Supplemental Materials

Where are the breaks in translation from theory to clinical practice (and back) in addressing depression? An empirical graph-theoretic approach

Greg J. Siegle, Angélique O.J. Cramer, Nees Jan van Eck, Philip Spinhoven, Steven D. Hollon, Johan Ormel, Marlene Strege, Claudi L. H. Bockting

Table of Contents

Supplement 1: Specifications for the Google News search

Supplement 1a: Search terms for study 1

Terms were derived from any terms containing word fragments:
'depress','mental_','mentally', 'psychol','psychia','neuros', 'neuropsy', 'therapy','patient'};
hand edited from 165 down to 101 to remove clearly bad terms like "fundamentally" (contains the search term “mental”) and "great depression".

ANTIDEPRESSANTS
Abnormal_Psychology
Antidepressant
Antidepressant_Drug
Applied_Psychology
Aromatherapy
Aversion_therapy
Behavior_Therapy
Clinical_Psychologist
Clinical_Psychologists
Clinical_Psychology
Cognitive_Neuroscience
Cognitive_Neurosciences
Cognitive_Psychology
Comparative_Psychology
DEPRESSED
DEPRESSING
DEPRESSION
Depress
Depressants
Depressingly
Depressions
Depressive
Depressive_Disorder
Depressive_Disorders
Depressives
Developmental_Psychology
Developmentally
Electroconvulsive_Therapy
Electroshock_therapy
Experimental_Psychology
Freudian_psychology
GROUP_THERAPY
Herbal_Therapy
Hydrotherapy
Hypnotherapy
Inpatient
Inpatients
Jungian_psychology
MENTALLY
MENTALLY_ILL
MENTAL_HEALTH
MENTAL_HOSPITAL
MENTAL_ILLNESS
Manic_Depression
Manic_Depressive_Illness
Mental_Anguish
Mental_Attitude
Mental_Capacity
Mental_Disease
Mental_Diseases
Mental_Disorder
Mental_Disorders
Mental_Hygiene
Mental_Illnesses
Mental_Institution
NEUROSCIENCE
NEUROSIS
NeuroSciences
Neuropsychiatric
Neuropsychiatry
Neuropsychological
Neuropsychology
Neuroscientist
Neuroscientists
Neuroses
OUTPATIENT
Occupational_Therapy
Outpatients
PATIENT
PATIENTS
PHYSICAL_THERAPY
PSYCHIATRIC
PSYCHIATRIST
PSYCHOLOGICAL
PSYCHOLOGIST
PSYCHOLOGY
Psychiatric_Hospital
Psychiatrists
Psychiatry
Psycholinguistics
Psychological_Disorders
Psychological_Medicine
Psychological_Science
Psychologically
Psychologies
Psychologists
Psychotherapy
Psychotic_depression
Shock_Therapy
Social_Psychology
Speech_Therapy
THERAPY
Temperamentally
Tricyclic_antidepressants
antidepressant_drugs
anxiety_neurosis
cognitive_neuroscientist
cognitive_neuroscientists
depressant
depresses
electrotherapy
mental_abnormality
mental_faculties
tricyclic_antidepressant
unipolar_depression

Supplement 1b: Calculation of metrics for the Google News Corpus

Corpus, word similarity weight assignment, and network parameters. The Google News corpus included ~1 billion words, with 200,000 words and phrases in common dictionaries and coded for content (https://code.google.com/archive/p/word2vec/#Pre-trained_word_and_phrase_vectors). We used the word2vec algorithm (Mikolov, T, Chen, K, Corrado, G, Dean, J 2013), which coded words on 300 algorithmically derived features and computed the similarity between each word pair as the cosine between the feature vectors. This coding yielded a contextual similarity metric (similarity of contexts in which words were mentioned).

Network layout. We obtained the network layout with the VOSviewer software (van Eck & Waltman 2010) (www.vosviewer.com) setting attraction and repulsion parameters to 1 and -1, respectively. Links were thresholded as "connected" if they had similarities of at least -2 standard deviations below the mean among the searched terms; see Supplement 7 for advantages and disadvantages of our use of this strategy.

Network structure metrics. Small world-ness was calculated as the ratio of mean shortest paths from one node to another in the network divided by the same index for a similar-sized random network (Humphries & Gurney 2008), via the SBEToolbox (Konganti et al. 2013). **Fractal dimensionality** was calculated using the Hausdorff dimensionality (Hausdorff 2001) for the "strong" connections in the adjacency matrix. This number represents the complexity of a binary image, ranging from 1 (simple line through an image) to 2 (random patterns).

Grouping terms: Terms were grouped into clusters using VOSviewer software clustering, a weighted variant of modularity-based clustering (Waltman et al. 2010) (resolution = 1, minimum cluster size = 5 items).

Metrics for inefficient communication between nodes. We used the SBEToolbox (Konganti et al. 2013) to estimate nodewise centrality and clustering values. *Degree centrality* represented the number of connections a node has to other terms. The *clustering coefficient* represented the extent to which a node clusters with other nodes. *Local average connectivity* reflected the extent to which a node is connected to its neighbors. Nodes that were low (<1 std below mean) on all three were considered orphans, with little connection to the rest of the graph. Nodes that were low (< 1 std below the mean) on any measure were considered "at risk". For groups of similar terms differing in tense (e.g., "antidepressant", "antidepressants") or specification of an irrelevant element (e.g., "antidepressant", "antidepressant drug"), if one term was not an orphan or "at risk", the whole group was given that status.

"Communication" and "lost in translation" indices. We considered terms not mentioned in conjunction with other modules to be at risk for being "lost in translation". We computed its "stabilizing" index (# connections within module) for each node versus its "communicating" index (# connections between modules) with each of the other modules. We computed a novel "lost in translation" index as the ratio of communicating connections to the stabilizing+communicating connections sum, with respect to each other module.

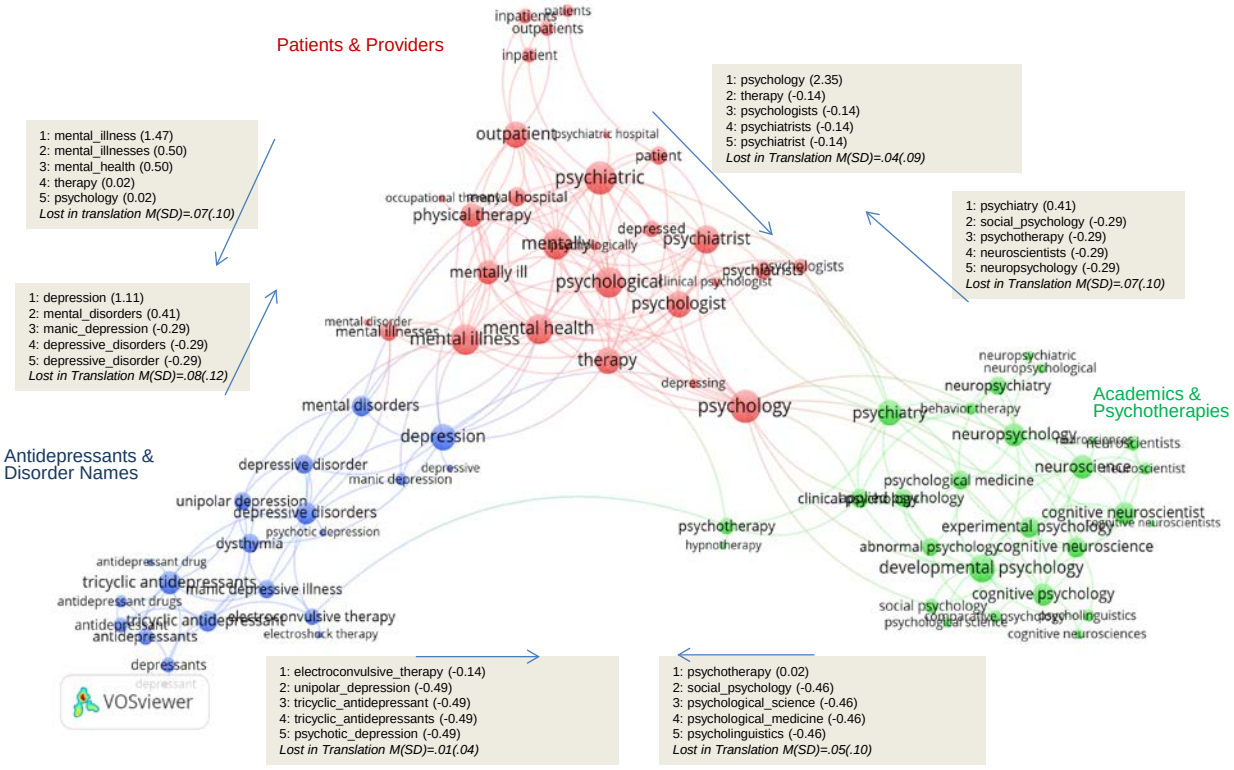
Metrics for determining communicative/bridging terms. We calculated the *Betweenness centrality*, the number of shortest paths through the network that involve a given node, representing the node's influence over the network. *Bridging centrality* represented the extent to which a node connects densely connected subsets of the graph, if it functions critically in propagating information from one area of the network to another. Nodes that were high (>1 std above the mean) on any property and had at least the mean number of connections (degree centrality >6) were considered "influential" and worthy of consideration as bridging terms. We also calculated the *brokering coefficient*, reflecting the extent to which a node connects nodes not otherwise connected in the network (Cai et al. 2010). This index performed largely as the

orphan/risk indices (because it accounted for nodes with low connectivity) and thus was analyzed separately. We calculated a “brokering influential” score as items scored as “influential” and also high (>1 std above the mean) in brokering. We considered nodes with high numbers of between-module connections as potential targets of remediating communication between modules (hence “communicative” or “bridge” nodes, (Cramer et al. 2010)).

Network quality: We considered words strongly associated with depression in the corpus regardless of whether they met our criteria.

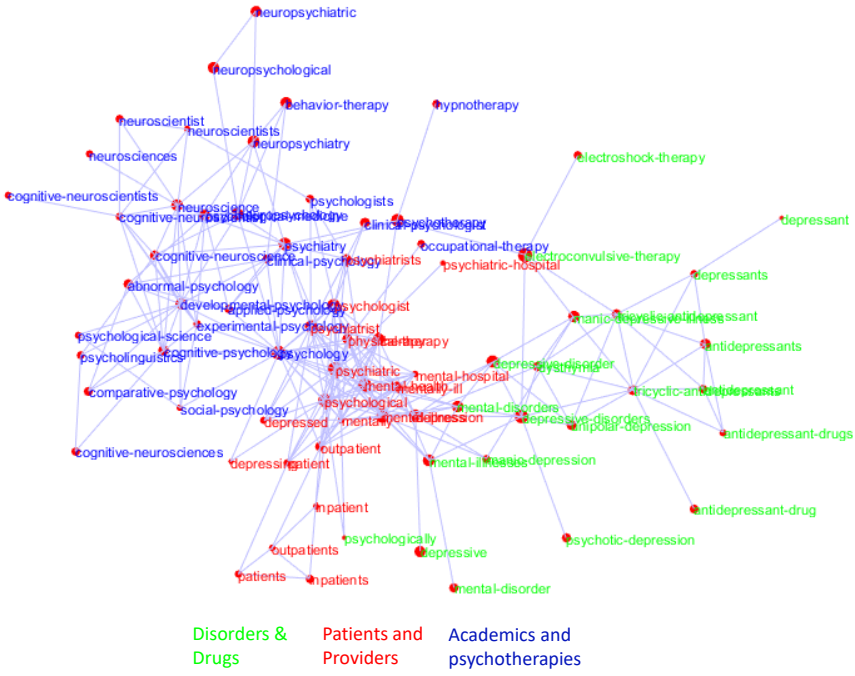
Supplement 2: Google News Network

Supplement Figure 2a. Color version of Figure 1, the Google News Network



Supplement Figure 2b. Modules derived using a second clustering algorithm

To examine the robustness of the VosViewer clustering solution we also clustered the giant component from the Google News Corpus using the well-known Louvain algorithm (Blondel et al. 2008) for modularity optimization. The layout was obtained using the Kamada-Kawai layout technique (Kamada and Kawai 1989). The visualization was obtained using the Pajek software package (de Nooy, Mrvar, and Batagelj 2011). Terms outside the giant component were the same as for VosViewer. Only 8 terms did not have the same cluster assignment in the two clustering approaches.



Supplement 3: Term-wise statistics

Supplement Table 3-1. Degree centrality, clustering, and node centrality statistics for each term with at least one connection.

Adjacent highlighted terms are considered groups of terms with identical meaning. Orphan and risk values which meet criteria described in the text but are part of a group of terms where at least one term is not an orphan or at-risk term are given values of 0.1.

	Detecting broken piplelines					Restoring broken piplelines				
Term	Degree Centrality	Clustering Coefficient	Local average Connectivity	Orphan	Risk	Betweenness Centrality	Bridging Centrality	Influential	Brokering Coefficient	Brokering Influential
abnormal psychology	6	0.533333	1.33333	0	0	17.1187	0.00047	0	-0.35512	0
antidepressant	4	0.833333	1.25	0	0	3.98108	0.00021	0	-0.55735	0
antidepressant drug	1	0	0	0.1	0.1	0	0	0	0.012423	0
antidepressant drugs	3	1	1	0	0	0	0	0	-0.65633	0
antidepressants	5	0.7	1.4	0	0	7.31441	0.00022	0	-0.47	0
applied psychology	6	0.666667	1.66667	0	0	29.6748	0.0012	0	-0.43851	0
behavior therapy	3	0.333333	0.333333	0	0	3	0.00039	0	-0.25087	0
clinical psychologist	2	1	0.5	0	0	0	0	0	-0.66846	0
clinical psychology	7	0.380952	1.14286	0	0	184.487	0.00391	0	-0.23889	0
cognitive neuroscience	8	0.571429	2	0	0	56.0304	0.00108	0	-0.35668	0
cognitive neurosciences	2	0	0	0	0.1	0.666667	0.00012	0	0.024693	0
cognitive neuroscientist	9	0.444444	1.77778	0	0	204.102	0.00163	0	-0.26112	0
cognitive neuroscientists	1	0	0	0.1	0.1	0	0	0	0.012423	0
cognitive psychology	9	0.416667	1.66667	0	0	283.682	0.00318	0	-0.2417	0
comparative psychology	3	1	1	0	0	0	0	0	-0.65633	0
depress	1	0	0	0.1	0.1	0	0	0	0.012423	0
depressant	1	0	0	0.1	0.1	0	0	0	0.012423	0
depressants	4	0.5	0.75	0	0	146	0.00405	0	-0.35668	0
depressed	5	1	2	0	0	0	0	0	-0.63252	0
depresses	1	0	0	0.1	0.1	0	0	0	0.012423	0
depressing	2	1	0.5	0	0	0	0	0	-0.66846	0
depression	13	0.358974	2.15385	0	0	966.416	0.00508	1	-0.15616	0
depressive	1	0	0	0.1	0.1	0	0	0	0.012423	0
depressive disorder	7	0.52381	1.57143	0	0	102.856	0.00274	0	-0.33733	0
depressive disorders	9	0.305556	1.22222	0	0	545.768	0.00458	1	-0.16002	0
developmental psychology	15	0.304762	2.13333	0	0	448.628	0.00172	1	-0.09417	0
dysthymia	7	0.571429	1.71429	0	0	328.66	0.00835	1	-0.3681	0
electroconvulsive therapy	5	0.1	0.2	0	1	421.583	0.00832	0	-0.03469	0
electroshock therapy	1	0	0	0.1	1	0	0	0	0.012423	0
experimental psychology	8	0.535714	1.875	0	0	93.2535	0.00195	0	-0.33369	0
hypnotherapy	1	0	0	1	1	0	0	0	0.012423	0
inpatient	4	0.666667	1	0	0	21.4813	0.00135	0	-0.46204	0
inpatients	4	0.666667	1	0	0	8.76404	0.00038	0	-0.46204	0
manic depression	3	0.333333	0.333333	0	0	44.0113	0.00768	0	-0.25087	0
manic depressive illness	6	0.333333	0.833333	0	0	66.153	0.00167	0	-0.21536	0
mental disease	1	0	0	0.1	0.1	0	0	0	0.012423	0
mental diseases	1	0	0	0.1	0.1	0	0	0	0.012423	0
mental disorder	1	0	0	0.1	0.1	0	0	0	0.012423	0
mental disorders	8	0.535714	1.875	0	0	234.368	0.00571	1	-0.33369	0
mental health	16	0.591667	4.4375	0	0	171.169	0.0012	0	-0.28246	0

mental hospital	7	1	3	0	0	0	0	0	-0.60927	0
mental illness	17	0.522059	4.17647	0	0	366.862	0.00198	0	-0.22738	0
mental illnesses	7	0.47619	1.42857	0	0	232.427	0.00351	0	-0.30558	0
mentally	16	0.583333	4.375	0	0	168.021	0.0009	0	-0.27721	0
mentally ill	11	0.709091	3.54545	0	0	47.5626	0.00083	0	-0.40713	0
neuropsychiatric	2	0	0	0	1	1	0.00012	0	0.024693	0
neuropsychiatry	6	0.4	1	0	0	139.749	0.00295	0	-0.26415	0
neuropsychological	2	0	0	0	1	15.9079	0.0021	0	0.024693	0
neuropsychology	10	0.355556	1.6	0	0	344.166	0.00326	0	-0.18643	0
neuroscience	11	0.363636	1.81818	0	0	339.747	0.00198	0	-0.18132	0
neurosciences	1	0	0	0.1	0.1	0	0	0	0.012423	0
neuroscientist	3	1	1	0	0	0	0	0	-0.65633	0
neuroscientists	4	0.5	0.75	0	0	36.2528	0.00183	0	-0.35668	0
neuroses	1	0	0	1	1	0	0	0	0.012423	0
neurosis	1	0	0	1	1	0	0	0	0.012423	0
occupational therapy	1	0	0	1	1	0	0	0	0.012423	0
outpatient	13	0.564103	3.38462	0	0	308.903	0.00252	0	-0.29674	0
outpatients	4	0.666667	1	0	0	8.76404	0.00038	0	-0.46204	0
patient	6	0.666667	1.66667	0	0	125.385	0.0051	0	-0.43851	0
patients	3	0.333333	0.333333	0	0	4.91356	0.00039	0	-0.25087	0
physical therapy	11	0.727273	3.63636	0	0	155.435	0.00129	0	-0.41771	0
psychiatric	20	0.4	3.8	0	0	679.252	0.00176	1	-0.11333	0
psychiatric hospital	1	0	0	1	1	0	0	0	0.012423	0
psychiatrist	15	0.485714	3.4	0	0	272.496	0.00147	0	-0.22405	0
psychiatrists	6	0.666667	1.66667	0	0	27.9974	0.00125	0	-0.43851	0
psychiatry	12	0.272727	1.5	0	0	740.361	0.00665	1	-0.1014	0
psycholinguistics	3	0.333333	0.333333	0	0	9.08175	0.00069	0	-0.25087	0
psychological	17	0.522059	4.17647	0	0	219.837	0.00087	0	-0.22738	0
psychological medicine	7	0.428571	1.28571	0	0	25.7314	0.00066	0	-0.27279	0
psychological science	2	1	0.5	0	0	0	0	0	-0.66846	0
psychologically	2	1	0.5	0	0	0	0	0	-0.66846	0
psychologist	13	0.564103	3.38462	0	0	147.745	0.00106	0	-0.29674	0
psychologists	4	0.5	0.75	0	0	80.8494	0.00571	0	-0.35668	0
psychology	20	0.321053	3.05	0	0	1609.41	0.00553	1	-0.05529	0
psychotherapy	5	0.1	0.2	0	1	540.356	0.01138	0	-0.03469	0
psychotic depression	1	0	0	1	1	0	0	0	0.012423	0
social psychology	4	0.666667	1	0	0	45.546	0.00243	0	-0.46204	0
therapy	13	0.641026	3.84615	0	0	221.914	0.00246	0	-0.34475	0
tricyclic antidepressant	8	0.392857	1.375	0	0	299.99	0.00402	0	-0.23605	0
tricyclic antidepressants	10	0.311111	1.4	0	0	545.606	0.00314	1	-0.15309	0
unipolar depression	6	0.6	1.5	0	0	97.5836	0.00345	0	-0.39768	0

Supplement Table 3-2: Connectivity coefficients for the network of depression-related disciplines

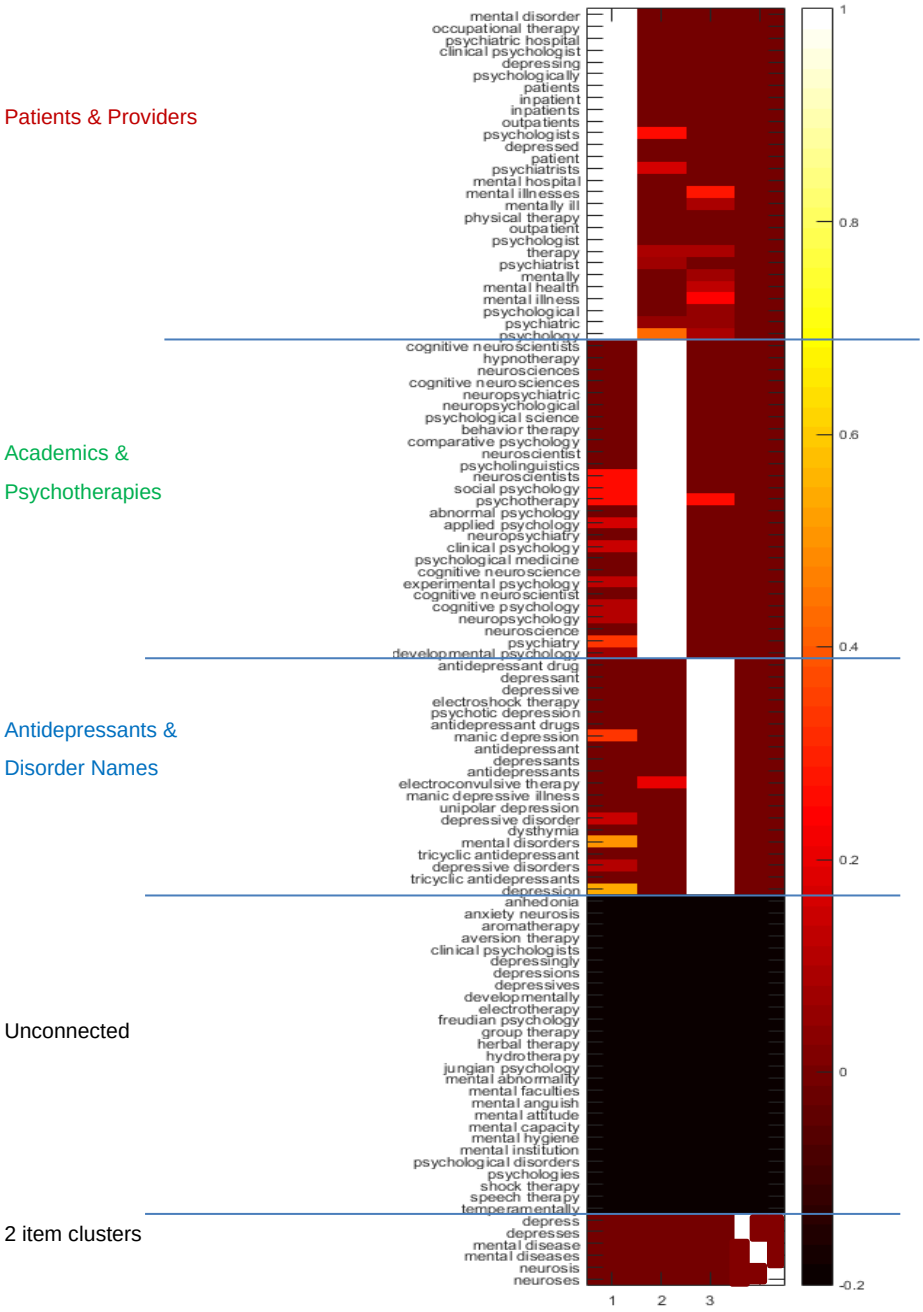
	Detecting broken piplelines					Restoring broken piplelines				
Term	Degree Centrality	Clustering Coefficient	Local average Connectivity	Orphan	Risk	Betweenness Centrality	Bridging Centrality	Influential	Brokering Coefficient	Brokering Influential
anesthesiology	61	0.92459	27.7377	0	0	6.6203	2.21E-05	0	-0.028011	0
anthropology	53	0.955733	24.8491	0	0	3.04463	1.39E-05	0	-0.107076	0
behavioral sc	67	0.854817	28.209	0	0	19.6694	5.07E-05	0	0.0536997	0
biochem&mol biol	66	0.84662	27.5152	0	0	21.5476	5.64E-05	1	0.0508027	0
biology	62	0.868324	26.4839	0	0	15.6054	4.72E-05	0	0.0092651	0
card&cardiov sys	64	0.880456	27.7944	0	0	14.0423	4.06E-05	0	0.0178301	0
chem,medicinal	48	0.955674	22.4583	0	0	2.83838	1.56E-05	0	-0.148545	0
clin neurology	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
comp sc,ai	28	0.970899	13.1071	0	1	0.498909	8.39E-06	0	-0.342018	0
crit care medic	57	0.946115	26.4912	0	0	4.14005	1.61E-05	0	-0.0701434	0
economics	53	0.952104	24.7547	0	0	3.74885	1.69E-05	0	-0.105219	0
educ,scient disc	56	0.933117	25.6607	0	0	6.19605	2.45E-05	0	-0.071347	0
educat,special	53	0.921626	23.9623	0	0	6.51428	2.85E-05	0	-0.0894824	0
endocrin&metabol	66	0.870396	28.2879	0	0	15.8375	4.29E-05	0	0.0380093	0
eng,biomedical	47	0.938945	21.5957	0	0	3.92534	2.22E-05	0	-0.148466	0
ethnic studies	48	0.952128	22.375	0	0	3.17314	1.74E-05	0	-0.146731	0
evolut biology	34	0.934046	15.4118	0	1	2.46867	2.65E-05	0	-0.263719	0
family studies	65	0.866346	27.7231	0	0	17.3221	4.78E-05	0	0.0327969	0
genetics&heredit	63	0.896057	27.7778	0	0	10.8911	3.31E-05	0	0.00207724	0
geriatr&gerontol	67	0.859792	28.3731	0	0	17.9152	4.66E-05	0	0.0510211	0
gerontology	66	0.87366	28.3939	0	0	14.8512	4.04E-05	0	0.0362661	0
hlth care sc&ser	69	0.83035	28.2319	0	1	25.801	6.10E-05	1	0.0814717	1
hlth pol&serv	69	0.830776	28.2464	0	1	25.7312	6.09E-05	1	0.0812389	1
humanities,mult	26	0.938462	11.7308	0	1	1.31597	2.42E-05	0	-0.346042	0
immunology	60	0.917514	27.0667	0	0	8.41908	2.86E-05	0	-0.0319904	0
infec disease	54	0.959469	25.4259	0	0	2.51984	1.11E-05	0	-0.100887	0
integr&compl med	59	0.93571	27.1356	0	0	5.31811	1.91E-05	0	-0.0491571	0
medical ethics	35	0.998319	16.9714	0	1	0.0350877	3.87E-07	0	-0.286841	0
medical informat	63	0.87148	27.0159	0	0	15.7563	4.64E-05	0	0.0151244	0
medicine,gen&int	69	0.830776	28.2464	0	1	25.7675	6.10E-05	1	0.0812389	1
medicine,legal	34	0.989305	16.3235	0	1	0.255434	2.96E-06	0	-0.29189	0
medicine,res&exp	67	0.859792	28.3731	0	0	17.9152	4.66E-05	0	0.0510211	0
multidiscipl sc	69	0.830776	28.2464	0	1	25.7312	6.09E-05	1	0.0812389	1
neuroimaging	54	0.932215	24.7037	0	0	5.42692	2.30E-05	0	-0.0868808	0
neurosciences	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
nursing	65	0.874519	27.9846	0	0	16.1012	4.48E-05	0	0.0284273	0
nutrition&diet	61	0.906011	27.1803	0	0	10.0082	3.26E-05	0	-0.0183105	0
obstetrics&gynec	61	0.926776	27.8033	0	0	6.43362	2.15E-05	0	-0.029146	0
oncology	63	0.907322	28.127	0	0	9.20477	2.84E-05	0	-	0

ophthalmology	46	0.972947	21.8913	0	0	1.22119	7.54E-06	0	-0.174433	0
pediatrics	66	0.868998	28.2424	0	0	16.4558	4.44E-05	0	0.0387574	0
periphl vasc dis	54	0.904263	23.963	0	0	9.27334	3.82E-05	0	-0.0723086	0
pharmacol&pharma	69	0.83035	28.2319	0	1	25.801	6.10E-05	1	0.0814717	1
philosophy	25	0.96	11.52	0	1	0.906538	1.86E-05	0	-0.367563	0
physiology	57	0.926065	25.9298	0	0	6.62651	2.51E-05	0	-0.0597873	0
prim hlth care	64	0.88244	27.7969	0	0	13.7002	3.97E-05	0	0.0167755	0
psychiatry	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
psychol,appl	61	0.883607	26.5082	0	0	13.6001	4.31E-05	0	-0.00648624	0
psychol,biolog	54	0.959469	25.4259	0	0	2.54787	1.12E-05	0	-0.100887	0
psychol,clinic	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
psychol,develop	68	0.840211	28.1471	0	1	23.5236	5.78E-05	1	0.0688784	0
psychol,educat	57	0.890351	24.9298	0	0	11.2922	4.12E-05	0	-0.0410706	0
psychol,exper	65	0.857212	27.4308	0	0	19.5205	5.31E-05	0	0.0377033	0
psychol,mathema	25	0.946667	11.36	0	1	0.820727	1.72E-05	0	-0.360737	0
psychol,multid	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
psychol,psanal	63	0.869944	26.9683	0	0	16.6916	4.88E-05	0	0.0159456	0
psychol,social	68	0.822651	27.5588	0	1	27.6747	6.65E-05	1	0.0784662	0
publ,env&occ hlt	70	0.816977	28.1857	0	1	29.828	6.75E-05	1	0.0959729	1
rehabilitation	64	0.881944	27.7813	0	0	14.656	4.23E-05	0	0.017039	0
religion	48	0.955674	22.4583	0	0	3.13313	1.72E-05	0	-0.148545	0
soc sc,biomedic	68	0.834504	27.9559	0	1	24.978	6.09E-05	1	0.0719843	0
soc sc,interdis	61	0.891803	26.7541	0	0	12.3745	3.94E-05	0	-0.0108284	0
soc sci,math m	35	0.961345	16.3429	0	1	1.25653	1.33E-05	0	-0.268165	0
social issues	51	0.950588	23.7647	0	0	4.03488	1.94E-05	0	-0.120836	0
social work	64	0.880456	27.7344	0	0	14.972	4.31E-05	0	0.0178301	0
sociology	59	0.90941	26.3729	0	0	9.1878	3.20E-05	0	-0.035477	0
statistics&proba	37	0.978979	17.6216	0	1	0.70832	6.81E-06	0	-0.258247	0
substance abuse	64	0.891369	28.0781	0	0	11.6649	3.43E-05	0	0.0120436	0
surgery	65	0.864904	27.6769	0	0	17.2225	4.75E-05	0	0.03357	0
virology	46	0.978744	22.0217	0	0	0.913792	5.65E-06	0	-0.177367	0
womens studies	60	0.920339	27.15	0	0	7.6826	2.62E-05	0	-0.0334625	0

Mean	57.54929577	0.898001592	25.13904225	Mean	12.45070222	3.60246E-05	-0.04541011
Std	12.22268235	0.051623978	4.553458942	Std	9.457551629	1.92486E-05	0.125811124
low cutoff	45.32661343	0.846377614	20.57958331	High cutoff	21.90825385	5.52732E-05	0.080401017

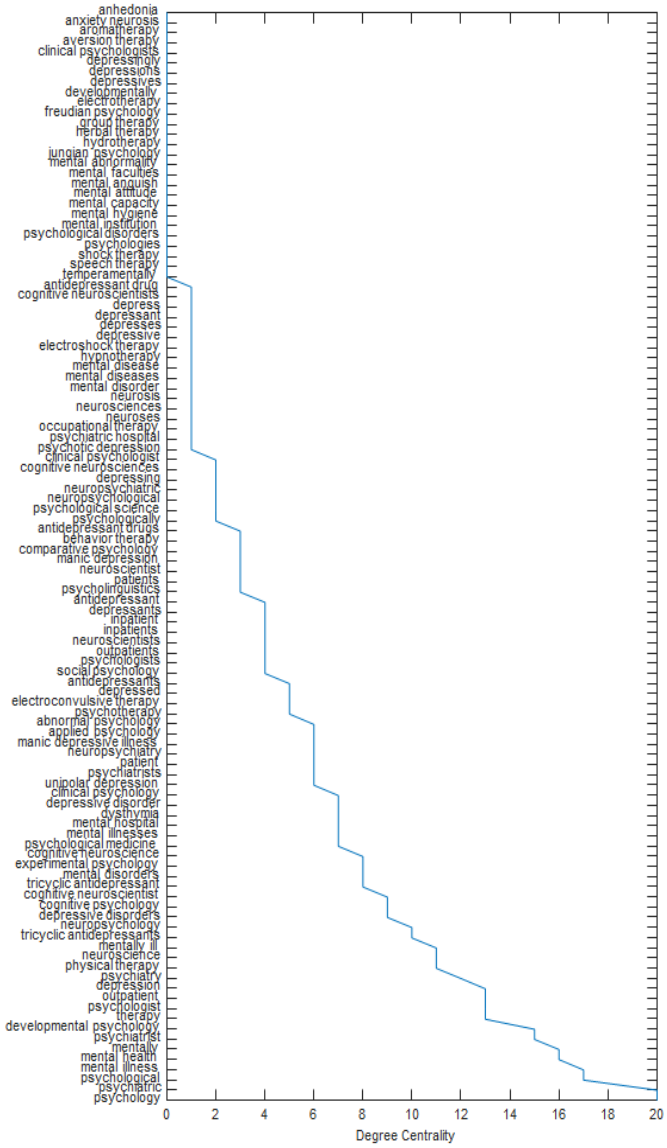
Supplement Figure 3-1: Lost-in-translation indices for each term, from each module to each other module.

Low levels of between module connectivity (deep red) are shown for nearly all terms.



Supplement Figure 3-2: Sorted degree centrality (# of connections) for each term in the Google News network

The low number of connections in general, particularly for terms commonly referenced in association with depression in academia (e.g., anhedonia has zero connections) could suggest specific breaks in translation.



Supplement 4: Analysis of a priori terms for the Google News corpus.

Motivation. To further understand how search terms of specific interest from the Google News corpus were communicating, we examined whether they were co-occurring in the same sources, or in sources that would lead to knowledge about both terms. For this we considered both the shortest path in the network to get from one term to the other, and how many routes between the terms there were with that shortest path length.

Method

Network metrics for a priori terms: Word- and word-pair based statistics including shortest path between terms, number of edges with shortest path distance, and 1-step topological overlap (via (Rubinov and Sporns 2010) were computed for 9 *a priori* terms chosen by GS as representative of domains-of-interest in which communication-breakdowns might be detected, from 101 term list. These included: “depression”,”patient”, “psychologist”, “psychology”, “psychiatrist”, “psychiatry”,’ “neuroscience”, “antidepressant”, “psychotherapy”. For targets for bridging, we also looked at the topological similarity (i.e., similarity of network links) between our *a priori* terms, to see whether influential terms were similarly connected or might have different influences on the network from each other.

Result.

Question 3: Are there terms that do not communicate strongly? Weak Paths: Table S2-1 shows bivariate data for paths between a priori terms grouped by module, z-scored with respect to all bivariate paths between terms in the giant-component network.

Table S2-1a shows the number of shortest paths between all of the chosen a priori terms; all were within 2 standard deviations of the network’s mean, though “psychology” was particularly well connected (Z>1.5) to “psychotherapy”, “psychiatry”, and “psychologist”, and “depression” was well connected only to “patient” with Z<-.5 connections to “psychologist”,”psychiatry”, and “psychotherapy”. Similarly, the lowest number of short paths (Z=-1.21) were between “psychologist”, “psychiatry”, “psychotherapy”, and “psychiatrist”. “Neuroscience” had approximately average connectivity to every examined term, with among the network’s stronger connections to “patient”.

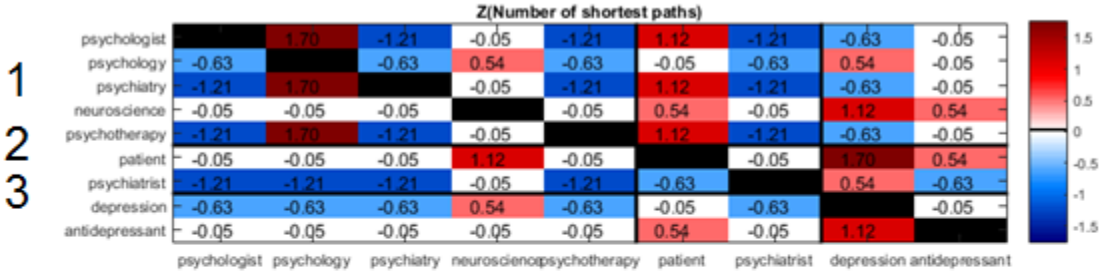
Table S2-1b shows the analogous graph for the distance of the shortest path between the a priori terms; shorter paths are highlighted in blue and longer paths are highlighted in red. As shown, “psychology”, “depression”, and “antidepressant” all had comparatively long “shortest paths” to the rest of the network. “Patient”, “psychiatrist”, and “neuroscience” had some of the shortest paths.

Topological Similarity (TS). Of the *a priori terms*, “psychology”, “psychiatry”, and “depression” were all influential. Table 4 shows the topological similarity between the *a priori* terms - the low values (most below 0.5) suggest there was strong differentiation among the connectivity of examined terms. As shown, some of the terms had similar networks (e.g., “psychiatry” and “psychology”, TS=.76). The network for depression was different from all of the other terms (8 TS’s<.3, TS psychology=.41) and the network for patients also diverged from ⅞ terms (TS<.4) with the lone similarity to “psychiatrist”. Thus increasing associations of depression specifically with psychiatry and psychology could have strong influence on the network.

Supplement Table S4-1: Z-scores for network inefficiencies

Scores are given for the number of shortest path edges between nodes and the length of the shortest path. Terms are grouped by module within the giant component.

Number of shortest paths between nodes



Shortest path distance

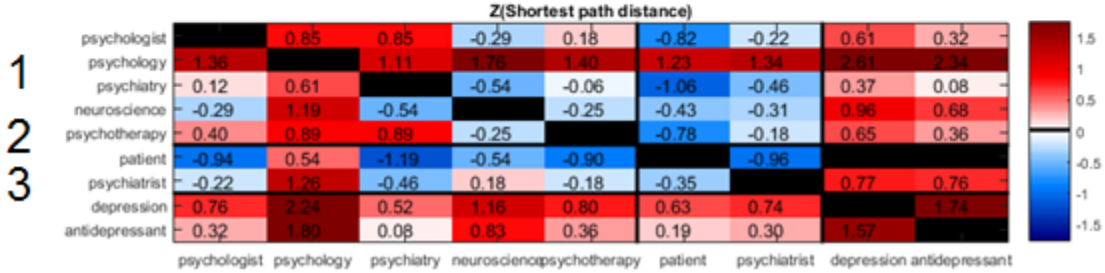
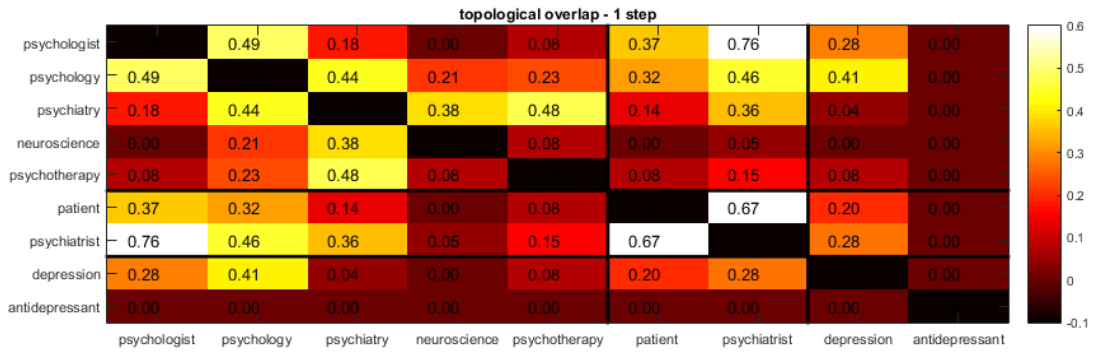


Table S4-2: Network similarities for potential use in increasing communication between domains in the Google News corpus.

Terms are grouped by module within the giant component.



Supplement 5: Terms used in the World of Science search

Supplement Table 5-1: Mental Health Science subdisciplines for the discipline-wise search

1. Psychiatry
2. Neurosciences
3. Clinical neurology
4. Pharmacology & Pharmacy
5. Psychology, Multidisciplinary
6. Psychology, Clinical
7. Medicine, General & Internal
8. Public, Environmental & Occupational Health
9. Geriatrics and Gerontology
10. Gerontology
11. Cardiac & Cardiovascular Systems
12. Psychology, Developmental
13. Health Care Sciences & Services
14. Pediatrics
15. Nursing
16. Oncology
17. Endocrinology & Metabolism
18. Multidisciplinary Sciences
19. Behavioral Sciences
20. Rehabilitation
21. Biochemistry & Molecular Biology
22. Surgery
23. Anesthesiology
24. Physiology
25. Health Policy & Services
26. Medicine, Research & Experimental
27. Obstetrics & Gynecology
28. Psychology, Social
29. Substance Abuse
30. Genetics & Heredity
31. Social Sciences, Biomedical
32. Peripheral Vascular Disease
33. Immunology
34. Family Studies
35. Nutrition & Dietetics
36. Primary Health Care
37. Biology
38. Critical Care Medicine
39. Social Work
40. Economics
41. Integrative & Complementary Medicine
42. Psychology, Experimental

43. Chemistry, Medicinal
44. Psychology, Psychoanalysis
45. Evolutionary Biology
46. Psychology, Applied
47. Womens Studies
48. Neuroimaging
49. Infectious Diseases
50. Social Sciences, Interdisciplinary
51. Ophthalmology
52. Sociology
53. Psychology, Educational
54. Psychology, Biological
55. Education, Special
56. Medical Informatics
57. Engineering, Biomedical
58. Anthropology
59. Social Issues
60. Education, Scientific Disciplines
61. Humanities, Multidisciplinary
62. Virology
63. Religion
64. Computer Science, Artificial Intelligence
65. Medicine, Legal
66. Statistics & Probability
67. Ethnic Studies
68. Social Sciences, Mathematical Methods
69. Psychology, Mathematical
70. Medical Ethics
71. Philosophy

Supplement Table 5-2: List of filtering terms for articles from the discipline-wise search

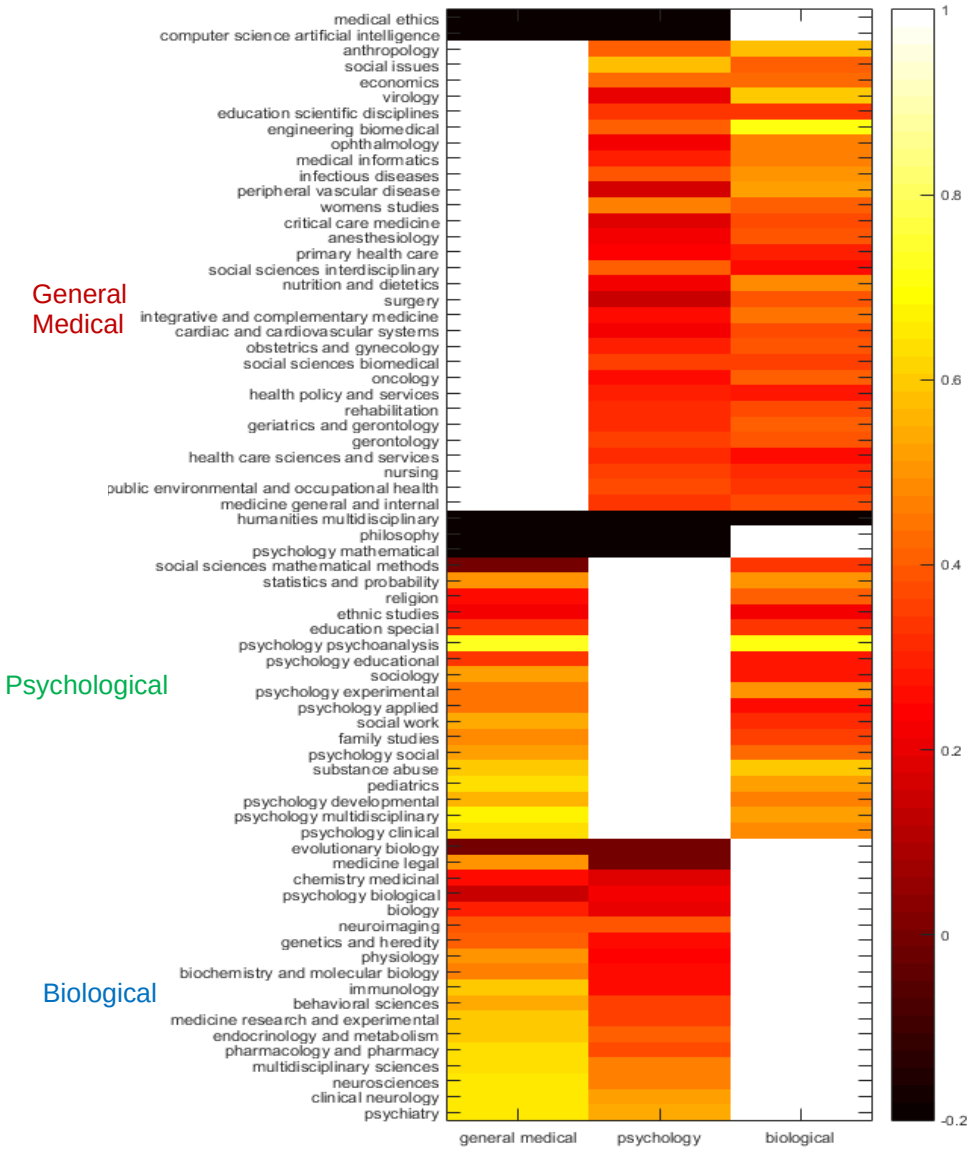
1. Respiratory depression
2. Synaptic depression
3. Long term depression
4. Depression in the context of bone fractures
5. Point depression
6. Shoulder depression
7. Myocardial depression
8. Metabolic depression
9. Depressor muscles
10. Depressed * scar
11. Depressed myofilament
12. Depressed cardiac
13. ST-depression
14. Depressed infection
15. Depressed * insulin
16. Depressed cognitive alertness
17. Depressed nutritional
18. Depressed mental status
19. Depressed * count
20. Depressed ejection fraction
21. Depressed ventricular function
22. Depressed contraction
23. Depressed stroke
24. Depressed * growth
25. Depressed * rate
26. Depressed current
27. Inbreeding depression
28. Cortical spreading depression
29. Karst depressions
30. River * depression
31. Depressed air
32. Depression years
33. Segment depression
34. Key-depression
35. Great depression

* Means “within three words”

Supplement 6: World of Science discipline-wise search

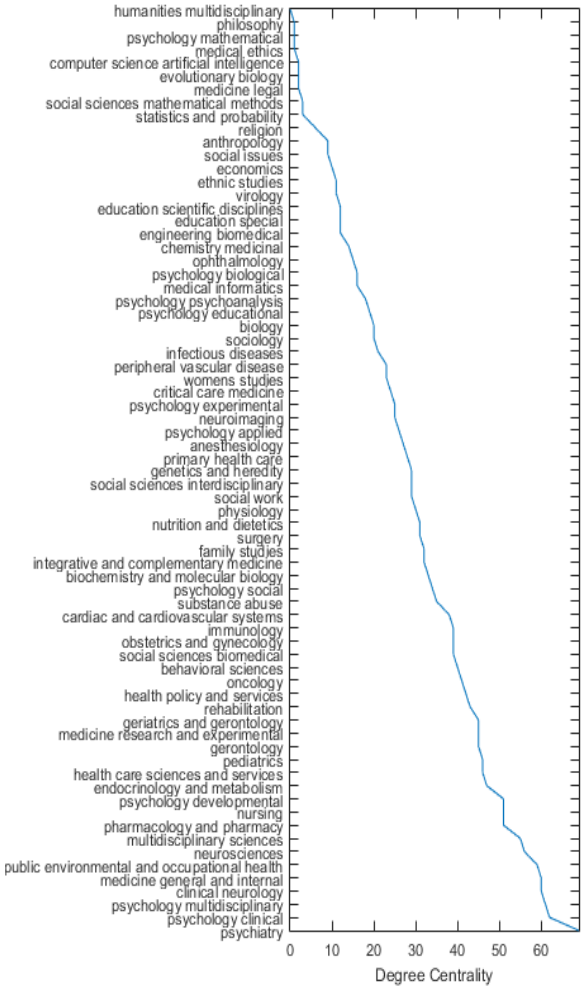
Supplement Figure 6-1. Lost in translation indices for each term in each module to each other module

Moderate levels of between module connectivity are shown for most terms.

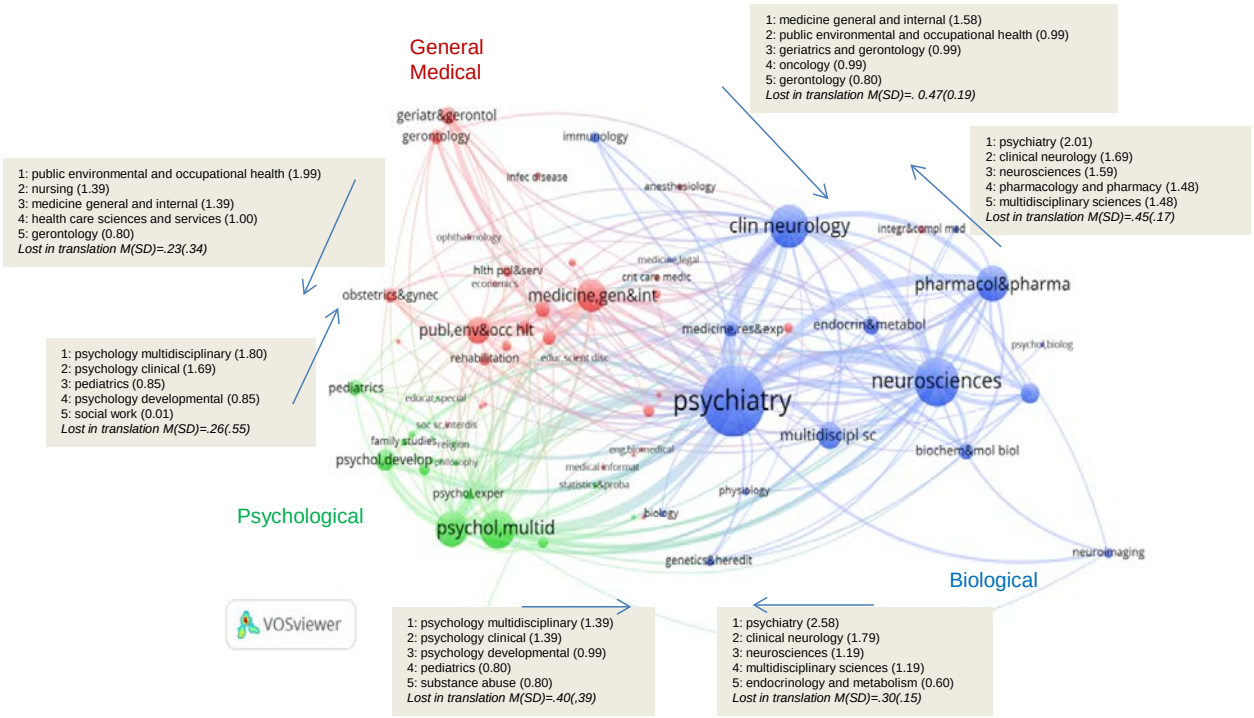


Supplement Figure 6-2. Sorted degree centrality (# of connections) for each discipline in the Google News network.

The moderate number of connections in general, and linear increase across the graph could represent a range of translational potential. The low number of connections associated with disciplines outside psychology/psychiatry (e.g., humanities, philosophy, medical ethics, computer science, evolutionary biology) could represent specific areas in which translational potential could be maximized.



Supplement Figure 6-3: Color version of Figure 2, the discipline-wise network

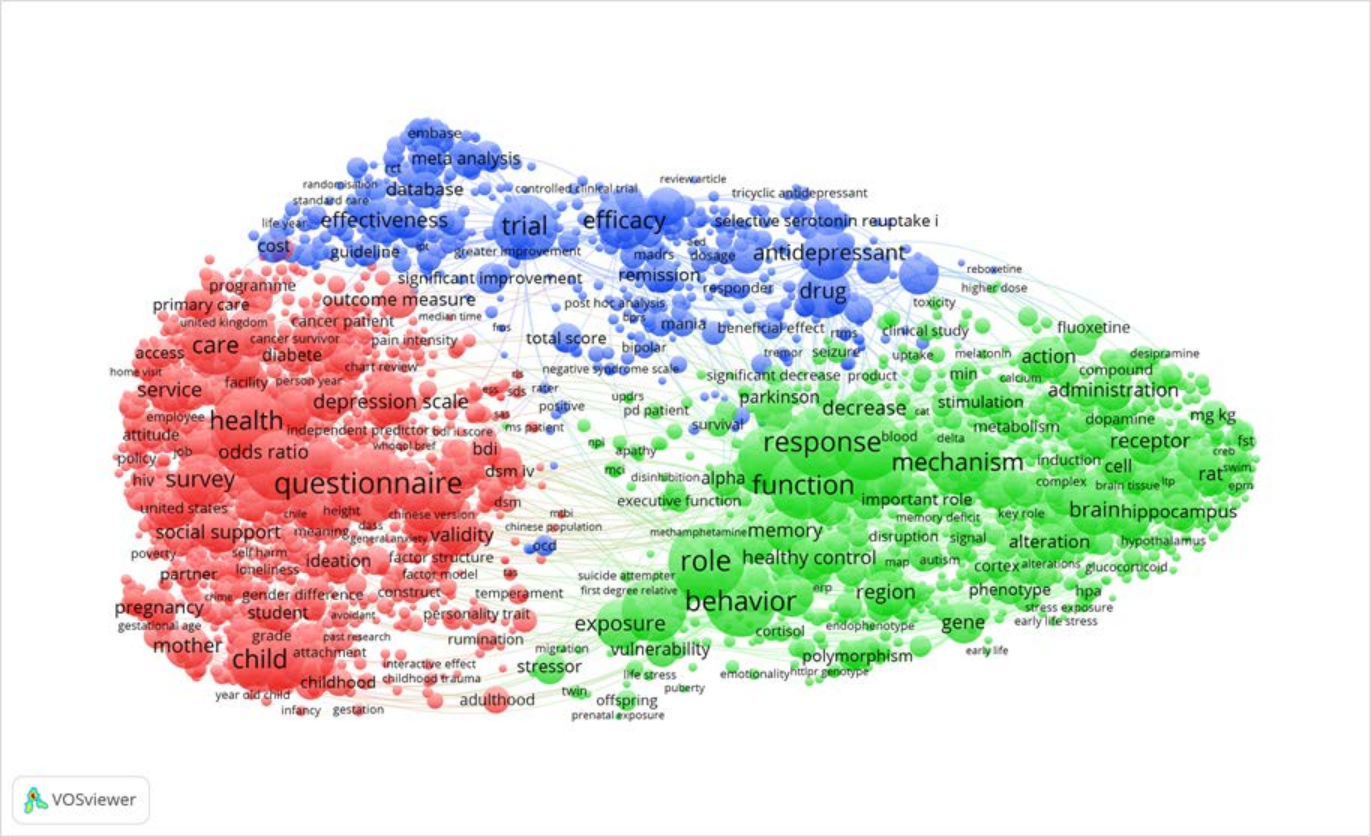


Supplement 7: On the use of binary weights.

A primary goal of this paper was to consider the extent that texts that mention some terms also measure other terms, e.g., whether texts that mention terms in one cluster were likely to mention terms from another cluster. Towards this end we used the fairly common approach of binarizing connection strengths such that connections deemed stronger than some threshold were considered to be co-mentioned (i.e., the connection was deemed "present" and the nodes were said to communicate) and connections deemed weaker than the threshold were considered to not be co-mentioned (the connection was deemed "absent" and the nodes were said to not communicate.) We then used metrics dependent on such binary weights such as centrality and bridging measures to make determinations of communication. This approach has advantages and disadvantages. Disadvantages stem from the fact that continuous variation in weights is ignored. In truth all nodes are connected, so binarized weights is an artificial concept. Our threshold for calling a connection present or absent is fairly arbitrary – lower thresholds would find more nodes to be connected, and thus would suggest higher inter-cluster connectivity. And finally, very strong connections, even if there are few, are not given more weight than somewhat strong connections. As such, even if there is a specific connection which serves as a "superhighway" between clusters, it is not given any more prominence in our characterizations than the other connections that are present or absent. The result is that we could suggest, for example, that clusters are not well connected despite the presence of one or more very strong connections. The advantages of this approach are that it captures our question, which more strongly regards the number of super-threshold connections than whether there are specific strong connections. That is, for questions such as "are texts that mention terms in cluster 1 also likely to mention terms in cluster 2" the binarized connections method answers whether the mention of most terms in cluster 1 are likely to co-occur with mention of terms in cluster 2 without regard for the relative frequency of mentions of the cluster 1 terms. The continuous weights approach would account for there being many publications which mention a single term in cluster 1, that all mentioned a single term in cluster 2; we did not want to emphasize such occurrences. In addition, the use of binarized weights is a notorious solution to the overfitting problem in which specific weights are often too dependent on specific sampling criteria – binary weights are likely more robust such that if we changed sampling criteria to include or not include any given nodes that has very strong weights, our conclusions would likely remain the same. Finally, the use of binary weights allows use of many statistics (e.g., centrality and bridging) that answer our primary questions; these statistics are not developed for continuous weight approaches. While this last argument would not generally be a reason we would use the binary weights approach – we could likely have redeveloped relevant statistics and published the associated methods and validation papers before submitting the current work – the other arguments for preserving binary weights convinced us this detour would likely not be of sufficient effort to warrant pursuing it.

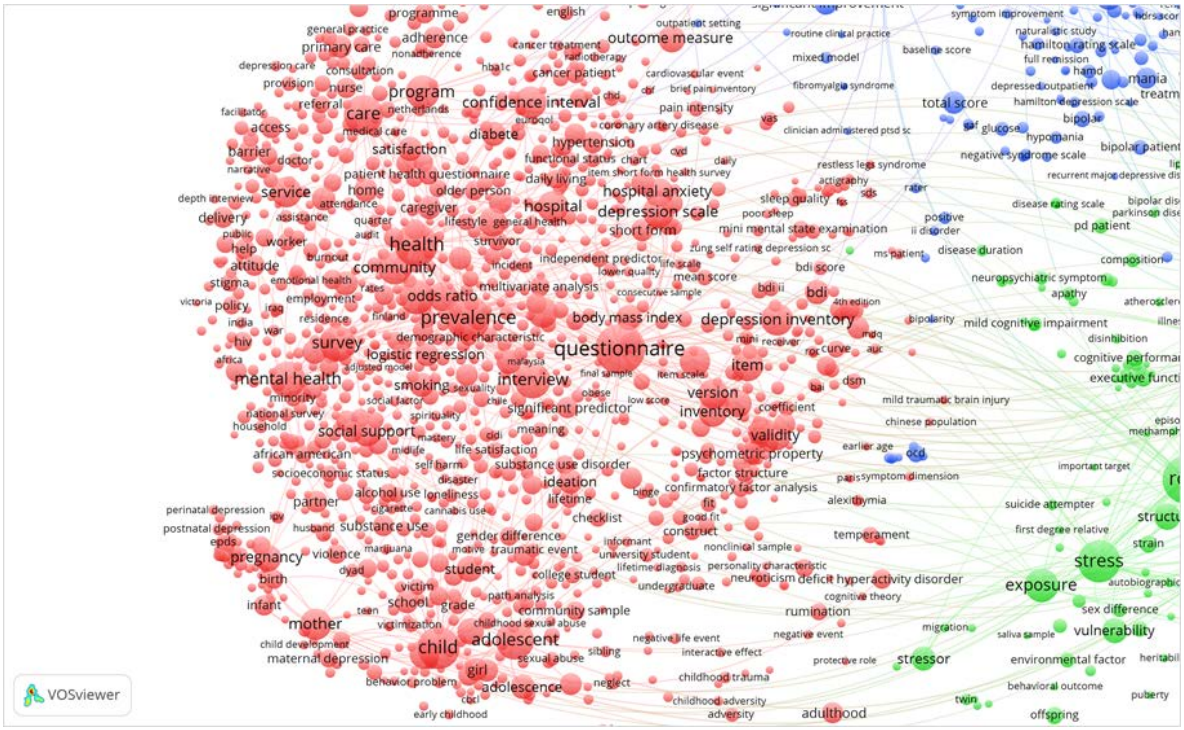
Supplement 8: Noun Phrase Co-occurence Network details and statistics

Supplement Figure 8-1: Color version of Figure 3, the Noun Phrase Co-occurrence Network

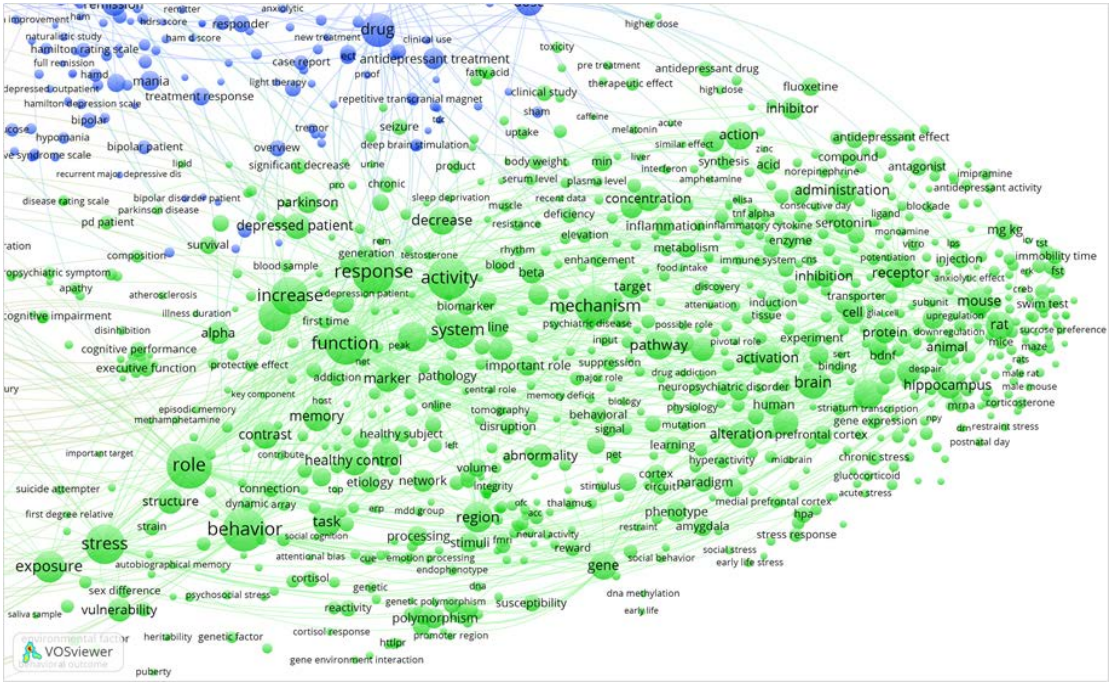


Supplement Figure 8-2: Detail for Noun Phrase Co-occurrence Network

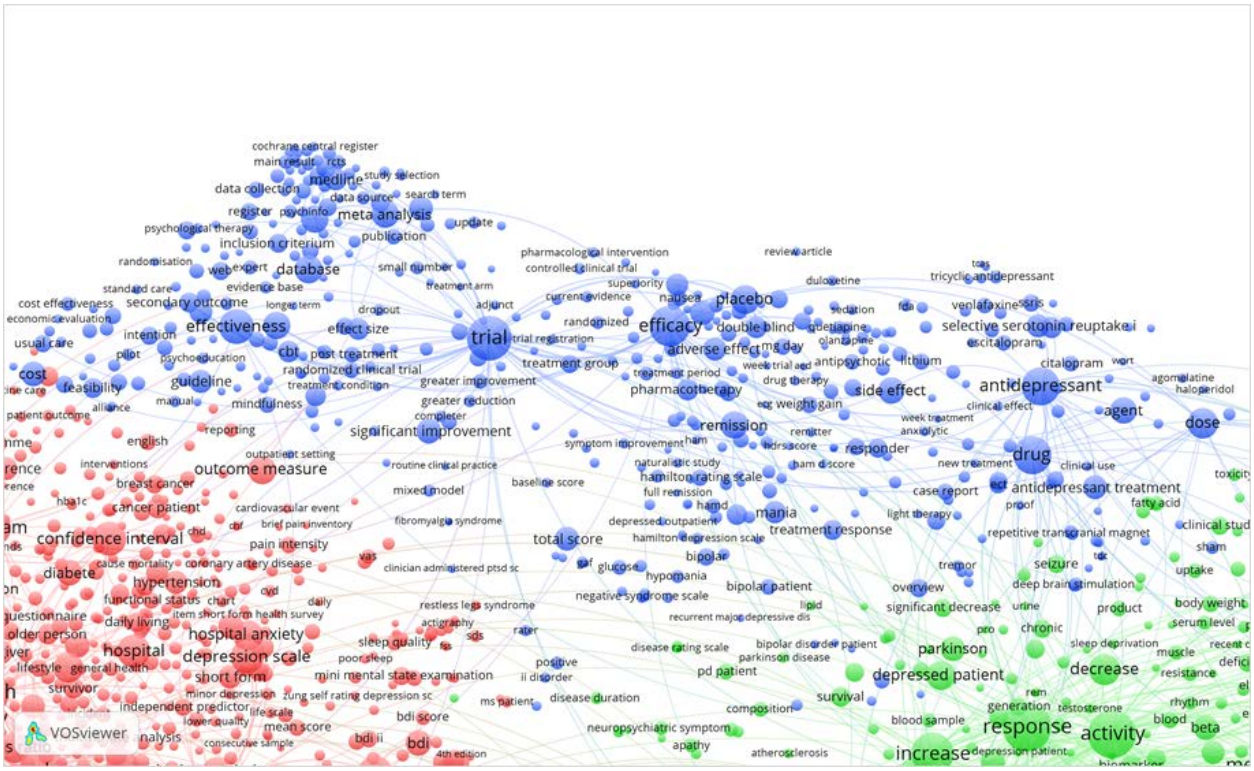
Cluster 1:



Cluster 2:

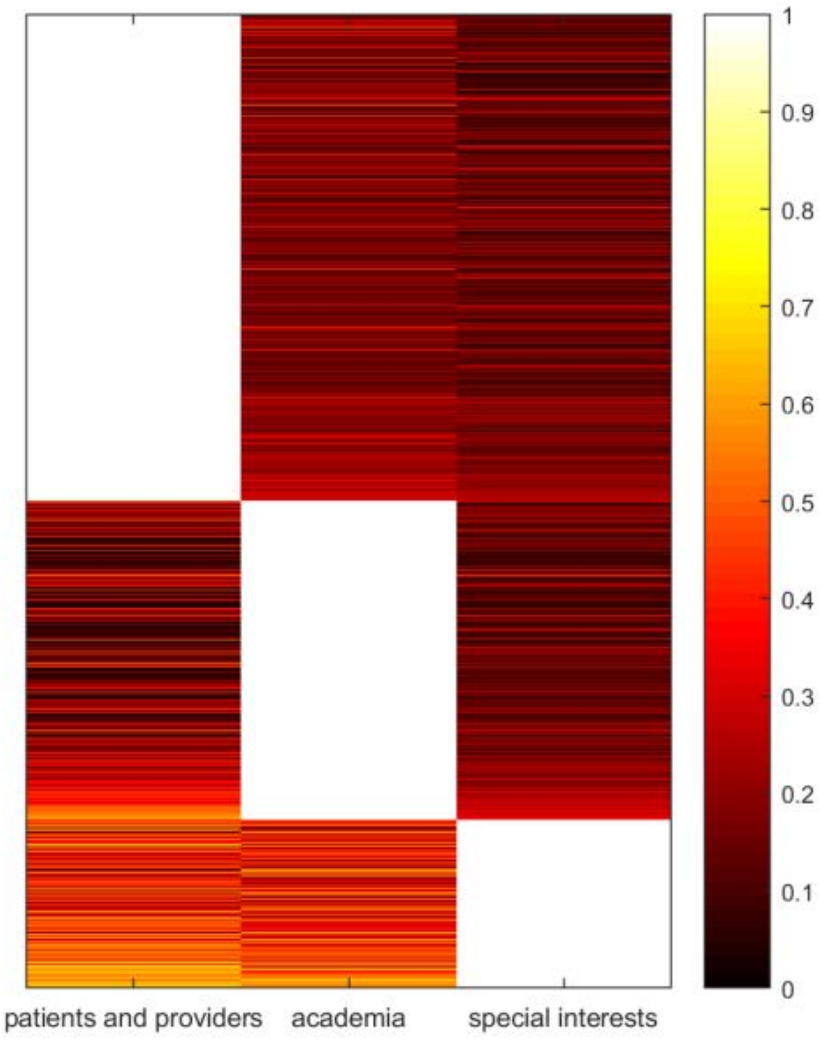


Cluster 3:


cochrane central register
main result
rcts
medline
study selection
data collection
search term
data source
register
psychinfo
meta analysis
update
psychological therapy
publication
inclusion criterium
randomisation
web
expert
database
small number
pharmacological intervention
review article
controlled clinical trial
tcas
tricyclic antidepressant
duloxetine
standard care
evidence base
treatment arm
superiority
cost effectiveness
secondary outcome
longer term
dropout
adjunct
current evidence
nausea
placebo
sedation
fda
venlafaxine
ssris
economic evaluation
effectiveness
effect size
trial
trial registration
randomized
efficacy
double blind
quetiapine
selective serotonin reuptake i
usual care
intention
olanzapine
escitalopram
pilot
psychoeducation
cbt
post treatment
adverse effect
mg day
antipsychotic
lithium
citalopram
wort
cost
randomized clinical trial
treatment group
week trial aed
agomelatine
haloperidol
feasibility
guideline
treatment condition
greater improvement
treatment period
drug therapy
pharmacotherapy
side effect
antidepressant
manual
mindfulness
greater reduction
ecg weight gain
clinical effect
agent
dose
patient outcome
alliance
completer
remitter
week treatment
anxiolytic
remission
reporting
significant improvement
symptom improvement
ham
hdrs score
responder
english
outpatient setting
natural istic study
ham d score
new treatment
drug
clinical use
interventions
routine clinical practice
hamilton rating scale
outcome measure
baseline score
full remission
case report
ect
antidepressant treatment
breast cancer
mixed model
hamd
proof
fatty acid
hba1c
cancer patient
cardiovascular event
depressed outpatient
mania
light therapy
clinical stud
cbf
brief pain inventory
fibromyalgia syndrome
hamilton depression scale
treatment response
repetitive transcranial magnet
confidence interval
pain intensity
total score
bipolar
sham
vas
clinician administered ptsd sc
glucose
tremor
seizure
uptake
diabete
coronary artery disease
hypomania
hypertension
bipolar patient
overview
deep brain stimulation
cvd
negative syndrome scale
functional status
chart
daily
restless legs syndrome
lipid
significant decrease
urine
product
body weight
questionnaire
item short form health survey
recurrent major depressive dis
daily living
actigraphy
pro
chronic
serum level
older person
hospital anxiety
rater
sds
sleep quality
fes
disease rating scale
bipolar disorder patient
parkinson
sleep deprivation
muscle
hospital
depression scale
poor sleep
parkinson disease
positive
decrease
lifestyle
general health
ii disorder
pd patient
depressed patient
resistance
short form
mini mental state examination
survival
rem
survivor
minor depression
zung self rating depression sc
ms patient
disease duration
composition
generation
testosterone
rhythm
independent predictor
bdi score
neuropsychiatric symptom
blood sample
response
activity
blood
beta
mean score
bdi ii
bdi
apathy
atherosclerosis
increase
depression patient
4th edition
VOSviewer

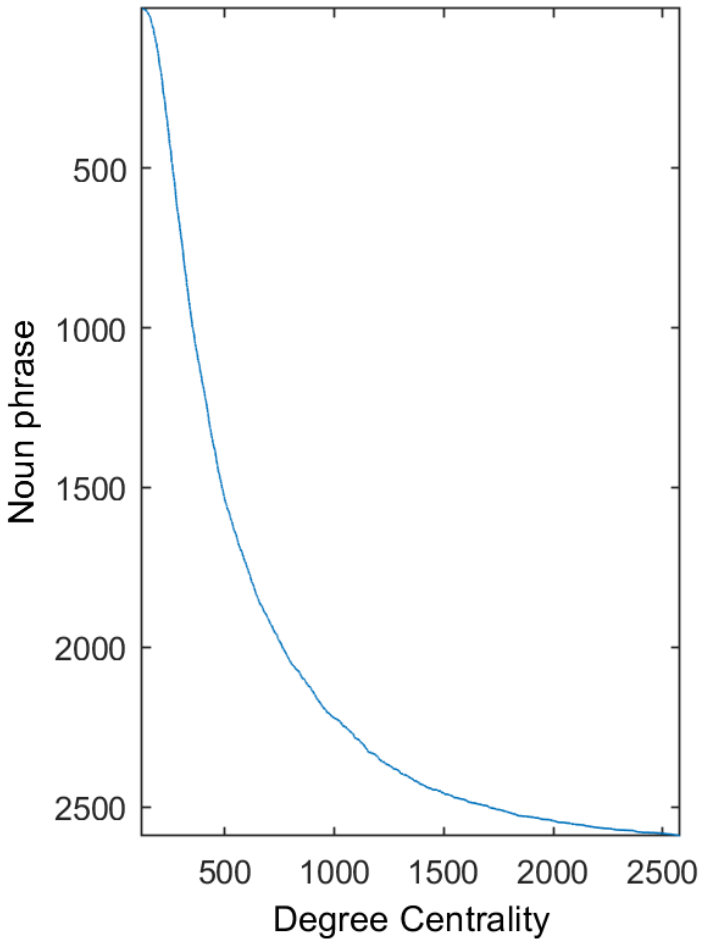
Supplement Figure 8-2: Lost in Translation Indices for each noun phrase from each module to each other module in the network.

Low levels of connectivity for the first two modules (patients and providers and academia; deeper reds) and moderate connectivity for the special interests module (yellows) are shown. Terms are unlabeled so as to keep the figure readable, given the large number of terms.



Supplement Figure 8-3: Sorted degree centrality for each network term

The relatively high number of connections associated with most terms in the network could suggest moderate translational potential if the network were to be capitalized upon in planning communications.



Supplement 9: Network of top 50 associations for “side effect”

Terms are plotted in three modules, with module indicated by the color of the circle next to the item.