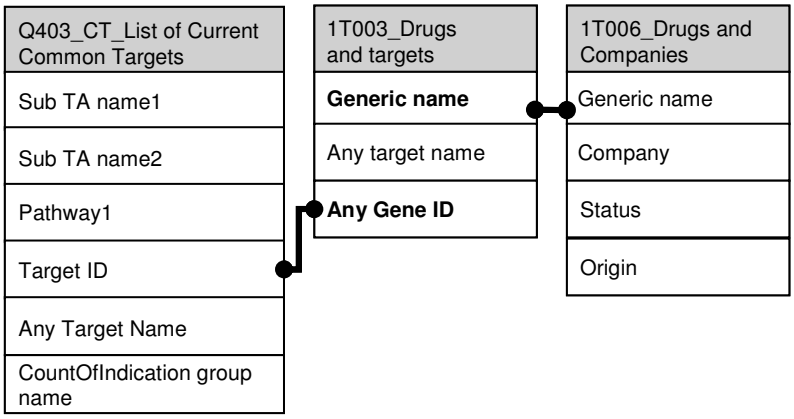
Die Abfrage verknüpfte die Liste der indikationsübergreifenden Zielmoleküle „ Q403_CT_List of Current Common Targets" über die Inputtabelle 1T003_Drugs and targets mit der Inputtabelle 1T002_Indications and drugs.

Pro indikationsübergreifenden Zielmolekül („Target ID"/"Any target name") wurden alle Substanzen („Generic name") mit allen ihren Indikationen („Any indication") und dem entsprechenden Entwicklungsstatus („Any Indication Status") aufgelistet. Im Gegensatz zu allen vorhergehenden Abfragen, wurden jetzt alle Substanzen mit ihren Indikationen aufgelistet und nicht nur diejenigen, die zu bestimmten therapeutischen Gebieten zugeordnet werden konnten.



Field:	Pathway1	TargetID
Table:	T401_CT_Step1	1T010_Targets and pathways
Total:	Group By	Group By
Sort:	Ascending	
Show:	✓	✓
Criteria:		

Abb. 10.43: „Select"-Abfrage „All KEGG targets" listete alle Zielmoleküle, die entsprechend der KEGG Rohdaten Teil eines verbindenden Signalwegs waren
Die Abfrage verknüpfte die Inputtabelle der aufbereiteten KEGG-Rohdaten „1T010_Targets and pathways" mit der Tabelle „T401_CT_Step1" aus der alle Namen von verbindenden Signalwegen gezogen wurden.
Die tabellarische Übersicht zeigt, dass allen Signalwegen, die wichtig für eine Verbindung waren („Pathway") diejenigen Zielmoleküle („Target ID") zugeordnet wurden, die entsprechend der KEGG Rohdaten Teil des Signalwegs sind. Dabei spielte es keine Rolle, ob die Zielmoleküle pharmazeutisch aktiv waren oder nicht.



Field:	Target ID	Any Target Name	Generic name	Company
Table:	Q403_CT_List of Current Common Targets	Q403_CT_List of Current Common Targets	1T006_Drugs and companies	1T006_Drugs and companies
Total:	Group By	Group By	Group By	Group By
Sort:	Ascending		Ascending	Ascending
Show:	✓	✓	✓	✓
Criteria:				

Field:	Origin
Table:	1T006_Drugs and companies
Total:	Group By
Sort:	
Show:	✓
Criteria:	

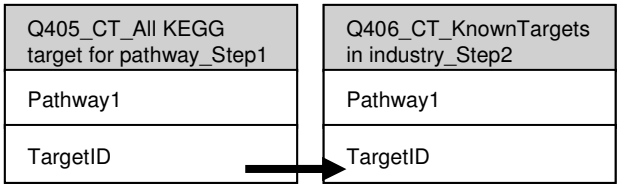
Abb. 10.44: „Select"-Abfrage „Detailed info 2" gab Entwickler und mögliche Lizenznehmer der Substanzen an

Die Abfrage verknüpfte die Liste der indikationsübergreifenden Zielmoleküle „ Q403_CT_List of Current Common Targets" über die Inputtabelle „1T003_Drugs and targets" mit der Inputtabelle „1T006_Drugs and companies".
Pro indikationsübergreifendem Zielmolekül („Target ID"/"Any target name") wurden alle Substanzen („Generic name") aufgelistet. Zu jeder Substanz wurde der Name der Firma, die an der Substanz beteiligt ist, genannt („Company"), sowie außerdem in der nächsten Spalte („Origin") erläutert, ob es sich bei der Firma um den Entwickler („Originator") oder einen Lizenznehmer („Licensee") handelt.

Q402_CT_Step2
Sub TA name1
Sub TA name2
Pathway1
Target ID
Any Target Name
CountOfIndication group name

Field:	Pathway1	TargetID
Table:	Q402_CT_Step2	Q402_CT_Step2
Total:	Group By	Group By
Sort:	Ascending	
Show:	✓	✓
Criteria:		

Abb. 10.45: **„Select"-Abfrage „Known targets in industry" listete alle Zielmoleküle, die bereits für die beiden Indikationsgruppen pharmazeutisch adressiert werden**
Die Abfrage nutzte das Ergebnis der Abfrage aus Abb. 10.40, um eine zweite Liste im selben Format wie das Ergebnis der Abfrage aus Abb. 10.43 zu erstellen, die alle Zielmoleküle enthielt, die für die eine oder beide untersuchten Indikationsgruppen bereits pharmazeutisch adressiert werden. Die Abfrage stellte damit lediglich eine andere Darstellung des Ergebnisses der Abfrage aus Abb. 10.40 dar, war aber nötig, um den folgenden Abfrageschritt zu ermöglichen.
Die tabellarische Übersicht zeigt, dass das Ergebnis der Abfrage aus zwei Spalten bestand, in denen alle Signalwege („Pathway1") aufgelistet waren, die zur Verbindung der Indikationsgruppen beitragen, sowie die Zielmoleküle, die in mindestens einer der beiden Indikationsgruppen adressiert werden („Target ID").



Field:	Pathway1	TargetID	TargetID
Table:	Q405_CT_All KEGG target for pathway_Step1	Q405_CT_All KEGG target for pathway_Step1	Q406_CT_KnownTargets in industry_Step2
Total:			
Sort:	Ascending		
Show:	✓	✓	
Criteria:			Is Null

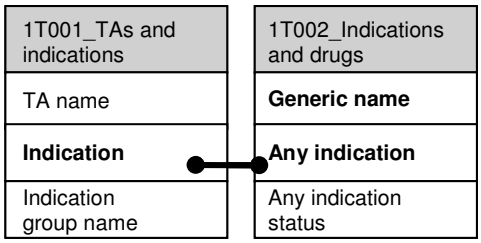
Abb. 10.46: „Select"-Abfrage mit Differenz Bildung entfernte alle bereits bekannten Zielmoleküle aus der Gesamtliste und erstellte damit eine Liste neuer Zielmoleküle

Ausgang der Differenz waren die Tabellen der Abfrage „All KEGG targets" (Abb. 10.43) und „Known targets in industry" (Abb. 10.45). Der Pfeil in der graphischen Darstellung und die letzte Spalte der tabellarischen Darstellung („Target ID" aus Q406 „Is Null") zeigen die Bildung der Differenz zwischen den beiden Tabellen.

Durch diese Operation der relationalen Algebra wurden alle Zeilen (Tupel), die im Ergebnis der Abfrage Q406 enthalten sind, aus dem Ergebnis der Abfrage Q405 entfernt. Dies führte zu einer Residualliste von Zielmolekülen („Target ID"), die zu einem verbindenden Signalweg („Pathway1") zwischen zwei Indikationsgruppen gehören, jedoch noch für keine der beiden Indikationsgruppen adressiert werden. Dies war die Liste der neuen Zielmoleküle.

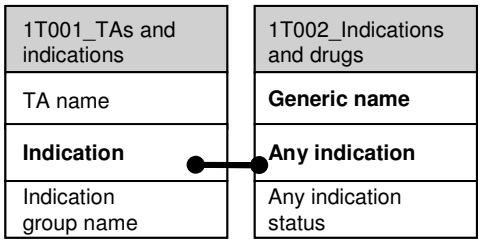
10.1.8 Abfragen zur Charakterisierung der Datenbasis

10.1.8.1 Eigenschaften der therapeutischen Gebiete



Field:	TA name	Generic name
Table:	1T001_TAs and indications	1T002_Indications and drugs
Total:	Group By	Count
Sort:		
Show:	✓	✓
Criteria:		

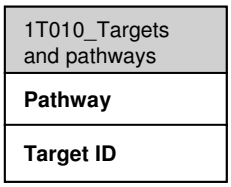
Abb. 10.47: „Select"-Abfrage ermittelte die Anzahl der Forschungsprojekte pro therapeutischem Gebiet
Ausgang der Abfrage waren die beiden Inputtabellen „1T001_TAs and indications" und „1T002_Indications and drugs", die über das gleichartige Feld „Indication"/"Any indication" miteinander verbunden waren.
Die Abfrage zählte mit Hilfe einer „Count"-Funktion die Anzahl der Einträge von Substanzen („Generic name"). Jeder Eintrag einer Substanz entspricht einem Forschungsprojekt in dem entsprechenden therapeutischen Gebiet.



Field:	Any Indication Status	TA name	Any Indication
Table:	1T001_TAs and indications	1T001_TAs and indications	1T003_Indications and drugs
Total:	Group By	Group By	Count
Sort:	Ascending		
Show:	✓	✓	✓
Criteria:			

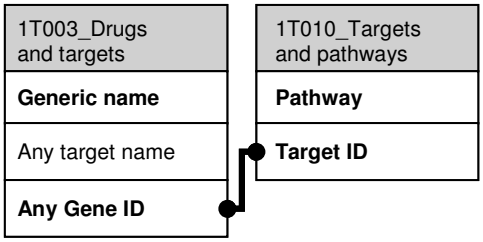
Abb. 10.48: „Select"-Abfrage zeigte die Verteilung von Forschungsprojekte auch unterschiedliche Entwicklungsphasen
Ausgang der Abfrage waren erneut die beiden Inputtabellen „1T001_TAs and indications" und „1T002_Indications and drugs", die über das gleichartige Feld „Indication"/"Any indication" miteinander verbunden waren.
Die tabellarische Übersicht zeigt die Struktur des Abfrageergebnisses: Für jeden Entwicklungsstatus („Any indication status") wurde pro therapeutisches Gebiet („TA name") mit einer Zählfunktion die Anzahl der Indikationseinträge ermittelt („Count" von „AnyIndication"). Jeder Eintrag einer Indikation stellt ein Forschungsprojekt dar, d.h. das Vorkommen der Substanz im Zusammenhang mit einer Indikation in dem therapeutischen Gebiet.

10.1.8.2 Eigenschaften der Signalwege



Field:	Pathway	Target ID
Table:	1T010_Targets and pathways	1T010_Targets and pathways
Total:	Group By	Count
Sort:		Descending
Show:	✓	✓
Criteria:		

Abb. 10.49: „Select"-Abfrage zählte die Anzahl der Zielmoleküle pro Signalweg
Diese Abfrage analysierte die Rohdatentabelle „1T010_Targets and pathways". Für jeden Signalweg („Pathway") wurde über eine Zählfunktion die Anzahl der Zielmoleküle („Count" von „Target ID") ermittelt. Die Sortierfunktion („Descending") sortierte die Ergebnisansicht so, dass die Signalwege mit der größten Zahl von Zielmolekülen als erste aufgelistet wurden.



Field:	Pathway	Any Gene ID
Table:	1T010_Targets and pathways	1T003_Drugs and targets
Total:	Group By	Group By
Sort:	Ascending	
Show:	✓	✓
Criteria:		

Abb. 10.50: „Select"-Abfrage erstellte eine Liste aller pharmazeutisch adressierten Zielmoleküle pro Signalweg
Diese Abfrage verwendete die Rohdatentabellen „1T003_Drugs and targets" und „1T0101_Targets and pathways", die über die gleichartigen Felder „Any Gene ID" und „Target ID" verbunden waren.
Für jeden Signalweg („Pathway") wurden die dazugehörigen Zielmoleküle („Any Gene ID") aufgelistet. Dadurch, dass das Zielmolekül aus der Tabelle „1T003_Drugs and targets" gezogen wurde, konnte gewährleistet werden, dass im Ergebnis nur pharmazeutisch adressierte Zielmoleküle enthalten waren. Zielmoleküle, für die es noch keine Substanz gibt, sind nicht in der aus Pharmaprojects befüllten Rohdatentabelle enthalten. Die Sortierfunktion listete die Signalwege alphabetisch aufsteigend.

Q502B_# of addressed targets per pathway_step1
Pathway
Any Gene ID

Field:	Pathway	Any Gene ID
Table:	Q502B_# of addressed targets per pathway_step1	Q502B_# of addressed targets per pathway_step1
Total:	Group By	Count
Sort:	Ascending	
Show:	✓	✓
Criteria:		

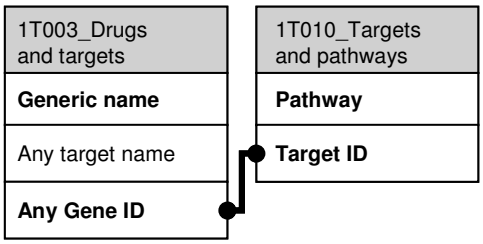
Abb. 10.51: „Select"-Abfrage zählte die Anzahl der pharmazeutisch adressierten Zielmoleküle pro Signalweg
Diese Abfrage baute auf der zuvor erstellten Liste pharmazeutisch adressierter Zielmoleküle pro Signalweg auf (vergleiche Abb. 10.50).
Mit Hilfe einer Zählfunktion („Count") wurde die Anzahl der Zielmoleküle („Any Gene ID") pro Signalweg („Pathway") ermittelt. Die Sortierfunktion listete die Signalwege alphabetisch aufsteigend.

T100_TAs and Pathways
TA name
Pathway

Field:	Pathway	CountOfTA name1: TA name
Table:	T100_TAs and Pathways	T100_TAs and Pathways
Total:	Group By	Count
Sort:		Descending
Show:	✓	✓
Criteria:		

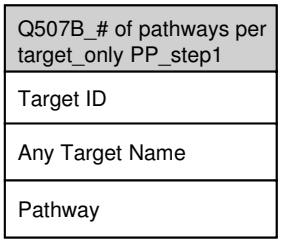
Abb. 10.52: „Select"-Abfrage zählte die Anzahl der therapeutischen Gebiete, in denen ein Signalweg eine Rolle spielt
Diese Abfrage nutzte die Intermediärtabelle „T100_TAs and Pathways", die im Rahmen der Vernetzungsanalyse der therapeutischen Gebiete berechnet worden war und die Zuordnung von Signalwegen zu therapeutischen Gebieten enthält.
Für jeden Signalweg („Pathway") wurde mit Hilfe einer Zählfunktion („Count") die Anzahl der therapeutischen Gebiete ermittelt, in denen er eine Rolle spielt. Diese Abfrage stellt die Umkehrung der Intermediärtabelle T100 dar. Die Sortierfunktion („Descending") ermöglichte, dass die Signalwege, die in den meisten Gebieten von Bedeutung sind, zuerst aufgelistet wurden.

10.1.8.3 Eigenschaften der Zielmoleküle



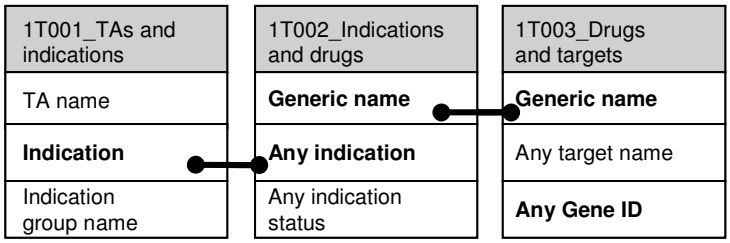
Field:	Target ID	Any Target Name	Pathway
Table:	1T010_Targets and pathways	1T003_Drugs and targets	1T010_Targets and pathways
Total:	Group By	Group By	Group By
Sort:			Descending
Show:	✓	✓	✓
Criteria:			

Abb. 10.53: „Select"-Abfrage ordnete pharmazeutisch aktiven Zielmolekülen Signalwege zu
Grundlage der Abfrage waren die Rohdatentabellen „1T003_Drugs and targets" und „1T010_Targets and pathways", die durch Verbindung der gleichartigen Felder „Any Gene ID" und „Target ID" verknüpft waren.
Die Abfrage listete in den ersten beiden Spalten die pharmazeutisch aktiven Zielmoleküle, d.h. die Zielmoleküle, die in der Tabelle 1T003 gemeinsam mit einer pharmazeutisch aktiven Substanz vermerkt waren. Jedem Zielmolekül wurden dann die Signalwege zugeordnet, in denen es entsprechend der KEGG Rodaten vorkommt Kommt ein Zielmolekül in mehr als einem Signalweg vor, so wurde es in mehreren Zeilen mit den entsprechenden Signalwegen aufgeführt.



Field:	Target ID	Any Target Name	Pathway
Table:	Q507B_# of pathways per target_only PP_step1	Q507B_# of pathways per target_only PP_step1	Q507B_# of pathways per target_only PP_step1
Total:	Group By	Group By	Count
Sort:			Descending
Show:	✓	✓	✓
Criteria:			

Abb. 10.54: „Select"-Abfrage zählte die Anzahl der Signalwege für jedes pharmazeutisch aktive Zielmolekül
Grundlage der Abfrage war die Zuordnung von Signalwegen zu pharmazeutisch aktiven Zielmolekülen aus der vorhergehenden Abfrage (Abb. 10.53).
Mit Hilfe einer Zählfunktion wurde für jedes Zielmolekül („Target ID"/"Any Target Name") die Zahl der Signalwege bestimmt, zu denen es in der vorhergehenden Abfrage zugeordnet worden war („Count" von „Pathways"). Die Sortierfunktion „Descending" sortierte das Ergebnis so, dass Zielmoleküle mit vielen Signalwegen zuerst angezeigt wurden.



Field:	TA name	Any Gene ID	Any Target Name	Generic name
Table:	1T001_TAs and indications	1T003_Drugs and targets	1T003_Drugs and targets	1T002_Indications and drugs
Total:	Group By	Group By	Group By	Count
Sort:	Ascending	Ascending	Ascending	
Show:	✓	✓		✓
Criteria:				

Abb. 10.55: „Select"-Abfrage erstellte pro therapeutisches Gebiet eine Liste aller Substanzen pro Zielmolekül
Grundlage der Abfrage sind die drei Rohdatentabellen 1T001, 1T002 und 1T003. Die drei Tabellen standen über gleichartige Felder miteinander in Beziehung.
Pro therapeutisches Gebiet („TA name") und Zielmolekül („Any Gene ID"/"Any Target Name") wurde mit Hilfe einer Zählfunktion die Zahl der Substanzen bestimmt („Generic name"). Durch Verknüpfung der Rohdaten wurden dabei nur die Substanzen gezählt, die über eine Indikation für das entsprechende therapeutische Gebiet eine Rolle spielen. Die Sortierfunktion listete die ersten drei Spalten alphabetisch aufsteigend.

Q505_# of drugs per target_step1
TA name
Any Gene ID
Any Target Name
CountOfGenericName

Field:	TA name	CountOfGenericName
Table:	Q505_# of drugs per target_step1	Q505_# of drugs per target_step1
Total:	Group By	Avg
Sort:		
Show:	✓	✓
Criteria:		

Abb. 10.56: „Select"-Abfrage berechnete pro therapeutisches Gebiet die durchschnittliche Anzahl von Substanzen pro Zielmolekül

Grundlage der Abfrage war das Ergebnis des ersten Berechnungsschritts (siehe Abb. 10.55) in dem die Anzahl der Substanzen für jedes Zielmolekül innerhalb eines therapeutischen Gebiets bestimmt worden war.
Mit Hilfe einer Durchschnittsfunktion („Avg") wurde in zweiten Berechnungsschritt pro therapeutischem Gebiet („TA name") der Durchschnitt über die Zahl der Substanzen pro Zielmolekül („CountOfGenericName") gebildet.

10.1.8.4 Eigenschaften der Substanzen

1T001_TAs and indications
TA name
Indication
Indication group name

1T002_Indications and drugs
Generic name
Any indication
Any indication status

Field:	TA name	Generic Name	Any indication
Table:	1T001_TAs and indications	1T002_Indications and drugs	1T002_Indications and drugs
Total:	Group By	Group By	Count
Sort:	Ascending	Ascending	
Show:	✓	✓	✓
Criteria:			

Abb. 10.57: „Select"-Abfrage erstellte pro therapeutisches Gebiet eine Liste aller Substanzen und zählte die Indikationen pro Substanz

Grundlage der Abfrage waren die beiden Rohdatentabellen 1T001 und 1T002. Die beiden Tabellen standen über die Felder „Indication"/"Any indication" miteinander in Beziehung.
Pro therapeutisches Gebiet („TA name") wurden alle Substanzen aufgelistet („Generic name") und mit Hilfe einer Zählfunktion die Anzahl der Indikationen ermittelt („Count" von „Any indication"). Durch Verknüpfung der Rohdaten wurden dabei nur die Indikationen und Substanzen berücksichtigt, die in das entsprechende therapeutische Gebiet gehören. Die Sortierfunktion listete die ersten beiden Spalten alphabetisch aufsteigend.

Q503_# of indications per generic name_step1
TA name
Generic Name
CountOfAny Indication

Field:	TA name	CountOfAny Indication
Table:	Q503_# of indications per generic name_step1	Q503_# of indications per generic name_step1
Total:	Group By	Avg
Sort:		
Show:	✓	✓
Criteria:		

Abb. 10.58: „Select"-Abfrage berechnete pro therapeutisches Gebiet die durchschnittliche Anzahl von Indikationen pro Substanz

Grundlage der Abfrage war das Ergebnis des ersten Berechnungsschritts (vergleiche Abb. 10.57), in dem die Anzahl der Indikationen für jede Substanz innerhalb eines therapeutischen Gebiets bestimmt worden war.

Mit Hilfe einer Durchschnittsfunktion („Avg") wurde in zweiten Berechnungsschritt pro therapeutischem Gebiet („TA name") der Durchschnitt über die Zahl der Indikationen pro Substanz („CountOfAny Indication") gebildet.

10.2 Identifikation und Priorisierung von Signalwegen in der Verbindung von Indikationsgruppen (weitere Analysen)

Im Folgenden werden Ergebnisse der Identifikation und Priorisierung von Signalwegen in der Verbindung von weiteren Indikationsgruppen dargestellt. Dieser Anhang ist damit eine Ergänzung zum Abschnitt 7.4.3.

10.2.1.1 Signalwegpriorisierung zwischen Brustkrebs und rheumatoider Arthritis

Zum Vergleich mit dem in der Abb. 7.33 untersuchten „Hot-Spot“ der Verbindungsmatrix der therapeutischen Gebiete Onkologie und Bewegungsapparat wurde ein weiterer Verbindungspunkt derselben Matrix im Detail untersucht. Die Verbindung zwischen Brustkrebs und rheumatoider Arthritis unterscheidet sich von der zuvor beschriebenen nur durch einen der beiden Verbindungspartner und verfügt ebenfalls über 40 gemeinsame Signalwege. Bei Gewichtung mit einfacher Durchschnittsbildung weist die Verbindung von Brustkrebs mit rheumatoider Arthritis die maximale Verbindungsstärke in der Matrix auf.

Abb. 10.59 zeigt die Rangliste der Signalwege. Im Vergleich zu Abb. 7.33 lassen sich nur geringe Veränderungen der Reihenfolge der Signalwege feststellen. Dies ist nicht verwunderlich, da es sich bei der hier untersuchten Verbindung ebenfalls um die Verbindung einer Krebserkrankung mit rheumatoider Arthritis handelt – lediglich die Art der Tumoren unterscheidet sich. Auch in der hier gezeigten Untersuchung führen die Cytokine mit ihren Rezeptoren („Cytokine-cytokine receptor interaction“) die Rangliste an. Im Vergleich zu der Rangliste der Verbindung von Lungenkrebs mit rheumatoider Arthritis haben jedoch die Signalwege „Focal adhesion“ und „Toll-like receptor signaling pathway“ eine etwas stärkere Bedeutung.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Cytokine-cytokine receptor interaction**	**11**	**10**	**9**	**40**	**20**
2.	**Focal adhesion**	**6**	**7**	**3**	**26**	**15**
3.	**MAPK signaling pathway**	**5**	**6**	**7**	**18**	**10**
4.	**Jak-STAT signaling pathway**	**5**	**3**	**3**	**22**	**16**
5.	**Toll-like receptor signaling pathway**	**5**	**4**	**6**	**18**	**24**
6.	**Hematopoietic cell lineage**	**4**	**4**	**5**	**16**	**24**
7.	**Natural killer cell mediated cytotoxicity**	**4**	**3**	**6**	**13**	**14**
8.	**ErbB signaling pathway**	**4**	**4**	**4**	**14**	**21**
9.	**Apoptosis**	**4**	**4**	**5**	**13**	**21**
10.	**Adipocytokine signaling pathway**	**3**	**3**	**5**	**8**	**18**
11.	T cell receptor signaling pathway	3	3	4	9	14
12.	Neuroactive ligand-receptor interaction	3	2	0	15	6
13.	Regulation of actin cytoskeleton	3	3	2	11	6
14.	VEGF signaling pathway	3	7	2	11	19
15.	Gap junction	3	4	2	10	12
16.	p53	3	1	2	13	22
17.	Fc epsilon RI signaling pathway	2	3	4	7	15
18.	GnRH signaling pathway	2	3	3	7	10
19.	Calcium signaling pathway	2	4	0	13	7
20.	Cell adhesion molecules (CAMs)	2	1	1	10	8
21.	Insulin signaling pathway	2	1	3	5	6
22.	B cell receptor signaling pathway	2	1	2	6	13
23.	Leukocyte transendothelial migration	2	2	1	9	9
24.	Long-term depression	2	1	3	5	11
25.	Melanogenesis	2	1	3	4	7
26.	mTOR signaling pathway	2	1	2	5	15
27.	Cell cycle	2	1	2	7	8
28.	Axon guidance	1	1	1	5	5
29.	ECM-receptor interaction	1	1	0	9	10
30.	Wnt signaling pathway	1	1	1	7	5
31.	Long-term potentiation	1	1	2	3	7
32.	Tight junction	1	1	0	6	5
33.	PPAR signaling pathway	1	0	0	5	7
34.	Adherens junction	1	3	1	4	7
35.	Dorso-ventral axis formation	1	3	1	4	18
36.	TGF-beta signaling pathway	1	1	1	4	6
37.	Notch signaling pathway	1	0	3	0	6
38.	Phosphatidylinositol signaling system	1	0	0	4	5
39.	Antigen processing and presentation	1	0	1	2	4
40.	Complement and coagulation cascades	1	0	0	3	4

Abb. 10.59: Signalwegrangliste für die Verbindung von Brustkrebs mit rheumatoider Arthritis

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Brustkrebs und rheumatoide Arthritis wichtig sind.

(*Fortsetzung der Bildunterschrift auf der nächsten Seite).*

(Fortsetzung der Bildunterschrift von Abb. 10.59)
„Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.
Im Vergleich zur Signalwegrangliste für die Verbindung von Lungenkrebs mit rheumatoider Arthritis sind nur geringe Unterschiede erkennbar. Dies ist zu erwarten, da es sich ebenfalls um eine Verbindung einer Krebserkrankung mit rheumatoider Arthritis handelt.

10.2.1.2 Signalwegpriorisierung zwischen Alzheimer und Demenz und Thrombose

Abb. 10.60 zeigt die Rangliste der verbindenden Signalwege zwischen den Indikationsgruppen Alzheimer und Demenz und Thrombose. Diese Verbindung ist der „Hot-Spots“ der Vernetzungsmatrix der therapeutischen Gebiete Zentralnervensystem und Herz-Kreislauf-System.

Die 31 verbindenden Signalwege wurden analog den zuvor beschriebenen Untersuchungen in einer Rangliste sortiert. Zwei Signalwege – „Neuroactive ligand-receptor interaction“ und „Calcium signaling pathway“ – stellen für die Verbindung dieser beiden Indikationsgruppen beinahe ein Drittel der Gesamtverbindungsstärke. Während der „Neuroactive ligand-receptor interaction“ Signalweg mit und ohne Berücksichtigung der Substanzanzahl einen gleich hohen Anteil an der Gesamtverbindungsstärke hat, steigt die Bedeutung des „Calcium signaling pathway“ bei Gewichtung mit einfacher Durchschnittsbildung. Ein ähnlicher Effekt ist bei dem Signalweg feststellbar, der die Komponenten der Blutgerinnung enthält („Complement and coagulation cascades“). Die erhöhte Bedeutung für die Verbindung bei einfacher Durchschnittsbildung entsteht durch Berücksichtigung der Anzahl der Substanzen pro verbindendem Zielmolekül.

Offenbar verfügen der „Calcium signaling pathway“ und der „Complement and coagulation cascades“ Signalweg über zahlreiche Substanzen pro Zielmolekül. Dies ist ein Hinweis auf die hohe Bedeutung dieser beiden Signalwege in der Entwicklung von Medikamenten und auf dem pharmazeutischen Markt. Im Bereich des Gerinnungssystem ist darüber hinaus der hohe relevante Anteil des Signalwegs auffällig: Mit 20 indirekt verbindenden Zielmolekülen und einem direkt indikationsübergreifenden Molekül ist beinahe ein Drittel des gesamten Gerinnungssignalwegs für die Verbindung relevant. Ob dieser hohe relevante

Anteil tatsächlich für eine indikationsübergreifende Funktion der Zielmoleküle des Signalwegs spricht, müsste auf Ebene der Zielmoleküle untersucht werden.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Neuroactive ligand-receptor interaction**	**17**	**17**	**5**	**33**	**17**
2.	**Calcium signaling pathway**	**13**	**16**	**2**	**28**	**18**
3.	**Complement and coagulation cascades**	**9**	**14**	**1**	**20**	**32**
4.	**MAPK signaling pathway**	**7**	**8**	**2**	**15**	**7**
5.	**GnRH signaling pathway**	**5**	**8**	**1**	**10**	**12**
6.	**Regulation of actin cytoskeleton**	**5**	**8**	**0**	**14**	**7**
7.	**VEGF signaling pathway**	**4**	**2**	**2**	**8**	**17**
8.	**Focal adhesion**	**4**	**3**	**0**	**14**	**7**
9.	**Long-term depression**	**4**	**1**	**1**	**10**	**16**
10.	**Gap junction**	**3**	**5**	**1**	**6**	**8**
11.	Long-term potentiation	3	8	2	5	13
12.	Fc epsilon RI signaling pathway	3	1	1	7	12
13.	Cytokine-cytokine receptor interaction	3	1	1	9	4
14.	Leukocyte transendothelial migration	3	1	2	5	8
15.	Cell adhesion molecules (CAMs)	2	1	0	6	5
16.	ErbB signaling pathway	2	1	0	5	6
17.	PPAR signaling pathway	1	1	1	2	4
18.	Tight junction	1	0	0	5	6
19.	Jak-STAT signaling pathway	1	1	0	4	3
20.	Toll-like receptor signaling pathway	1	0	2	1	5
21.	Insulin signaling pathway	1	0	1	3	4
22.	Apoptosis	1	0	1	3	6
23.	Axon guidance	1	0	0	3	2
24.	Natural killer cell mediated cytotoxicity	1	0	0	3	2
25.	Wnt signaling pathway	1	0	1	2	6
26.	B cell receptor signaling pathway	1	0	0	3	2
27.	Adipocytokine signaling pathway	1	0	1	1	5
28.	Hedgehog signaling pathway	1	0	0	3	4
29.	T cell receptor signaling pathway	1	0	1	1	3
30.	p53	0	0	0	2	4
31.	mTOR signaling pathway	0	0	1	0	3

Abb. 10.60: Signalwegrangliste für die Verbindung von Alzheimer und Demenz mit Thrombose

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Alzheimer und Demenz und Thrombose wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Die Abbildung zeigt, dass vor allem die beiden Signalwege „Neuroactive ligand-receptor interaction“ und „Calcium signaling pathway“ für die Verbindung von Alzheimer und Demenz mit Thrombose eine Rolle spielen. Darüber hinaus fällt der „Complement and coagulation cascades“ Signalweg auf, der über einen hohen relevanten Anteil verfügt.

10.2.1.3 Signalwegpriorisierung zwischen Schizophrenie und Bluthochdruck

Die Untersuchung der verbindenden Signalwege zwischen Schizophrenie und Bluthochdruck unterscheidet sich von den vorhergehenden Analysen, da dieser Verbindungspunkt keinen „Hot-Spot“ der Verbindung der therapeutischen Gebiete ZNS und Herz-Kreislauf-System darstellt. Mit nur neun gemeinsamen Signalwegen wurde diese Verbindung vor allem untersucht, um mögliche Effekte einer geringen Verbindungsstärke auf die Signalwegrangliste sehen zu können.

	Name des Signalwegs	**Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke,** in Prozent: Bei doppelter Durchschnittsbildung	Bei einfacher Durchschnittsbildung	**Anzahl der Zielmoleküle, die zur Verbindung beitragen:** Direkt	Indirekt	**Anteil des Signalwegs, der an der Verbindung beteiligt ist,** in Prozent
1.	Neuroactive ligand-receptor interaction	46	36	5	45	46
2.	Calcium signaling pathway	31	34	4	28	31
3.	Gap junction	11	16	4	5	11
4.	Long-term depression	3	1	0	5	3
5.	Regulation of actin cytoskeleton	3	1	0	4	3
6.	Long-term potentiation	2	11	0	4	2
7.	Cytokine-cytokine receptor interaction	2	0	0	3	2
8.	Hematopoietic cell lineage	2	1	0	3	2
9.	Jak-STAT signaling pathway	1	0	0	2	1

Abb. 10.61: Signalwegrangliste für die Verbindung von Schizophrenie mit Bluthochdruck
Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Schizophrenie und Bluthochdruck wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg.
Für die Verbindung von Schizophrenie mit Bluthochdruck gibt es nur neun gemeinsame Signalwege. Die Rangliste zeigt, dass nur die beiden großen Signalwege „Neuroactive ligand-receptor interaction“ und „Calcium signaling pathway“ über direkt indikationsübergreifende Zielmoleküle verfügen. Die Tatsache, dass diese beiden Signalwege mehr als zwei Drittel der Gesamtverbindungsstärke ausmachen, zeigt, dass die beiden Indikationsgruppen nur über wenige allgemeine Zellfunktionen verbunden sind.

Abb. 10.61 zeigt, dass vor allem große Signalwege, die grundlegende Zellfunktionen beschreiben, für die Verbindung wichtig sind. Die Signalwege „Neuroactive ligand-receptor interaction“ und „Calcium signaling pathway“ stellen gemeinsam deutlich mehr als zwei Drittel der Verbindungsstärke. Dies deutet darauf hin, dass die beiden Indikationsgruppen vor allem durch allgemeine zelluläre Prinzipien, nicht aber durch weitere spezifische Signalwege, verbunden sind. Während die Anzahl der relevanten Moleküle für die beiden großen

Signalwege vergleichbar mit den vorherigen Analysen ist (fünf bzw. vier direkt verbindende und 44 bzw. 28 indirekt verbindende Moleküle), können bei allen anderen Signalwegen keine direkten Verbindungsmoleküle identifiziert werden. Dies unterstreicht die schwache Verbindung der Indikationsgruppen.

10.2.1.4 Signalwegpriorisierung zwischen Alzheimer und Demenz und generalisierten Immunerkrankungen

Die in Abb. 10.62 dargestellte Analyse zeigt die Rangliste der verbindenden Signalwege eines der beiden „Hot-Spots“ der Verbindungsmatrix der therapeutischen Gebiete ZNS mit Transplantationssystem und Immunsystem. Wie bereits im Abschnitt 7.3.3 beschrieben wurde, ist diese Kombination von therapeutischen Gebieten nur mäßig durch gemeinsame Signalwege verbunden. Wird bei der Gewichtung die Anzahl der Substanzen pro Zielmolekül berücksichtigt, so liegt der „Hot-Spot“ der Matrix bei der Verbindung der Indikationsgruppen Alzheimer und Demenz mit generalisierten Immunerkrankungen.

Abb. 10.62 zeigt die Rangliste der 34 Signalwege, die sich die beiden Indikationsgruppen teilen. Im Vergleich zu den zuvor untersuchten „Hot-Spots“ von stark verbundenen therapeutischen Gebieten, gibt es in der Verbindung von Alzheimer und Demenz mit generalisierten Immunerkrankungen relativ wenige direkt verbindende Zielmoleküle. Lediglich der große Neurotransmitter-Signalweg („Neuroactive ligand receptor interaction“) verfügt über drei direkt indikationsübergreifende Zielmoleküle. Potential für die Entdeckung möglicher neuer indikationsübergreifender Zielmoleküle könnte aber in den zahlreichen indirekt verbindenden Molekülen liegen. Im Vergleich der anteiligen Verbindungsstärke bei Gewichtung mit einfacher beziehungsweise doppelter Durchschnittsbildung fällt auf, dass der „VEGF signaling pathway“ durch Berücksichtigung der Substanzanzahl bei einfacher Durchschnittsbildung deutlich an Bedeutung für die Verbindung gewinnt. Offenbar gibt es innerhalb dieses Signalwegs eine hohe Anzahl von Substanzen pro Zielmolekül. Dies erscheint in Anbetracht des kommerziellen Interesses alleine am VEGF-A Molekül nicht verwunderlich (vergleiche Abschnitt 3.3.2).

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Neuroactive ligand-receptor interaction**	**15**	**17**	**3**	**27**	**12**
2.	**Cytokine-cytokine receptor interaction**	**9**	**7**	**0**	**26**	**10**
3.	**Calcium signaling pathway**	**9**	**11**	**2**	**15**	**10**
4.	**MAPK signaling pathway**	**6**	**4**	**2**	**15**	**7**
5.	**Jak-STAT signaling pathway**	**6**	**3**	**0**	**14**	**9**
6.	**Toll-like receptor signaling pathway**	**4**	**4**	**1**	**14**	**15**
7.	**Natural killer cell mediated cytotoxicity**	**4**	**3**	**1**	**10**	**8**
8.	**Apoptosis**	**3**	**2**	**0**	**11**	**13**
9.	**VEGF signaling pathway**	**3**	**18**	**3**	**5**	**11**
10.	**Gap junction**	**3**	**3**	**1**	**6**	**7**
11.	GnRH signaling pathway	3	2	3	3	6
12.	Regulation of actin cytoskeleton	3	2	1	7	4
13.	Long-term potentiation	3	2	1	6	10
14.	Leukocyte transendothelial migration	3	1	2	6	7
15.	T cell receptor signaling pathway	3	2	0	9	10
16.	Cell adhesion molecules (CAMs)	2	2	0	9	7
17.	Fc epsilon RI signaling pathway	2	2	2	5	9
18.	Long-term depression	2	2	1	5	8
19.	Focal adhesion	2	1	1	6	4
20.	Taste transduction	2	1	0	4	8
21.	B cell receptor signaling pathway	2	1	0	7	11
22.	Adipocytokine signaling pathway	2	2	0	6	8
23.	ErbB signaling pathway	2	1	1	5	7
24.	Insulin signaling pathway	1	1	0	6	4
25.	Complement and coagulation cascades	1	1	0	4	6
26.	Melanogenesis	1	1	1	2	3
27.	mTOR signaling pathway	1	1	0	4	9
28.	Phosphatidylinositol signaling system	1	0	1	2	4
29.	Tight junction	1	0	1	2	3
30.	Wnt signaling pathway	1	1	1	2	2
31.	Axon guidance	1	1	0	3	2
32.	Cell cycle	1	1	0	3	3
33.	p53	1	0	0	2	3
34.	Hedgehog signaling pathway	0	0	0	2	4

Abb. 10.62: Signalwegrangliste für die Verbindung von Alzheimer und Demenz mit generalisierten Immunerkrankungen

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Alzheimer und Demenz und generalisierte Immunerkrankungen wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Die Abbildung zeigt unter anderem, dass es zwar zahlreiche gemeinsame Signalwege gibt, diese jedoch kaum über direkt verbindende Zielmoleküle verfügen. Wichtigster Signalweg ist der große „Neuroactive ligand receptor interaction“ Signalweg.

10.2.1.5 Signalwegpriorisierung zwischen multipler Sklerose und generalisierten Immunerkrankungen

Abb. 10.63 zeigt die Signalwegrangliste des zweiten „Hot-Spots“ der Verbindungsmatrix der therapeutischen Gebiete ZNS und Transplantationssystem und Immunsystem, der Verbindung zwischen multipler Sklerose und generalisierten Immunerkrankungen. Beide Indikationsgruppen teilen sich 31 Signalwege und bei Gewichtung mit doppelter Durchschnittsbildung hat diese Kombination von Indikationsgruppen den höchsten normierten Verbindungswert.

Die Rangliste der Signalwege dieses Verbindungspunkts zeigt – im Gegensatz zu der aus Abb. 10.62 – nur geringe Unterschiede der anteiligen Verbindungsstärke in Abhängigkeit von einfacher und doppelter Durchschnittsbildung. Offenbar sind alle verbindenden Signalwege durch eine ähnliche Anzahl von Substanzen pro Zielmolekül gestützt. Die beiden Signalwege, die auch für die Verbindung von Alzheimer und Demenz mit der Gruppe der generalisierten Immunerkrankungen besonders wichtig waren (vergleiche Abschnitt 10.2.1.4), belegen auch hier – allerdings in umgekehrter Reihenfolge – die Rangplätze eins und zwei.

Obwohl einer der beiden Verbindungspartner (generalisierte Immunerkrankungen) konstant geblieben ist, lassen sich verhältnismäßig viele Veränderungen in der Signalwegpriorisierung im Vergleich zu Abb. 10.62 feststellen: Die Verbindungen sind nicht nur durch eine deutlich höhere Anzahl an direkt verbindenden Zielmolekülen gestützt, einzelne Signalwege tauchen auch völlig neu in der Rangliste auf. Ein Beispiel dafür ist der „Hematopoetic cell lineage“ Signalweg, der mit einem relevanten Anteil von 22 % und 19 relevanten Zielmolekülen den vierten Rangplatz der wichtigsten Signalwege belegt und bei der Verbindung von Alzheimer und Demenz mit generalisierten Immunerkrankungen nicht einmal verbindend wirksam war. Ein anderes Beispiel ist der VEGF-Signalweg, der vom neunten Rangplatz nun völlig von der Verbindungsrangliste verschwindet. Dieser Vergleich ist ein gutes Indiz für die Spezifität der Signalwegranglisten und ihre Abhängigkeit von beiden Indikationsgruppenpartnern.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Cytokine-cytokine receptor interaction**	**15**	**13**	**11**	**24**	**14**
2.	**Neuroactive ligand-receptor interaction**	**10**	**11**	**1**	**17**	**7**
3.	**Jak-STAT signaling pathway**	**8**	**8**	**5**	**13**	**12**
4.	**Hematopoietic cell lineage**	**8**	**7**	**4**	**15**	**22**
5.	**Toll-like receptor signaling pathway**	**7**	**9**	**5**	**13**	**18**
6.	**Cell adhesion molecules (CAMs)**	**5**	**5**	**5**	**8**	**10**
7.	**MAPK signaling pathway**	**5**	**4**	**2**	**12**	**5**
8.	**Natural killer cell mediated cytotoxicity**	**4**	**6**	**3**	**9**	**9**
9.	**Regulation of actin cytoskeleton**	**4**	**4**	**2**	**8**	**5**
10.	**Apoptosis**	**3**	**3**	**2**	**7**	**11**
11.	Calcium signaling pathway	3	5	0	7	4
12.	Focal adhesion	3	3	2	7	5
13.	Leukocyte transendothelial migration	3	3	2	6	7
14.	T cell receptor signaling pathway	3	2	2	6	9
15.	Adipocytokine signaling pathway	2	2	3	2	7
16.	ECM-receptor interaction	2	3	2	3	6
17.	Fc epsilon RI signaling pathway	2	2	1	5	8
18.	Insulin signaling pathway	2	1	1	4	4
19.	ErbB signaling pathway	1	1	1	3	5
20.	B cell receptor signaling pathway	1	1	0	5	8
21.	Long-term depression	1	1	0	4	5
22.	mTOR signaling pathway	1	1	1	2	6
23.	Phosphatidylinositol signaling system	1	1	0	4	5
24.	Gap junction	1	0	0	4	4
25.	Long-term potentiation	1	0	0	3	4
26.	p53	1	0	0	2	3
27.	Tight junction	1	0	0	3	3
28.	Melanogenesis	1	0	0	3	3
29.	TGF-beta signaling pathway	1	1	1	0	1
30.	Axon guidance	1	1	1	0	1
31.	Antigen processing and presentation	0	0	1	0	1

Abb. 10.63: Signalwegrangliste für die Verbindung von multipler Sklerose mit generalisierten Immunerkrankungen

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für multiple Sklerose und generalisierte Immunerkrankungen wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Die Rangliste der 31 gemeinsamen Signalwegen wird durch den „Cytokine-cytokine receptor interaction“ Signalweg angeführt, der 11 direkt verbindende Zielmoleküle hat. Vier weitere Signalwege verfügen über vier bzw. fünf indikationsübergreifende Zielmoleküle und sind damit besonders interessant für weitere Analysen.

10.2.1.6 Signalwegpriorisierung zwischen Thrombose und rheumatoider Arthritis

Abb. 10.64 zeigt die Signalwegrangliste der Verbindung von Lungenkrebs und Thrombose, dem „Hot-Spot“ der Verbindung der therapeutischen Gebiete

Onkologie und Herz-Kreislauf-System bei Gewichtung mit doppelter Durchschnittsbildung.

Die Verbindung ist durch 35 gemeinsame Signalwege gekennzeichnet. Bei Betrachtung der Rangliste fällt auf, dass je nach Art der Gewichtung nicht nur der relative Anteil der einzelnen Signalwege an der Gesamtverbindungsstärke variiert, sondern sich sogar die Reihenfolge der Signalwege in der Rangliste ändert. Während bei doppelter Durchschnittsbildung der „Cytokine-cytokine receptor interaction“ Signalweg mit einem direkt und 39 indirekt verbindenden Zielmolekülen den größten Anteil an der Verbindung hat, sind „Calcium signaling pathway“ und „Complement and coagulation cascades“ bei einfacher Gewichtung wichtiger. Dieser Effekt muss durch die unterschiedliche Anzahl von Substanzen pro Zielmolekül verursacht sein. Das bedeutet, dass es einzelne Zielmoleküle geben muss, die für die Indikationsgruppen Thrombose und/oder rheumatoide Arthritis durch eine besonders hohe Anzahl von Substanzen adressiert werden und damit bei einfacher Durchschnittsbildung die Verbindungsstärken verändern.

Die bisherigen Analysen haben gezeigt, dass insbesondere im therapeutischen Gebiet Herz-Kreislauf-System hohe Substanzanzahlen zu einer Veränderung des Vernetzungsbildes führen können (vergleiche Abschnitt 7.2.3.3 und Abschnitt 7.3.4). Daher liegt die Vermutung nahe, dass die hier beobachtete Veränderung der Rangfolge durch Substanzen mit Indikationen im Bereich Thrombose verursacht wird.

Neben den oben genannten Signalwegen fällt vor allem der Signalweg „MAPK signaling pathway“ auf, aus dem ein Drittel aller Moleküle für die Verbindung indirekt oder direkt relevant sind. Dieser hohe relevante Anteil legt eine spezifische Funktion in den beiden Indikationsgruppen nahe und bietet einen möglichen Ansatzpunkt für Forschung und Entwicklung.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Cytokine-cytokine receptor interaction**	**10**	**6**	**1**	**39**	**17**
2.	**Neuroactive ligand-receptor interaction**	**8**	**9**	**3**	**25**	**12**
3.	**Calcium signaling pathway**	**6**	**10**	**2**	**19**	**10**
4.	**Complement and coagulation cascades**	**6**	**10**	**0**	**23**	**27**
5.	**MAPK signaling pathway**	**6**	**9**	**2**	**22**	**33**
6.	**Hematopoietic cell lineage**	**5**	**4**	**0**	**24**	**13**
7.	**Jak-STAT signaling pathway**	**4**	**2**	**0**	**17**	**10**
8.	**Focal adhesion**	**4**	**2**	**2**	**15**	**17**
9.	**Regulation of actin cytoskeleton**	**4**	**6**	**1**	**14**	**11**
10.	**Natural killer cell mediated cytotoxicity**	**4**	**2**	**0**	**15**	**8**
11.	Toll-like receptor signaling pathway	4	3	2	13	11
12.	VEGF signaling pathway	3	6	2	8	15
13.	GnRH signaling pathway	3	6	1	10	12
14.	Fc epsilon RI signaling pathway	3	3	1	10	17
15.	Cell adhesion molecules (CAMs)	3	1	1	9	16
16.	ECM-receptor interaction	3	2	1	11	16
17.	Leukocyte transendothelial migration	2	2	3	5	13
18.	Long-term depression	2	1	2	8	10
19.	Apoptosis	2	2	1	9	8
20.	Adipocytokine signaling pathway	2	2	1	7	13
21.	T cell receptor signaling pathway	2	1	1	6	9
22.	ErbB signaling pathway	1	0	0	6	9
23.	Gap junction	1	3	0	5	7
24.	Insulin signaling pathway	1	0	0	5	7
25.	B cell receptor signaling pathway	1	0	1	4	5
26.	p53	1	0	0	5	4
27.	Long-term potentiation	1	5	0	5	4
28.	TGF-beta signaling pathway	1	1	0	4	7
29.	Tight junction	1	0	0	4	4
30.	Axon guidance	1	0	2	1	4
31.	PPAR signaling pathway	1	0	0	3	5
32.	Adherens junction	1	0	1	2	4
33.	mTOR signaling pathway	1	0	0	3	6
34.	Wnt signaling pathway	0	0	0	2	2
35.	Cell Communication	0	0	0	3	1

Abb. 10.64: Signalwegrangliste für die Verbindung von Thrombose mit rheumatoider Arthritis

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Thrombose und rheumatoide Arthritis wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Auffälligstes Merkmal der Rangliste ist der starke Einfluss der Veränderung der Gewichtungsart auf die Rangfolge der Signalwege. Offenbar gibt es verschiedene Zielmoleküle, die bei Thrombose und/oder rheumatoider Arthritis durch besonders viele Substanzen adressiert werden und daher bei Berücksichtigung der Substanzanzahl bei einfacher Durchschnittsbildung zu einer Verschiebung der Rangliste führen.

10.2.1.7 Signalwegpriorisierung zwischen Lungenkrebs und Thrombose

Abb. 10.65 zeigt die Signalwegrangliste der Verbindung der Indikationsgruppen Lungenkrebs und Thrombose. Diese Verbindung verfügt über 35 gemeinsame Signalwege und ist der „Hot-Spot“ der Verbindungsmatrix der therapeutischen Gebiete Onkologie und Herz-Kreislauf-System bei doppelter Durchschnittsbildung.

Es fällt auf, dass diese Verbindung durch viele verschiedene Signalwege in ähnlichem Maß getrieben wird – kein Signalweg hat bei Gewichtung mit doppelter Durchschnittsbildung mehr als 7 % Anteil an der Gesamtverbindungsstärke. Die relevanten Anteile der Signalwege, die auf den ersten beiden Rangplätzen liegen („Cytokine-cytokine receptor interaction“ und „Focal adhesion“), von 14 beziehungsweise 15 % sind verhältnismäßig niedrig und deuten eher nicht auf eine spezifische Funktion in der Verbindung dieser Indikationsgruppen hin. Mit sieben direkt verbindenden und 22 indirekt verbindenden Zielmolekülen, könnte aber der „Focal adhesion“ Signalweg ein interessanter Forschungskandidat sein.

Ähnlich wie bei der Rangliste der Verbindung von Thrombose mit rheumatoider Arthritis (vergleiche Abb. 10.64) kann auch in der Rangliste in Abb. 10.65 ein Effekt durch die hohe Substanzanzahl für einzelne Zielmoleküle beobachtet werden. Der Signalweg „Calcium signaling pathway“ verfügt über kein einziges direkt verbindendes Zielmolekül; dennoch rückt er bei Gewichtung mit einfacher Durchschnittsbildung auf Platz eins der Rangliste. Dies wird vermutlich dadurch verursacht, dass es ein oder mehrere Zielmoleküle im Signalweg gibt, für die es ein hohes kommerzielles Interesse für die Indikationsgruppe Thrombose gibt. Dies verzerrt die Berechnung der Verbindungsstärke des Signalwegs, wenn nicht durch doppelte Durchschnittsbildung um die Substanzanzahl korrigiert wird.

	Name des Signalwegs	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei doppelter Durchschnittsbildung	Anteil des Signalwegs an der gewichteten Gesamtverbindungsstärke, in Prozent: Bei einfacher Durchschnittsbildung	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Direkt	Anzahl der Zielmoleküle, die zur Verbindung beitragen: Indirekt	Anteil des Signalwegs, der an der Verbindung beteiligt ist, in Prozent
1.	**Cytokine-cytokine receptor interaction**	**7**	**6**	**3**	**31**	**14**
2.	**Focal adhesion**	**6**	**7**	**7**	**22**	**15**
3.	**MAPK signaling pathway**	**6**	**8**	**2**	**29**	**12**
4.	**Calcium signaling pathway**	**6**	**11**	**0**	**27**	**16**
5.	**Regulation of actin cytoskeleton**	**5**	**7**	**4**	**19**	**11**
6.	**Neuroactive ligand-receptor interaction**	**5**	**7**	**0**	**23**	**9**
7.	**Complement and coagulation cascades**	**5**	**8**	**1**	**20**	**30**
8.	**Hematopoietic cell lineage**	**4**	**3**	**1**	**19**	**23**
9.	**ErbB signaling pathway**	**4**	**3**	**1**	**17**	**21**
10.	**VEGF signaling pathway**	**4**	**3**	**2**	**14**	**23**
11.	GnRH signaling pathway	4	6	1	15	16
12.	Gap junction	3	6	1	14	15
13.	Jak-STAT signaling pathway	3	1	0	16	10
14.	Fc epsilon RI signaling pathway	3	1	0	15	20
15.	Long-term depression	3	1	1	13	18
16.	Toll-like receptor signaling pathway	3	1	1	11	12
17.	Natural killer cell mediated cytotoxicity	3	1	0	13	10
18.	Apoptosis	2	1	1	13	17
19.	Cell adhesion molecules (CAMs)	2	1	2	8	8
20.	Adipocytokine signaling pathway	2	1	1	11	17
21.	ECM-receptor interaction	2	2	3	8	13
22.	Leukocyte transendothelial migration	2	1	3	6	8
23.	T cell receptor signaling pathway	2	1	1	8	10
24.	p53	2	1	0	11	16
25.	Adherens junction	2	2	3	4	9
26.	PPAR signaling pathway	2	1	0	7	10
27.	Long-term potentiation	2	5	0	8	12
28.	Insulin signaling pathway	2	1	0	7	5
29.	Axon guidance	1	1	3	3	5
30.	Tight junction	1	1	1	6	6
31.	B cell receptor signaling pathway	1	1	1	4	8
32.	Wnt signaling pathway	1	0	0	5	3
33.	mTOR signaling pathway	1	1	1	3	9
34.	TGF-beta signaling pathway	1	0	0	6	7
35.	ABC transporters	1	0	0	2	5

Abb. 10.65: Signalwegrangliste für die Verbindung von Lungenkrebs mit Thrombose

Die Sortierung aller gemeinsamen Signalwege erfolgt entsprechend ihres Anteils an der Gesamtverbindungsstärke bei doppelter Durchschnittsbildung (erste Spalte). Zum Vergleich wird ebenfalls der Anteil des Signalwegs bei einfacher Durchschnittsbildung angegeben (zweite Spalte). In der dritten und vierten Spalte steht die Zahl der Zielmoleküle, die zur Signalwegverbindung beitragen. Unter „direkt“ sind die Moleküle aufgeführt, die bereits heute für Lungenkrebs und Thrombose wichtig sind. „Indirekt“ nennt die Anzahl der Moleküle, die Teil des Signalwegs sind, aber erst für eine der beiden Indikationsgruppen genutzt werden. In der letzten Spalte steht der Anteil des Signalwegs, der für die Verbindung wichtig ist. Dieser errechnet sich aus der Gesamtzahl der relevanten Zielmoleküle dividiert durch die Gesamtanzahl von Molekülen im Signalweg. Die zehn wichtigsten Signalwege sind grau unterlegt.

Die Verbindung von Lungenkrebs mit Thrombose ist der „Hot-Spot“ der Verbindung der therapeutischen Gebiete Onkologie und Herz-Kreislauf-System und wird durch 35 gemeinsame Signalwege gestützt, die in ähnlichem Maß zur Verbindung beitragen. Kein Signalweg kann bei doppelter Durchschnittsbildung mehr als 7 % zur Verbindungsstärke beisteuern. Die hohe Bedeutung des Signalwegs „Calcium signaling pathway“ bei einfacher Durchschnittsbildung wird vermutlich durch hohes kommerzielles Interesse an Anti-Thrombose-Medikamenten verursacht.

.

10.3 Abkürzungsverzeichnis

Abkürzung	*Bedeutung*
Abb.	Abbildung
Akt1/PKB	Proteinkinase B
bzw.	beziehungsweise
ca.	circa
cGMP	cyclisches Guanosinmonophosphat
COPD	chronisch obstruktive Lungenerkrankung
COX-2	Cyclooxygenase 2
d.h.	das heißt
DNA	Desoxyribonukleinsäure
DR4	Death receptor 4
EMEA	Europäische Arzneimittelagentur („European Medicines Agency“)
ERM	Entity-Relationship-Modell
F&E	Forschung und Entwicklung
FDA	Amerikanische Arzneimittelagentur („Food and Drug Administration“)
FLS	synoviale Fibroblasten („Fibroblast-like synoviocytes“)
GCP-2	„Granulocyte-chemotactiv protein 2“
IL1R1	Interleukin 1 Rezeptor Typ I
IL1-RA	Interleukin 1 Rezeptor Antagonist
IL1β	Interleukin 1 β
IND	„Investigational new drug“
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAPK	„mitogen-activated protein kinase”
MDSC	„Myeloid-derived suppressor cells“
MIF	„Macrophage migration inhibitory factor“
mRNA	Messenger Ribonukleinsäure
mTOR	„mammalian Target of Rapamycin”

NDA	„New drug application“
NO	Stickstoffmonoxid
NSCLC	nicht-kleinzelliges Lungenkarzinom („Non small cell lung cancer“)
PDE	Phosphodiesterase
PKC	Proteinkinase C
PTHrP	„parathyroid hormone related protein“
SCLC	kleinzelliges Lungenkarzinom („Small cell lung cancer“)
TA	therapeutisches Gebiet („Therapeutic area“)
TNFRSF10A	„Tumor necrosis factor related apoptosis inducing ligand (TRAIL) receptor superfamily member 10 A” (alias DR4)
TRAIL	„Tumor necrosis factor related apoptosis inducing ligand”
VEFFR	„Vascular endothelial growth factor“ Rezeptor
VEGF	Gefäßwachstumsfaktor („Vascular endothelial growth factor“)

10.4 Verzeichnis der akademischen Lehrer

Meine akademischen Lehrer waren die folgenden Damen und Herren in Marburg:

Aigner, Aumüller, Bastians, Bauer, Berndt, Besedovsky, Boekhoff, Braun, Czubayko, Daut, del Rey, Eilers, Elsässer, Feuser, Frenking, Garten, Grzeschik, Gebhard, Gromes, Gudermann, Hasilik, Hassel, Heeg, Jungclas, Kirchner, Koolman, Kunz, Lang, Lill, Löffler, Lutz, Moll, Müller, Plant, Renz, Renkavitz-Pohl, Roeper, Röhm, Schäfer, Schulz, Seitz, Suske, Weihe, Westermann

10.5 Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Frank Czubayko und Herrn Prof. Dr. Thomas Gudermann. Sie haben mir nicht nur die Möglichkeit gegeben, am Institut für Pharmakologie und Toxikologie meine Doktorarbeit anzufertigen, sondern auch durch ihre umfassende wissenschaftliche Betreuung sowie ihre wertvollen Anregungen und Diskussionen einen erfolgreichen Abschluss dieser Arbeit ermöglicht.

Meinen besonderen Dank möchte ich Herrn Dr. Jakub Pijewski aussprechen. Er hat mich mit vielen Diskussionen, Anregungen und Rat und Tat nicht nur zum Thema Microsoft Access unterstützt. Ich möchte mich außerdem bei Herrn Dr. Thomas Rudolph für seine Unterstützung im Vorfeld der Doktorarbeit bedanken.

Der Firma McKinsey & Company danke ich für die Möglichkeit der Nutzung der Pharmaprojects-Rohdaten.

Herrn Dr. Frank Langer danke ich für seine Hilfe bei statistischen Fragestellungen.

Meinen ehemaligen Mit-Doktoranden Frau Dorke Meyer und Herrn Dr. Hermann Kalwa danke ich für Rat und Tat bei Fragen aller Art und vor allem dafür, dass ich mich immer in Marburg willkommen fühlen durfte.

Ein großes Dankeschön für ihre unschätzbare persönliche und praktische Unterstützung geht außerdem an meine Freundin Yuldon GyanaTshang.

Schließlich möchte ich noch ganz besonders meinen Eltern Heidi und Ludwig Fehr und meinem Freund Sebastian Moebus danken. Ihr Rückhalt, ihr Verständnis und ihr Vertrauen waren meine größte Hilfe. Mein Vater verdient hier meinen ganz besonderen Dank. Er hat sich im Verlauf dieser Arbeit nicht nur durch viele hundert Seiten Präsentationen gearbeitet, sondern auch alle Entwürfe in kürzester Zeit Korrektur gelesen.

Papa, deine Kritik, deine Anregungen und dein Feedback haben entscheidend zum Erfolg dieser Arbeit beigetragen.