

Research Update on P01 Specific Aim “Health Impacts of Medical Research”

Packalen and Bhattacharya

July 2015

Basic Science Roots of Medical Practice

Research Question

- What areas of basic science have contributed to advances in medicine?

Motivation

- Increased pressure to fund translational research
 - Can help make the case for funding basic research
- Some advances in medicine have come from unexpected basic science advances
 - Knowledge of thermophilic bacteria led to advance in PCR automation (which contributed to breast cancer treatments)

Analysis

1. What ideas have been relevant to medical practice?
2. Which applied biomedical research papers have introduced/explored those ideas?
3. Which basic research papers are linked to those applied papers?
 - What basic science areas dominate?
 - What basic science areas are funded?
 - What are the time delays?

Methodology

1. What ideas have been relevant to medical practice?
 - Harrison's Internal Medicine, 17 editions 1950-present
 - The #1 standard textbook of internal medicine training in medical school
 - Practice oriented; also used as a reference book by doctors
 - Contains ideas that are not just new but deemed most relevant to patient care (by editors...)

Methodology (cont.)

- Using textbooks alone
 - Index new N-grams in textbooks; more emphasis on N-grams mentioned more often
 - Index new citations in textbooks to biomedical literature (approx. 20% of citations to review articles => use also Thompson Reuters citation data)

Methodology (cont.)

1. What ideas have been relevant to medical practice?
 - Main approach: Using textbooks and a metathesaurus
 - Unified Medical Language System (“UMLS”) metathesaurus
 - 1M+ concepts
 - 5M+ concept names (synonyms)
 - Constructed from 150+ vocabularies
 - 127 concept categories (e.g. “Fungus”, “Bacterium”, “Disease or Syndrome”, “Medical Device”)

“Organ or Tissue Function”

C0302600 | angiogenic | angiogenesis | angiogenicprocess

“Amino Acid, Peptide or Protein”

**C1171892 | vascularendothelialgrowthfactor_A_human | vascularendothelialgrowthfactor | _V
EGFA_ | _VPF_ | _VEGF148_ | vascularendothelialgrowthfactorahuman | vascularendothelialgro
wthfactorhuman | _VEGF_ | vascularpermeabilityfactor | _VEGF_proteinhuman | _VEGFA_prote
inhuman | vascularendothelialgrowthfactor_A_**

“Pharmacologic Substance”

**C0796392 | _BEVACIZUMAB__UNIDENTIFIED_|_BEVACIZUMAB_|recombinanthumanizedanti
vegfm monoclonalantibody | antivegf | bevacizumabbiosimilar_BEVZ92_| bevacizumab | immuno
globulin_G1_humanmousemonoclonalrhumabvegfgammachainantihumanvascularendotheli
algrowthfactor disulfidewithhumanmousemonoclonalrhumabvegflightchaindimer | antivegf
monoclonalantibody | antivegfrhumab | monoclonalantibodyantivegf | antivegfhumanizedmo
noclonalantibody | moab_VEGF_|rhumabvegf**

“Clinical Drug”

**C1586267 | avastin25mgmlintravenoussolution | bevacizumab25_MGML_injectablesolution | a
vastin25_MGML_injectablesolution | bevacizumab25_MGML_intravenoussolution | _BEVACIZ**

“Bacterium”

C0995303 | thermusaquaticusbrockandfreeze1969 | thermusaquaticus

“Laboratory Procedure”

C0200931 | polymerasechainreaction | _PCR_analysis | polymerasechainreactionanalysis

Methodology (cont.)

- Main approach: Using textbooks and UMLS
 - Index which UMLS concepts appear in Harrison and when
 - Emphasis on concepts in treatment/diagnosis oriented categories such as
 - “Therapeutic or Preventative Procedure”
 - “Medical Device”
 - “Pharmacologic Substance”
 - “Laboratory Procedure”

Methodology (cont.)

2. Which applied biomedical research papers have introduced/explored those ideas that showed up in textbooks?
 - MEDLINE Data on 20M+ Biomedical Publications
 - Expect many applied papers to explore and develop each concept before appears in Harrison
 - Medical Subject Headings (“MeSH”) vocabulary
 - Each MEDLINE paper indexed with MeSH terms
 - Paper deemed applied if indexed with a “Diseases [C]” MeSH code

Methodology (con't.)

3. Which basic research papers are linked to those applied papers?
- MEDLINE, MeSH, THOMPSON REUTERS Citation Data
 - First determine which ideas appear in applied research papers that were linked to changes in medical practice;
 - N-grams
 - UMLS concepts in basic categories such as “Bacterium”, “Fungus”, “Amino Acid, Peptide or Protein”
 - Citations

Methodology (con't.)

- Next, determine which basic research papers linked to those ideas
 - By mentioning relevant N-grams
 - By mentioning relevant UMLS concepts
 - By citation

Methodology (cont.)

- Next, determine which basic research papers linked to those ideas
 - By mentioning relevant N-grams
 - By mentioning relevant UMLS concepts
 - By citation
- Then, determine frequency of such papers for each basic science area; areas defined by
 - MeSH code for biomedical papers
 - Thompson Reuters research fields?

Anatomy [A]

[Body Regions \[A01\] +](#)

[Musculoskeletal System \[A02\] +](#)

[Digestive System \[A03\] +](#)

[Respiratory System \[A04\] +](#)

[Urogenital System \[A05\] +](#)

[Endocrine System \[A06\] +](#)

[Cardiovascular System \[A07\] +](#)

[Nervous System \[A08\] +](#)

[Sense Organs \[A09\] +](#)

[Tissues \[A10\] +](#)

[Cells \[A11\] +](#)

[Fluids and Secretions \[A12\] +](#)

[Animal Structures \[A13\] +](#)

[Stomatognathic System \[A14\] +](#)

[Hemic and Immune Systems](#)

[\[A15\] +](#)

[Embryonic Structures \[A16\] +](#)

[Integumentary System \[A17\] +](#)

[Plant Structure \[A18\] +](#)

[Fungal Structure \[A19\] +](#)

[Bacterial Structure \[A20\] +](#)

[Viral Structure \[A21\] +](#)

Organisms [B]

[Eukaryota \[B01\] +](#)

[Archaea \[B02\] +](#)

[Bacteria \[B03\] +](#)

[Viruses \[B04\] +](#)

[Organism Forms \[B05\] +](#)

Phenomena and Processes [G]

[Physical Phenomena \[G01\] +](#)

[Chemical Phenomena \[G02\] +](#)

[Metabolic Phenomena \[G03\] +](#)

[Cell Physiological Phenomena \[G04\] +](#)

[Genetic Phenomena \[G05\] +](#)

[Microbiological Phenomena \[G06\] +](#)

[Physiological Phenomena \[G07\] +](#)

[Reproductive and Urinary Physiological Phenomena \[G08\]](#)

[Circulatory and Respiratory Physiological Phenomena \[G09\]](#)

[Digestive System and Oral Physiological Phenomena \[G10\]](#)

[Musculoskeletal and Neural Physiological Phenomena \[G11\]](#)

[Immune System Phenomena \[G12\] +](#)

[Integumentary System Physiological Phenomena \[G13\] +](#)

[Ocular Physiological Phenomena \[G14\] +](#)

[Plant Physiological Phenomena \[G15\] +](#)

[Biological Phenomena \[G16\] +](#)

Methodology: Recap

1. What ideas show up in textbooks?
2. Which applied papers linked to these ideas?
3. How do these papers link to basic science?
 - Citations from applied papers with textbook link to basic science papers
 - Text analysis:
 - Idea A is found in textbook and in an applied paper C
 - Idea A & B are both found in applied paper C → idea B complements idea A
 - Search for which basic science papers idea B showed up
 - Determine which areas of science idea B showed up most

Main Output

- Share of papers in each basic science area that are linked to clinical advances in mentioned in textbooks
- NIH funding levels by basic science area

Trends in Medical Knowledge

Trends Medicine: Background

- There has been much more progress in effectively treating some disease (e.g. heart disease) than others (e.g. some cancers)
 - Maybe medicine in general has stagnated (Le Fanu, 2000)
 - Debate in cancer about stagnation

Trends in Medicine: Motivation (cont.)

- Many policies, including Medicare and Social Security expenditures depend on expected changes in the effectiveness of medical treatment
 - If we knew what ‘stagnant’ medicine looks like, we could better predict future changes based on what current medicine looks like.

Trends in Medicine: Motivation (cont.)

- You can have lots of changes in research that never translate over to changes in medical practice or in health

Trends in Medicine: Research Question

- What are the characteristics of changes in medical knowledge in cancer vs. heart disease?
- Historically, what does this look like for other areas of medicine?
- How much persistence is there in medical treatment?
 - Jumps and starts or gradual?
 - Lag structure in textbook changes v. mortality changes relationship

Data

- Harrison's Internal Medicine, 17 editions
- Surgery
 - Bailey and Love 1932 – 2013 (26 editions)
- Medical Oncology
 - Cavalli and Kaye 1997 – 2014 (4 editions)
- Cardiology
 - Braunwald's Heart Disease 1980- 2014 (10 editions)

Data (cont.)

- MEDLINE Publications Data
- UMLS Metathesaurus
- Thompson Reuters Citation Data
- Mortality data
 - Multiple-cause of death file

Methodology: Linking textbooks to diseases

- Section or chapter headings in textbooks
- Tag paragraphs using UMLS thesaurus disease terms

Methodology: Metrics to Examine for Each Disease

- Number of paragraphs linked to the disease
- New n-grams as a fraction of total n-grams
- New concepts in in treatment/diagnosis oriented UMLS categories such as “Therapeutic or Preventative Procedure”, “Medical Device”
 - Textbooks may change for reasons that need not be relevant to patient care (e.g. a discussion of proteins)
- Citation patterns
 - Links to breakthrough papers
 - Recency of citations

Trends in Medicine: Output

- Extent of change in textbook content by clinical area over time
- Link between change in knowledge and mortality

Medical Research
(Publications)

Other Factors
(e.g. Technological
Inventions)

Medical Knowledge
(Textbooks, clinical guides)

Health Outcomes

