The Segmentation of Therapeutic Populations in Oncology

Mark R. Trusheim, Ernst R. Berndt

Abstract

The emergence of targeted therapies in oncology has enormous potential to improve care for individual patients and specific subpopulations. A consequence of these advances is the segmentation of oncology into numerous sub-types, often with small patient populations. In this research we document the growth of candidate therapeutics along with the resultant segmentation of oncology treatments, based on a querying of the PharmaProjects® database for all anti-cancer indications from 1995 through 2011, augmented with Centers for Disease Control and Prevention (CDC) incidence information. We find that the number of marketed oncology treatments has doubled since 1995 from 71 to 146, while the proportion of molecular entities in Phase II/III programs with unidentified targets has declined from 56 percent to 28 percent. The number of anti-cancer compounds marketed or in active programs increased to 1,428 from 538, while active molecular targets under investigation have expanded from 99 to 265, with 590 distinct total targets explored over this period. We consider clinical, economic and development challenges emanating from the increasingly segmented oncology marketplace.

1. INTRODUCTION AND HISTORICAL BACKGROUND

With an estimated 570,000 deaths in 2010, cancer remains the second leading cause of death, accounting for 25 percent of total U.S. mortality. Since President Nixon declared the “War on Cancer” with passage of the National Cancer Act of 1971 and an appropriation of an extra $100 million for the National Cancer Institute (NCI), cancer treatments have splintered into numerous major sub-types, with a cumulative 146 distinct therapeutic anti-cancer agents approved for marketing and launched in at least one country by 2011.

In recent years, personalized medicine -- utilizing molecular targeted therapies alone or in combination -- has received much attention, accelerating a decades-long history of cancer

---

3 Number of active approved therapeutics based on authors’ analysis in the PharmaProjects® data base as of May 2011. See Methods and Data for further details.

Mark R. Trusheim is a Visiting Scientist at MIT.
Ernst R. Berndt is the Louis E. Seley Professor in Applied Economics at MIT.
The number of cancer sub-types and associated therapies has grown remarkably since Rudolph Virchow [1858] first characterized microscopic pathology to phenotypic cancer in the 1850s. Chemotherapy was introduced in the 1940s when Huggins and Hodges [1941] proposed stilbestrol for prostate cancer treatment and Department of Defense-sponsored pharmacologists discovered that intravenous mustard gas (mustine) temporarily reduced non-Hodgkin’s lymphoma tumor masses. Reasoning from the ability of folic acid to treat anemia, Farber pursued a molecular target approach to treating leukemia by testing an experimental Lederle folic acid inhibitor, aminopterin (AMT), to reduce bone marrow blood cell creation.

Development times have traditionally been long in cancer. Farber, Diamond, Mercer et al. (1948) reported that AMT treatment led to temporary remission of leukemia, but it required another five years before Lederle received FDA approval, who marketed AMT from 1953–1964 for pediatric cases. By 1971 combination therapy using chemotherapeutic agents was established, based on the breakthrough six years earlier by Frei, Karon, Levin et al. (1965) who reasoned that just as antibiotics with different mechanisms of action (MOA) proved more effective in treating tuberculosis than monotherapy, so too should combining chemotherapy agents with differing MOAs prove superior in fighting cancer. Even so, in 1971 only about 30 drugs were approved for cancer and fewer than 100 oncologists were estimated to have been practicing in the U.S. From its earliest days, medicinal oncology research has explored how differences in cancer MOAs could be exploited in treatments that segmented cancer into new sub-types.

---

4 For a more detailed history of cancer diagnoses and treatment, see Mukherjee [2010].

5 Goodman, Wintrobe, Dameshek et al. [1946].


7 Goodman, Wintrobe, Dameshek et al. [1946,1984].

8 Vanchieri [2007], p. 342.

9 Bloch [2004], p. 98.
Cancer management typically uses therapeutics delivered in short treatment courses for particular cancer sub-types. Traditionally, these sub-types have been based on the stage and organ of origin of the primary tumor. Thus, a drug might have been specific to lung cancer or to breast cancer and to a specific treatment stage such as first line, metastatic or adjuvant (an additional treatment to assist in the action of the main ingredient or treatment). For example, the chemotherapeutic the NCI describes as the most well-known natural-source cancer drug in the United States\textsuperscript{10}, paclitaxel (Taxol, Bristol-Myers Squibb), is not FDA approved for all cancers but instead is approved only for a patchwork of sub-types consisting of an AIDS induced sarcoma and three carcinoma organ sub-types further delineated by stage and prior treatment protocol: first line and subsequent therapy for ovarian cancer; adjuvant treatment of node-positive breast cancer following doxorubicin-containing chemotherapy; metastatic breast cancer after chemotherapy failure or six month relapse; first line treatment in combination with cisplatin of non-small cell lung cancer in patients not candidates for curative surgery and/or radiation therapy; and second-line treatment of AIDS-related Kaposi’s sarcoma.\textsuperscript{11}

In the past 15 years, cancer treatment has accelerated its migration toward molecular sub-types, with therapeutics targeted to particular molecular profiles of patient or tumor. Beyond breast cancer, which has been segmented into different therapeutic categories based on expression levels of HER2\textsuperscript{12} and hormone receptors\textsuperscript{13}, among others a KRAS colorectal cancer

\textsuperscript{10} National Cancer Institute [not dated].
\textsuperscript{13} Burstein, Prestrud, Seidenfeld et al. [2010].
(CRC) sub-type\textsuperscript{14} and an ALK-EML4 non-small cell lung cancer (NSCLC) sub-type\textsuperscript{15} have also emerged. The National Cancer Institute’s Human Cancer Genome Atlas is further classifying genetic sub-types in brain, ovarian, breast, and thirteen other cancer types\textsuperscript{16}. One can plausibly conclude that the segmentation of oncology types and therapeutic options along molecular lines currently stands at an early stage with substantial potential for further growth.

Both general cytotoxic and molecularly targeted therapies currently gain approval through trials on individual organ types. As measured by the number of supplemental New Drug Application (sNDA) approvals, on average recent molecular targeted agents have received regulatory approval for only several cancer sub-types.\textsuperscript{17} Moreover, even the most successful molecular targeted therapeutics do not benefit all organ types that display the biomarker. For example, as shown in Figure 2, imatinib (Gleevec: Novartis) has been proven beneficial in Philadelphia chromosome positive CML and acute lymphatic leukemia, Kit (CD117) positive gastrointestinal cancer and mastocytosis with specific c-Kit mutations and PDGFR (platelet derived growth factor receptor) gene rearrangements, but has not been successful in biomarker sub-types of CRC, NSCLC, liver, renal, ovarian, thyroid and head and neck cancers.\textsuperscript{18} Trastuzumab (Herceptin: Roche/Genentech), initially targeted HER2 3+ overexpressing breast cancer. In the 12 years since its approval, its use has only been extended to one other organ type: metastatic gastric or GEJ (Gastroesophageal junction) cancer (in October 2010)\textsuperscript{19} while demonstrating no benefit in Phase II trials in CRC, NSCLC, bladder, renal and pancreatic cancers.\textsuperscript{20} Also, while KRAS has proven a powerful companion diagnostic in CRC for EGFR inhibitors such as cetuximab (Erbitux: Bristol-Myers Squibb, Eli Lilly & Co., Merck Serono) and panitumumab (Vectibix: Amgen), it has not proven an effective companion diagnostic in other EGFR inhibitor indications such as squamous cell carcinoma head and neck (SCCHN)\textsuperscript{21} and NSCLC.\textsuperscript{22} The clinical experiences of these four drugs illustrate that rather than shifting from organ of origin sub-typing to molecular sub-typing, instead interactions of both sub-types must be considered, thereby segmenting cancer into ever more numerous and smaller patient sub-populations- imatinib (Gleevec: Novartis) has been proven beneficial in Philadelphia chromosome positive CML and acute lymphatic leukemia, Kit (CD117) positive gastroIntestinal cancer and mastocytosis with specific c-Kit mutations and PDGFR (platelet derived growth factor receptor) gene rearrangements, but has not been successful in biomarker sub-types of CRC, NSCLC, liver, renal, ovarian, thyroid and head and neck cancers.\textsuperscript{21} Trastuzumab (Herceptin: Roche/Genentech), initially targeted HER2 3+ overexpressing breast cancer. In the 12 years since its approval, its use had only been extended to one other organ type: metastatic gastric or GEJ (Gastroesophageal junction) cancer (in October 2010)\textsuperscript{15} while demonstrating no

\textsuperscript{14} Adam, Haller Poston et al. [2010].
\textsuperscript{15} Soda, Choi, Enomoto et al. [2007].
\textsuperscript{16} National Cancer Institute [2011].
\textsuperscript{17} Trusheim, Aitken, Berndt [2010].
\textsuperscript{18} Authors’ analysis of PharmaProjects® imatinib drug profile report.
\textsuperscript{20} Authors’ analysis of PharmaProjects® trastuzumab drug profile report.
\textsuperscript{21} European Medicines Agency [2008].
\textsuperscript{22} Mack, Holland, Redman et al. [2009]; O”Byrne, Bondarenko, Barrios et al. [2009].
benefit in Phase II trials in CRC, NSCLC, bladder, renal and pancreatic cancers\textsuperscript{21}. Also, while KRAS has proven a powerful companion diagnostic in CRC for EGFR inhibitors such as cetuximab (Erbitux: Bristol-Myers Squibb, Eli Lilly & Co., Merck Serono) and panitumumab (Vectibix: Amgen), it has not proven an effective companion diagnostic in other EGFR inhibitor indications such as squamous cell carcinoma head and neck (SCCHN)\textsuperscript{22} and NSCLC\textsuperscript{23,24}. The clinical experiences of these four drugs illustrate that rather than shifting from organ of origin sub-typing to molecular sub-typing, instead interactions of both sub-types must be considered, thereby segmenting cancer into ever more numerous and smaller patient sub-populations.

**Figure 2: Molecularly Targeted Therapies Unreliably Translate Across Organ of Origin Sub-Types**

<table>
<thead>
<tr>
<th>Generic</th>
<th>Brand</th>
<th>Companion Diagnostic/Marker</th>
<th>Successful Organ of Origin Type</th>
<th>Failed Organ of Origin Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>imatinib</td>
<td>Gleevec</td>
<td>Philadelphia Chromosome c-Kit PDGFR</td>
<td>CML GIST mastocytosis</td>
<td>CRC, NSCLC, liver, renal, ovarian, thyroid, head and neck</td>
</tr>
<tr>
<td>trastuzumab</td>
<td>Herceptin</td>
<td>HER2 PDGFR</td>
<td>breast cancer, GIST</td>
<td>All others tested</td>
</tr>
<tr>
<td>cetuximab</td>
<td>Erbitux</td>
<td>EGFR then KRAS</td>
<td>CRC</td>
<td>SCCHN, NSCLC</td>
</tr>
<tr>
<td>panitumumab</td>
<td>Vectibix</td>
<td>KRAS</td>
<td>CRC</td>
<td></td>
</tr>
</tbody>
</table>

Using annual 1995-2011 data, here we document the segmentation of cancer by quantifying the distribution of treatments and those under development over time. We examine the degree of segmentation by organ of origin and molecular target. In interpreting these trends, we examine simple correlations between sub-type incidence levels and the number of approved therapeutic agents, thereby evaluating the “directed technological change” hypothesis linking targeted R&D efforts to market size.\textsuperscript{23}

2. Methods and Data

The core data are taken from the PharmaProjects\textsuperscript{\textregistered} (Citeline) industry pipeline database, augmented with 2006 cancer incidence information from the CDC’s U.S. Cancer Statistics Working Group [2010]. PharmaProjects\textsuperscript{\textregistered} collects international molecular therapy pipeline information based on data from public sources ranging from peer-reviewed publications and government filings to company press releases, conference presentations and interviews. PharmaProjects\textsuperscript{\textregistered} does not collect information on research programs from academic, government, or non-profit settings, except to the extent those programs are conducted in collaboration with industry. Since few therapeutics have been approved without industry participation, this plausibly introduces minimal bias for late stage and marketed products. However, early stage candidate therapeutics may be undercounted leading to an understatement of segmentation. While other data base sources such as ClinicalTrials.gov include selected information on non-industry pipeline research programs, they do not provide comparatively detailed clinical trial information on molecular targets and mechanisms of action for either

\textsuperscript{23} Acemoglu and Linn [2004]; Finkelstein [2004].
industry or non-industry research programs. Because of our need to quantify segmentation by molecular target, we utilize the PharmaProjects® data base.

We queried the PharmaProjects® Trends database for all compounds with anti-cancer primary indications in active development, registration or marketing in one or more countries, not necessarily the U.S., from 1995 through 2011. The Trends database archives a snapshot of selected database fields in May of each year. We then join these data across the PharmaProjects® master file, which associates each compound with its molecular targets where known. PharmaProjects® designates compounds without assignment as “not applicable” (e.g., cell therapies, vaccines and modified viruses), “not available on Locus Link/Entrez Gene” (target described by developer could not be matched), or “unspecified” (such as general cytotoxic agents). We excluded from product counts generic drugs with indications similar to the original compound.

Adjuvant, pain, nausea, infection and related therapies including antibiotics, anabolic steroids, and growth stimulants are clearly critical oncology-related treatments. However, because this study focuses on primary anti-cancer treatments, we excluded all such products and candidate products from our analyses.

3. Results

Between 1995 and 2011, the number of anti-cancer molecular entities developed and launched has at least doubled. Figure 3 displays time trends for launched products, late stage development (Phase II, Phase III, pre-registration and registered but not launched) candidate products, early stage development (preclinical and Phase I) candidate products and total in All Phases. By 1995, the initial year analyzed, 71 distinct molecular entities were approved and marketed as anti-cancer treatments. In 2011 this number doubled to 146, an average annual growth rate (AAGR) of 4.6 percent.

The number of anti-cancer molecular entities in late stage development more than tripled from 98 in 1995 to 333 in 2011 (7.9 percent AAGR), increasing substantially from 109 to 165 molecular entities in development between 1998 and 2001, growing to 182 molecules in 2003, then stabilizing for several years, and growing more rapidly since 2006, reaching 333 unique molecular entities in 2011.
Figure 3: Number of anti-cancer molecular entities marketed and in development

The quantity of unique anti-cancer molecular entities in early stage development grew consistently from 369 candidates in 1995 to a peak of 1,053 in 2009. However, since then the volume of early stage compounds has declined to 949 in 2011—the only declines in the 17 year period. (Note that each year the snapshot of pipeline products is taken in May.)

The recorded total unique molecular entities in all development phases and commercialization increased from 538 in 1995 to 1,428 in 2011, a 165 percent absolute increase and a 6.1 percent AAGR. The aggregate number of programs trend is dominated by preclinical compounds. The increase in all candidates under development 1998-2001 grew by only by 32 percent, compared with a 51 percent increase for compounds in late stage development over the same time period. An even larger growth divergence occurred from 2009 to 2011, when late stage candidate products grew by 22.4 percent while the number of early stage candidate products declined by 9.9 percent.

4. Segmentation by Organ Type

Figure 4 juxtaposes the US incidence of selected cancer types as reported by the CDC with the number of distinct therapeutic agents marketed for that indication. Ordered left-to-right in descending rank by the number of therapeutics, Figure 4 displays a general trend positively
correlating incidence with the number of available agents: all the highest incidence sub-types—lung, breast, prostate, and colorectal cancers—have relatively large numbers of available therapeutics, while lower-incidence sub-types have fewer. This is consistent with the Acemoglu-Linn [2004] and Finkelstein [2004] “directed technological change” hypothesis linking number of approved therapeutics to market size, and with R&D investments in “orphan diseases” having varying incidence, as reported by Yin [2008]. Notable exceptions include thyroid cancer, a higher incidence sub-type with relatively few treatment options, and chronic myelogenous leukaemia (CML), a low incidence sub-type with relatively numerous therapeutics.

**Figure 4:** Cancer organ of origin incidence with approved therapeutics for each

**Figure 4:** Number of therapies is the number of unique molecular entities approved and marketed in at least one national jurisdiction for that organ of origin cancer type. Annual U.S. incidence counts obtained from the Centers for Disease Control and Prevention and National Cancer Institute U. S. Cancer Statistics Working Group. U.S. Cancer Statistics Working Group [2010].
A low incidence cancer sub-type only rarely has a dedicated therapeutic, i.e., one exclusively approved for use in that sub-type. Biliary cancer, nasopharyngeal cancer and Hodgkin’s lymphoma each has only two regulatory approved therapeutics, both of which are approved for use in other cancer sub-types as well. Brain cancer, having an annual incidence of 17,918 US reported cases in 2006, represents more than double the number of cases of biliary cancer and Hodgkin’s lymphoma, and more than ten times the number of nasopharyngeal cancers. With this larger incidence, brain cancer has two dedicated therapeutics out of nine total approved therapeutics for its treatment. Even more strikingly, breast cancer is the only approved indication for 14 of the 43 total approved therapeutics for this high incidence cancer, which is diagnosed over 11 times more frequently than brain cancer.\textsuperscript{24}

The linear bivariate regression in Figure 5 documents a 0.84 correlation between cancer organ sub-type annual patient incidence and the number of therapeutics approved to treat that cancer sub-type ($R^2 = 0.71$). Notably, very small population sub-types average about five therapeutics, not zero. We can interpret this as implying that even small segments attract investment, but as described earlier, many of these approved therapeutics for small indications are spillovers (supplementary approvals) from other more incident cancer types. Moreover, the range is skewed upwards by several special chronic cancer sub-types such as CML.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Bivariate regression of segment patient incidence and available drug therapies}
\end{figure}

\textsuperscript{24} U.S. Cancer Statistics Working Group [2010].
5. Segmentation by Target

Recent scientific discoveries interact segmentation by organ type with molecular typing. Figure 6 displays the number of known molecular targets engaged by the therapeutic programs in Figure 3 from 1995 through 2011. The number of targets with at least one marketed anticancer product has grown from 11 in 1995 to 31 in 2011 (6.5 percent AAGR). Similarly, between 1995-2011, the quantity of molecular targets with at least one agent in late stage development has grown from 27 to 114 (9.4 percent AAGR). In comparison, the number of targets in early stage development grew from 80 in 1995 to a peak of 234 in 2010, but then declined to 209 in 2011.
From 1995 to 2011, the sum total of active targets rose from 99 to 265 (6.2 percent AAGR), with a total of 590 distinct targets investigated over the 17 year period by at least one therapeutic program. As might be expected, however, a high failure rate in preclinical drug discovery and early development yields a total number of targets explored in any given year far exceeding the number of eventually approved and marketed products.

Finally, oncology drug development has become more focused. Figure 7 documents that the percentage of late stage clinical development programs directed at unidentified targets declined substantially from 54 percent in 1995 to only 29 percent in 2011. The share of launched products with unknown targets has declined more slowly, due both to the legacy of drugs approved pre-1995 and the approval of drugs from the 1995-2001 cohorts when unidentified percentages exceeded 40 percent. Since program target information is based on current 2011 knowledge, and because programs with unidentified targets include cellular therapies and therapeutic vaccines which are generally new therapeutic modalities, this analysis likely
understates the number of unidentified target programs in 1995 and overstates them in 2011, making the true decline likely even greater than Figure 7 suggests.

6. Discussion

Building in part on the Human Genome Project, scientific advances in understanding the molecular basis of oncology have spurred substantial industry activity in developing new oncology treatments since 1995. As quantified here by the number of new molecular entities in development and the number of molecular targets being explored, activity has more than doubled over the period with nearly three quarters (71 percent) of late stage development candidates now attacking specific targets, often with associated segmented patient sub-populations. To date, the smallest incidence cancer indications have generally received as little as one quarter as many distinct therapeutic investigations as the largest incidence cancers (5 versus 23 to 46 for prostate, lung and breast cancer) and even lower percentages of unique, single cancer therapeutics. While the volume of candidates in late stage development continues to rise, since 2009 the early stage development pipeline has experienced its first declines, approaching ten percent, both in the number of active unique candidates and distinct molecular targets being explored.

Compared to the millions of patients with cardiovascular disease or diabetes, even the largest oncology indications, such as breast cancer with 209,000 new U.S. cases in 2010\textsuperscript{25}, are relatively small, particularly since incidence approximates annual treated patient populations as most drug regimens are short term rather than chronic/maintenance. As a benchmark, the U.S. Orphan Drug Act defines an orphan disease as one with fewer than 200,000 U.S. patients.\textsuperscript{26} Molecular typing further segments these diseases. For example, breast cancer currently divides into categories such as HER2 positive (\textasciitilde 25 percent prevalence)\textsuperscript{27}, hormone receptor positive (either estrogen or progesterone), or triple negative (none of the prior three, \textasciitilde 15 percent prevalence).\textsuperscript{28} The largely unmet need for triple negative breast cancer treatments represents a population of only about 35,000 new U.S. patients per year.

Unfortunately, to offset this segmentation, there is no guarantee that therapeutics developed for a molecular target in one organ sub-type will prove effective in other organ sub-types having the same biomarker. For example, the anti-angiogenesis drug bevacizumab (Avastin:Roche Genentech) targeting VEGF proved effective in metastatic colorectal cancer, but is less effective in breast cancer or early stage colon cancer.\textsuperscript{29} Clinical trials in each organ type for each molecular sub-type are likely necessary for the foreseeable future, which may strain clinicians and exhaust patient subject pools, delaying trials, postponing patient access and raising costs.

Our trend analysis of the oncology pipeline from 1995 to the present, emphasizing oncology segmentation by organ of origin and therapeutic molecular target in addition to net product approvals, provides an objective quantification of oncology drug treatment evolution. It

\textsuperscript{25} American Cancer Society [2010].
\textsuperscript{26} Pollack [1990].
\textsuperscript{27} Pegram, Pauleti, Slamon [1998].
\textsuperscript{28} Aggrawal and Swain [2008].
\textsuperscript{29} Siow, Baas, Wakelee [2010]; Roche [2010].
may also foreshadow the opportunities, pace of change and challenges facing oncology payers and providers from recent shifts in the drug development pipeline.

While in this analysis we have characterized observed R&D output, measured by the past 17 year pipeline of compounds in development and marketed, we have not discerned the causative input factors. Future work might attempt to correlate factors ranging from unmet medical need, market size and scientific understanding to funding availability, development processes, regulatory standards and economic incentives which, while not proof of causation, might nonetheless suggest the relative importance of these various factors.

The segmentation of oncology treatments creates both opportunities and challenges. The expanding number of late stage candidate oncology therapeutics addressing an increasing number of molecular targets provides growing hopes for patients and clinicians that one or more agents will match the specific characteristics of each patient’s condition. Additional benefits of such improved efficacy may also include faster, smaller clinical trials and higher success rates of candidate therapies in those trials, resulting in more rapid broad patient access and more effective care as fewer patients receive treatments that provide limited benefits. To the extent that the segmentation results in drugs having increased performance improvements over other treatments, the segmentation into more stratified medicines may facilitate premium pricing, reflecting in part the greater fixed costs of development of small-population drugs and their superior performance in treating life-threatening diseases, thereby incentivizing their development, although putting greater pressures on payers. Alternatively, the resulting segmentation of cancer into ever smaller population sub-types may eventually discourage future development of additional therapeutics as recent declines in early stage candidates may presage. Without changes in the procedures by which we develop, evaluate, disseminate and finance new medicines, even when science generates promising intervention hypotheses, if the patient sub-population falls below a certain threshold—whether a few thousand or a few hundred—it may prove clinically impractical and economically infeasible for developers to create the efficacy, safety and clinical benefit evidentiary data package required at drug prices payers can afford. To the extent science and clinical experience suggests combination treatments are preferable to monotreatments, the clinical evidence generation and economic feasibility challenges are potentially even more daunting.

30 Trusheim, Berndt and Douglas [2007]; Bach [2009].
31 Yin [2008].
32 Kummar, Chen, Wright, et al. [2010], Mayfield [2011], National Cancer Institute [2012].
7. References

The authors acknowledge funding support from Pfizer Inc. for this research. The opinions expressed herein are those of the authors, and not necessarily those of the research sponsor or the Massachusetts Institute of Technology.


Bloch RA [2004], “Cancer…there’s Hope”, R.A. Bloch Cancer Foundation, Kansas City, MO,


Finkelstein AM [2004], “Static and dynamic effects of health policy: evidence from the vaccine industry”, *Quart J Econ* 119(2):527-564

Frei III E, Karon M, Levin RH, et al. [1965], “The effectiveness of combinations of antileukemic


