nature REVIEWS DRUG DISCOVERY

SUPPLEMENTARY INFORMATION

In format as provided by the authors

Breaking Eroom's Law

Michael S. Ringel, Jack W. Scannell, Mathias Baedeker and Ulrik Schulze

https://doi.org/10.1038/d41573-020-00059-3

Supplementary Box 1 | Data and analysis

Breaking Eroom's Law

The count and value of NMEs relative to R&D spend comes from BCG's New Therapeutic Drug (NTD) Database, which is also the source of BCG's annual publication in *Nat. Rev. Drug. Discov.* showing trends in count and value over time.[1] FDA approvals are from FDA's Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER),[2] peak sales estimates are from EvaluatePharma®,[3] and R&D spend data are from BCG Value Science, inflation-adjusted using the standard global GDP-based inflator from the Economist Intelligence Unit.[4] For additional details on methodology, see Schulze, Baedeker, Chen and Greber.[5]

Eroom's Law is linear on a log scale through 2010, with an average increase approximately 12% per annum, or a halving of productivity approximately every seven years. If this holds true even after 2010, we would see a continuation of the linear development of the number of drugs approved per billion US$ R&D spending when using a logarithmic scale.

To demonstrate that the deviation from regression after 2010 is statistically relevant we assumed the following:

1) Eroom's law is true up to 2010.
2) Eroom's law is not true between 2010 and 2018.

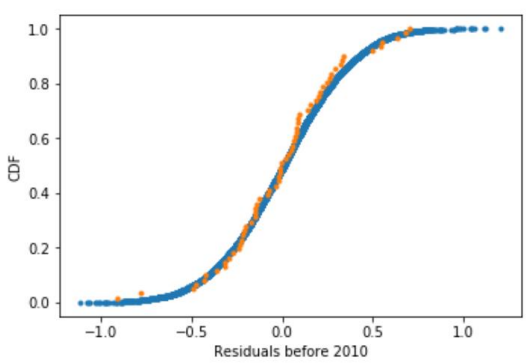
To confirm our assumptions, we used a log-linear regression to describe the number of new molecular entities (NMEs) approved by the US FDA per billion US$ R&D spending from 1950–2010. The *P* value, which describes the chance to reject the hypothesis incorrectly that there is no correlation between the data points, was $< 2 \times 10^{-39}$. The R^2 value, which describes how close the data points are to the assumed trend line, has a high value of over 95% and confirms the first visual impression that the data points follow a log-linear trend (see Figure 1a).

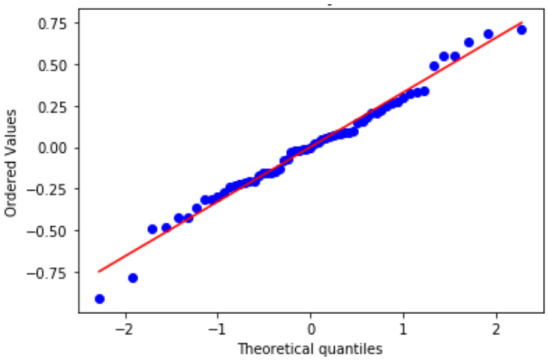
As a second step, we analysed the data after 2010 to see if the trend still holds true. If the trend continues, the distance of the data points from the trend line, the residuals, should follow the same random distribution as from 1950–2010. We used a *t* test to check the hypothesis that the residuals do follow the same trend. A pre-requirement for a *t* test is normally distributed residuals. This was confirmed using a probability plot (see Supplementary Figure 1a).

We find, as shown in Supplementary Figure 1b, that the residuals show a deviation after 2010. The *P* value here, which describes the chance to incorrectly reject this hypothesis, was $< 1.5 \times 10^{-10}$, and hence we conclude that the trend line does not describe the development after 2010.

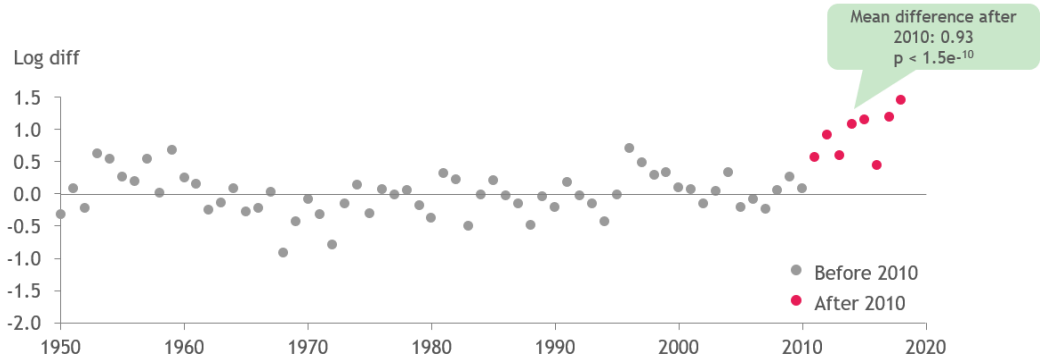
Note that we are unable to apply the *t* test methodology to the analysis of the value of NME approvals shown in Figure 1b, because even before 2010 the errors are not normally distributed (see probability plot in Supplementary Figure 1c). However, there is a mean difference of residuals after 2010 of 0.32 (see Supplementary Figure 1d), and we return to the assessment of this trendline using a bootstrapping method, below.

Supplementary figure 1a, cumulative distribution function and probability plot

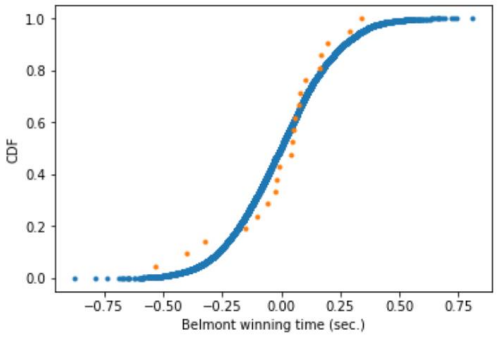


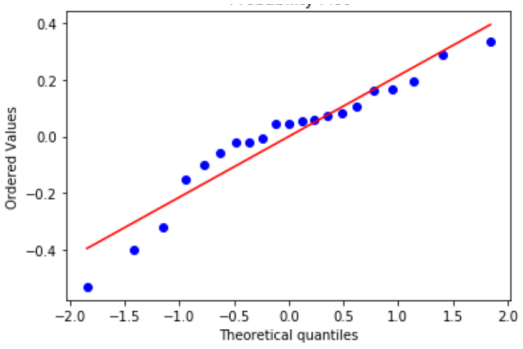


Supplementary figure 1b

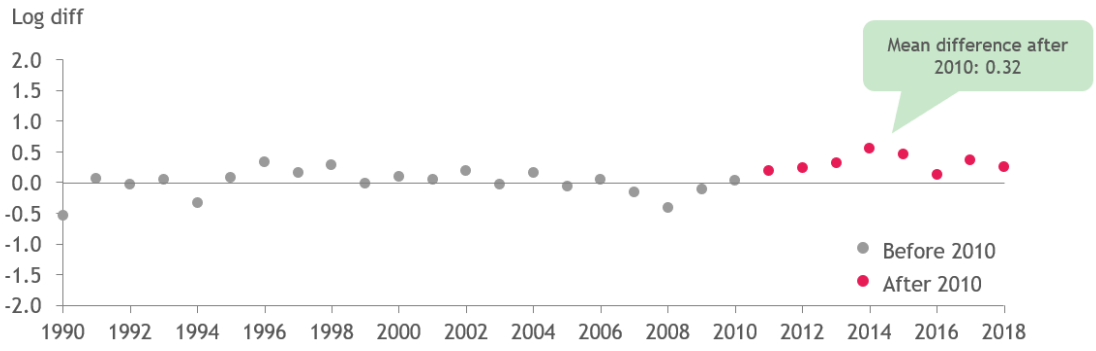


Supplementary figure 1c, cumulative distribution function and probability plot





Supplementary figure 1d

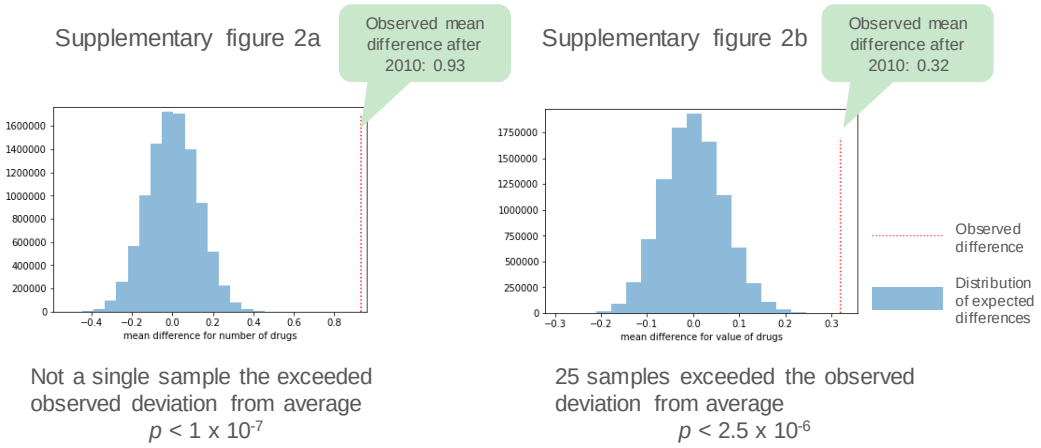


Supplementary Figure 1 | *t* test of deviation from Eroom's Law trend line. a | Cumulative distribution function and probability plot of residuals for count version of Eroom's Law up through 2010 show residuals are normally distributed, hence a *t* test is applicable. **b |** Residuals for count version of Eroom's Law show deviation from prior trend after 2010. **c |** Cumulative distribution function and probability plot of residuals for value version of Eroom's Law up through 2010 show residuals are not normally distributed, hence a *t* test cannot be applied. **d |** Residuals for value version of Eroom's Law, with mean difference after 2010 highlighted, but no *t* test applied.

We employed an additional bootstrapping analysis to both the assessment of the trend line of count of NMEs as well as value of NMEs (indeed, this was the only method employed for the latter analysis). Bootstrapping methods do not require an assumption of a normal distribution of residuals.

First, we determined the mean difference of the residuals from the trend line of the data points after 2010. Second, we compared this with the residuals before 2010. If both datasets belong to the same population, the order of the residuals should be independent. In this case, when the order is changed, the mean difference between the residuals should not change. However, if the two populations are different, when changing the order of data points, we should see a shift in the means of the residuals of the two populations. So the bootstrapping method assesses how likely we would be to see the observed mean difference of the data points after 2010 using only the data points before 2010 in any order. In our analysis, we re-ordered these residual data points 10 million times to get a distribution of the possible mean differences.

As shown in Supplementary Figures 2a and 2b, we find that for both the count version and the value version of Eroom's Law that the observed pattern of increasing residual deviation from the trend line is unlikely to be due to chance alone. For the count version of Eroom's law, the observed mean difference of 0.93 after 2010 was not exceeded even once in 10 million trials, yielding a bootstrapping value $P < 1 \times 10^{-7}$. For the value version of Eroom's law, the observed mean difference of 0.32 after 2010 was exceeded only 25 times in 10 million trials, yielding a value $P < 2.5 \times 10^{-6}$.



Supplementary Figure 2 | Bootstrapping test of deviation from Eroom's Law trend line. a | Bootstrapping test of count version of Eroom's Law shows deviation from trendline after 2010 ($p < 1 \times 10^{-7}$). **b |** Bootstrapping test of value version of Eroom's Law shows deviation from trendline after 2010 ($p < 2.5 \times 10^{-6}$).

Estimate of the share of development costs per NME between successful and failed projects

Figure 1c shows a directional estimate of how the cost of failure rose and then declined as a share of the total cost of development. This estimate is based on the following calculation:

First, we use a point estimate of the costs in each of phases I, II and III from DiMasi, Grabowski & Hansen.[6] This publication provides an estimate based on first-in-human candidates from 1995–2007. We assume the midpoint of this period and offset by half their estimated total phase I–phase III duration to ascribe their cost estimates to a point in time of 2006. To avoid confusion, we restate their phase cost estimates, reported in 2013$, to 2006$, using the US Bureau of Labor Statistics (BLS) CPI-U deflator.

Second, we extend these cost estimates to other years, using the US Bureau of Economic Analysis (BEA) Biomedical Research and Development Price Index (BRDPI), available at https://officeofbudget.od.nih.gov/gbipriceindexes.html. This approach accounts for the underlying change in costs of R&D (which have risen at a faster rate than CPI), but is admittedly a simplification, in that it ignores potential trends in cost by phase and overall due to changes in trial size, complexity, duration, and the exact timing for discounting of cash flows. We have annual data for the BRDPI for 2000 forward, but only 5-year data for 1990 to 1999, so we interpolate the annual change for the missing years.

Third, we estimate probabilities of success by phase for three time points: 1990, 2004 and 2013, using Pammolli, Magazzini and Riccaboni's[7] linear fit estimate for the first two time points (these were the first and last year provided in their analysis), and using Anderson, Wagner and Man[8] for 2013 (the midpoint of their 2011–2015 estimate). In both cases, we add regulatory attrition to phase III.

The results of this analysis are shown in the table below and in Figure 1c.

	US BEA BRPDI,	Cost ($M)			Probability of success			Starts required						
Year	Annual change (%)	P1	P2	P3	P1	P2	P3	P1	P2	P3	Total cost ($M)	Cost of molecule ($M)	Cost of failure ($M)	% in failure
1990	5.4	11.9	27.5	119.7	68%	58%	69%	3.7	2.5	1.4	285	159	126	44%
1991	5.0	12.5	29.0	126.2										
1992	4.6	13.1	30.4	132.5										
1993	4.3	13.7	31.8	138.7										
1994	3.9	14.3	33.2	144.6										
1995	3.5	14.9	34.5	150.2										
1996	3.5	15.4	35.7	155.5										
1997	3.6	16.0	36.9	161.0										
1998	3.6	16.5	38.3	166.7										
1999	3.7	17.1	39.6	172.8										
2000	3.7	17.7	41.1	179.1										
2001	3.3	18.4	42.6	185.7										
2002	3.3	19.0	44.0	191.8										
2003	3.5	19.6	45.5	198.2										
2004	3.7	20.3	47.1	205.1	48%	29%	35%	20.8	9.9	2.9	1474	272	1201	82%
2005	3.9	21.1	48.8	212.7										
2006	4.6	21.9	50.7	221										
2007	3.8	22.9	53.0	231.2										
2008	4.7	23.8	55.0	240.0										
2009	2.9	24.9	57.6	251.2										
2010	3	25.6	59.3	258.5										
2011	2.9	26.4	61.1	266.3										
2012	1.3	27.2	62.9	274.0										
2013	1.9	27.5	63.7	277.6	45%	33%	71%	9.5	4.3	1.4	926	369	557	60%

Scaling share of NMEs with a GWAS at least 5 years prior to approval by number of GWAS publications

We determined the target gene for all NME approvals from 2010 forward using the Target Entrez Gene ID from Citeline's Pharmaprojects® database (Informa, 2019). For entries where a Target Entrez Gene ID is available, we searched the Gene ID in NCBI's Gene database (https://www.ncbi.nlm.nih.gov/gene/) for corresponding NHGRI (National Human Genome Research Institute) GWAS Catalog entries. A threshold of one GWAS catalog entry as a minimum was used as

proxy for the availability of genetic validation. The number of NMEs with genetic validation (based on today's retrospective knowledge) divided by the total number of NMEs has been constant at approximately 50% (data not shown). In order to analyse the percentage of NME approvals that had known human genetic validation at the time of their entry into development, for each NHGRI GWAS Catalog entry (obtained as described above), we obtained the publication date of the corresponding research paper from NCBI PubMed (https://www.ncbi.nlm.nih.gov/pubmed/). Where more than one GWAS catalog entry and/or publication were available for a specific Gene ID, the earliest publication year was used as proxy for the first availability of human genetic validation. The number of NMEs with genetic validation at least 5 years prior to approval divided by the total number of NMEs has risen approximately linearly from 0% in 2011 to 45% in 2018 (data not shown).

For Figure 2a, we show the number of NMEs with genetic validation at least 5 years prior to approval scaled by the number of GWAS publications available, to account for the underlying growth in the number of GWASes available. These data are from the complete GWAS catalog at https://www.ebi.ac.uk/gwas/docs/file-downloads and are as follows:

Year	**GWAS publications available**
2007	81
2008	210
2009	406
2010	684
2011	1,008
2012	1,335
2013	1,675
2014	1,981
2015	2,278
2016	2,586
2017	2,939
2018	3,379

The regression analyses for Figures 2a, 2b, and 2c are performed on logit transformed data, given that the dependent variables are all percentage data.

Bootstrapping test of narrowing indications

We analysed whether there is a trend towards more narrow indications by examining the percentage of approvals at levels below ICD-10 level 3. As shown in Supplementary Figure 3a, ICD-10 codes have up to seven levels of increasingly narrow definitions of disease. The first three levels comprise the disease category, which historically was the main definition of the condition. In the example shown in Supplementary Figure 3a, this is C92, myeloid leukaemia. Levels below this define characteristics such as site, etiology, or other manifestation of the state of the disease. Information regarding the underlying disease biology would often be captured at these levels. In the example shown in Supplementary Figure 3a, this includes C92.1, Chronic myeloid leukaemia, BCR/ABL-positive, C92.6 Acute myeloid leukaemia with 11q23-abnormality, and C92.8, Acute myeloid leukaemia with multilineage dysplasia. The table below provides a comprehensive list of NME approvals in the period studied that have any approval below ICD-10 level 3, based on EvaluatePharma® data[3].

Company (current)	**Product (proprietary name)**	**FDA approval year**	**ICD-10 approval**
GlaxoSmithKline	Cutivate	1990	L23, L24, L20.8, L21 ,L20, L71.0, L30.1, I83.1
Pfizer	Synarel	1990	E30.1
Shire	Ethmozine	1990	I47.2

GlaxoSmithKline	Exosurf	1990	P22.0
Abbott Laboratories	eurodin	1990	F51.0, G47.0
Pfizer	Idamycin	1990	C92.0
Pfizer	Accupril	1991	I50.0
Novartis	Aredia	1991	E83.9
Sanofi	Ceredase	1991	E75.2
Pfizer	Zithromax	1991	A40, G00.1, H65, H66, J13, J20, J32, J40
Abbott Laboratories	Survanta	1991	P22.0
Johnson & Johnson	Supprelin	1991	E30.1
Sanofi	Fludara	1991	C91.1
AstraZeneca	Foscavir	1991	B20.2
Genta	Ganite	1991	E83.9
Hospira	Nipent	1991	C91.4
UCB	Actinex	1992	L57.0
Sanofi	Ambien	1992	F51.0, G47.0
Bayer	Betapace	1992	I47.2
Bristol-Myers Squibb	Vumon	1992	C91.0
Johnson & Johnson	Sporanox	1992	B35.1
Merck & Co	Propecia	1992	L63, L63.0, L63.1, L63.2, L63.8, L63.9, L64, L64.0, L64.8, L64.9
Novartis	Lamisil	1992	B35.1
Merck & Co	Claritin	1993	J30.1, J30.2
Pfizer	Tazocin	1993	V13.02
Boehringer Ingelheim	Orlaam	1993	F11.2
AstraZeneca	NeuTrexin	1993	B20.6
Sanofi	Lovenox	1993	I80.2
Johnson & Johnson	Leustatin	1993	C91.4
Sanofi	Cerezyme	1994	E75.2
UCB	Semprex-D	1994	J30.1, J30.2
AstraZeneca	Rhinocort	1994	J30.1, J30.2
Astellas Pharma	Prograf	1994	Z94.4
Shire	Oncaspar	1994	C91.0
Pfizer	Fragmin	1994	I80.2
Roche	CellCept	1995	Z94.0
Pfizer	Zyrtec	1995	J30.1, J30.2
GlaxoSmithKline	Valtrex	1995	A60.0
Sanofi	Rilutek	1995	G12.2
Sanofi	Allegra	1996	J30.1, J30.2
Meda	Astelin	1996	J30.1, J30.2
Acorda Therapeutics	Zanaflex	1996	R25.2
Jazz Pharmaceuticals	Cystadane	1996	E72.0
Gilead Sciences	Vistide	1996	B20.2
Medimmune	RespiGam	1996	J12.1
Merck & Co	Orgaran	1996	I80.2
Zambon	Monurol	1996	V13.02
Abbott Laboratories	Mavik	1996	I50.0
Shire	Agrylin	1997	D47.3, D75.2
3M	Aldara	1997	A63.0
Roche	Zenapax	1997	Z94.0
Novartis	Regranex	1997	E10.5
Wyeth	Normiflo	1997	I80.2
Amgen	Infergen	1997	B17.1
Merck & Co	Aggrastat	1998	I20.0
Abbott Laboratories	Zemplar	1998	E21.0, E21.1, E21.2, E21.3
Novartis	Vitravene	1998	B20.2

Pfizer	Viagra	1998	F52.2, F52.4
Pfizer	Detrol	1998	R32, N39.3, N39.4
Sanofi	Thymoglobulin	1998	Z94.0
AstraZeneca	Synagis	1998	J12.1
QOL Medical	Sucraid	1998	E74.3
Novartis	Simulect	1998	Z94.0
Sanofi	Renagel	1998	E83.3
bayer	Refludan	1998	D69.6
Teva Pharmaceutical Industries	Provigil	1998	G47.4
Forest Laboratories	Infasurf	1998	P22.0
GlaxoSmithKline	Wellferon	1999	C91.4
Chiesi	Curosurf Aerosol	1999	P22.0
Merck & Co	Temodar	1999	C71.9
Pfizer	Sonata	1999	F51.0, G47.0
Pfizer	Rapamune	1999	Z94.0
Otsuka Holdings	Pletal	1999	I73.9
Allergan	Ferrlecit	1999	D63.8
Sanofi	Hectorol	1999	E21.0, E21.1, E21.2, E21.3
Ikaria	INOmax	1999	P22.0
GlaxoSmithKline	Argatroban	2000	D69.6
Novartis	Visudyne	2000	H35.3
Teva Pharmaceutical Industries	Trisenox	2000	C92.0
Daiichi Sankyo	Evoxac	2000	M35.0
Pfizer	Prevnar	2000	A40, G00.1, H65, H66, J13, J20, J32, J40
Elan	Myobloc	2000	G24.3
Pfizer	Mylotarg	2000	C92.0
Celgene	Innohep	2000	I80.2
Amgen	Aranesp	2001	D63.8
GlaxoSmithKline	Arixtra	2001	I80.2
Bayer	MabCampath	2001	C91.1
Merck & Co	Clarinex	2001	J30.1, J30.2
Novartis	Zometa	2001	E83.9
Roche	Valcyte	2001	B20.2
GlaxoSmithKline	Twinrix	2001	B15.9
Johnson & Johnson	Tracleer	2001	I27.0, I27.2
Novartis	Elidel	2001	L23, L24, L20.8, L21, L20, L71.0, L30.1, I83.1
Merck & Co	PEGIntron	2001	B17.1
Johnson & Johnson	Natrecor	2001	I50.0
Novartis	Gleevec	2001	C92.1
Baxter International	Aralast	2002	E88.0
Jazz Pharmaceuticals	Xyrem	2002	G47.4
Eli Lilly	Strattera	2002	F90.0
United Therapeutics	Remodulin	2002	I27.0, I27.2
Roche	Pegasys	2002	B17.1
Swedish Orphan Biovitrum	Orfadin	2002	E70.2
Sanofi	Aldurazyme	2003	E76.0
Eli Lilly	Cialis	2003	F52.2, F52.4
CSL	Zemaira	2003	E88.0
Johnson & Johnson	Zavesca	2003	E75.2
Johnson & Johnson	Velcade	2003	C90.0
Pfizer	Somavert	2003	E22.0
Sanofi	Fabrazyme	2003	E75.2

Bayer	Levitra	2003	F52.2, F52.4
Allergan	Campral	2004	F10.2
Sanofi	Clolar	2004	C91.0
Johnson & Johnson	Ventavis	2004	I27.0, I27.2
Amgen	Sensipar	2004	E21.0, E21.1, E21.2, E21.3
Shire	Fosrenol	2004	E83.3
Emmaus Life Sciences	NutreStore	2004	K91.2
Pfizer	Macugen	2004	H35.3
Pfizer	Lyrica	2004	M79.2
Sumitomo Dainippon Pharma	Lunesta	2004	F51.0, G47.0
Swedish Orphan Biovitrum	Kepivance	2004	K12.3
GlaxoSmithKline	Arranon	2005	C91.0
Astellas Pharma	Vaprisol	2005	E87.1
Takeda	Rozerem	2005	F51.0, G47.0
Novartis	Exjade	2005	T45.4
BioMarin Pharmaceutical	Naglazyme	2005	E76.2
Ipsen	Increlex	2005	E34.3
Takeda	Amitiza	2006	K59.0
Pfizer	Chantix	2006	F17.2
Novartis	Veregen	2006	A63.0
Pfizer	Sutent	2006	D37.9
Shire	Elaprase	2006	E76.1
Bristol-Myers Squibb	Sprycel	2006	C92.1
Merck & Co	RotaTeq	2006	A08.0
Sumitomo Dainippon Pharma	Omnaris AQ Nasal Spray	2006	J30.1, J30.2
Sanofi	Myozyme	2006	E74.0
Roche	Lucentis	2006	H35.3
GlaxoSmithKline	Veramyst	2007	J30.1, J30.2
UCB	Xyzal	2007	J30.1, J30.2
Shire	Vyvanse	2007	F90.0
Novartis	Tasigna	2007	C92.1
Ipsen	Somatuline	2007	E22.0
Alexion Pharmaceuticals	Soliris	2007	D59.5
Teva Pharmaceutical Industries	Nuvigil	2007	G47.3
Roche	Mircera	2007	D63.8
Gilead Sciences	Letairis	2007	I27.0, I27.2
BioMarin Pharmaceutical	Kuvan	2007	E70.0
Shire	Cinryze	2008	D84.1
Teva Pharmaceutical Industries	Treanda	2008	C91.1
GlaxoSmithKline	Rotarix	2008	A08.0
Salix Pharmaceuticals	Relistor	2008	K59.0
Johnson & Johnson	Nucynta	2008	R52.0
Eisai	Banzel	2008	G40.4
GlaxoSmithKline	Arzerra	2009	C91.1
CSL	Berinert P	2009	D84.1
Ipsen	Dysport	2009	G24.3
Allergan	Savella	2009	M79.7
Otsuka Holdings	Samsca	2009	E87.1
Shire	Kalbitor	2009	D84.1
Novartis	Ixiaro	2009	A83.0
Roche	Actemra	2010	D36.0
Recordati	Carbaglu	2010	E72.2

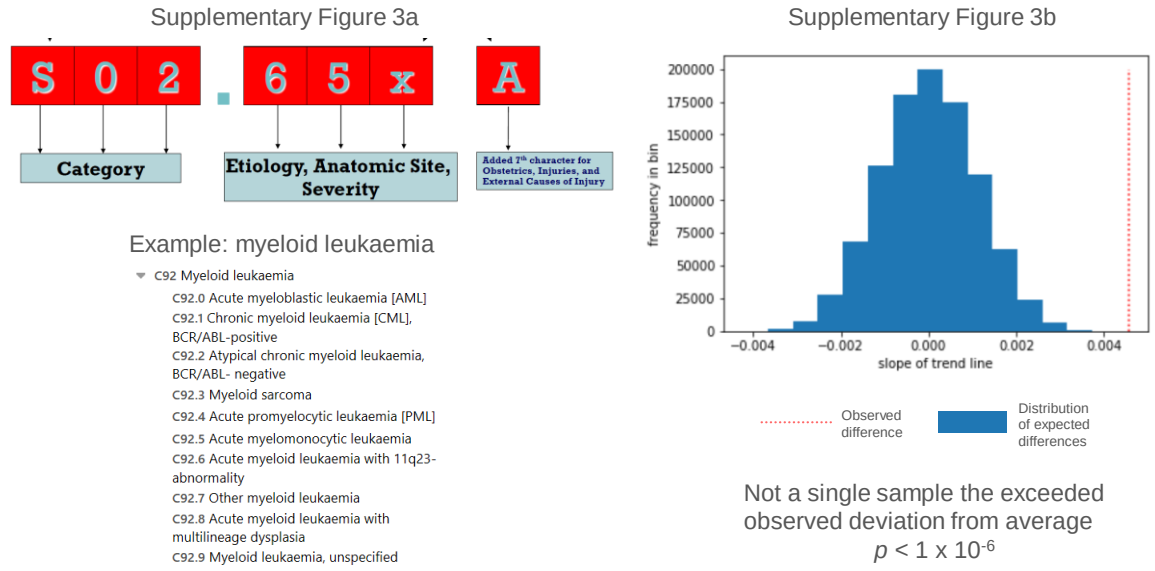
Auxilium Pharmaceuticals	Xiaflex	2010	M20.0
Merz	Xeomin	2010	G24.3
Shire	Vpriv	2010	E75.2
Merck KGaA	Egrifta	2010	E88.1
Pfizer	Prevnar 13	2010	A40, G00.1, H65, H66, J13, J20, J32, J40
Baxter International	Glassia	2010	E88.0
AstraZeneca	Brilinta	2011	I20.0
Bayer	Xarelto	2011	I80.2
Merck & Co	Victrelis	2011	B17.1
Merck & Co	Dificid	2011	A04.7
Jazz Pharmaceuticals	Erwinaze	2011	C91.0
Bayer	Eylea	2011	H35.3
Apotex	Ferriprox	2011	T45.4
Shire	Firazyr	2011	D84.1
Lundbeck	Onfi	2011	G40.4
Bristol-Myers Squibb	Nulojix	2011	Z94.0
ParaPRO	Natroba	2011	B85.0
GlaxoSmithKline	Horizant	2011	G25.8
Johnson & Johnson	Incivek	2011	B17.1
Novartis	Jakafi	2011	C94.4, D47.1, D47.4
LEO Pharma	Picato	2012	L57.0
Teva Pharmaceutical Industries	Synribo	2012	C92.1
Windtree Therapeutics	Surfaxin	2012	P22.0
Mitsubishi Tanabe Pharma	Stendra	2012	F52.2, F52.4
Pfizer	Bosulif	2012	C92.1
Shire	Gattex	2012	K91.2
Pfizer	Elelyso	2012	E75.2
Takeda	Omontys	2012	D63.8
Roche	Erivedge	2012	C44.9
Takeda	Iclusig	2012	C92.1
Amgen	Kyprolis	2012	C90.0
Bristol-Myers Squibb	Eliquis	2012	I80.2
Novo Nordisk	Tretten	2013	D68.2
Gilead Sciences	Sovaldi	2013	B17.1
Johnson & Johnson	Opsumit	2013	I27.0, I27.2
Johnson & Johnson	Olysio	2013	B17.1
Roche	Gazyva	2013	C91.1
Bayer	Adempas	2013	I27.0, I27.2
Emergent BioSolutions	BAT	2013	A05.1
Celgene	Pomalyst	2013	C90.0
Vanda Pharmaceuticals	Hetlioz	2014	G47.2
BioMarin Pharmaceutical	Vimizim	2014	E76.2
Novelion Therapeutics	Myalept	2014	E88.1
Celgene	Otezla	2014	M07.0, M07.1, M07.2, M07.3, L40.5
Stallergenes	Oralair	2014	J30.1, J30.2
Merck & Co	Grastek	2014	J30.1, J30.2
Merck & Co	Ragwitek	2014	J30.1, J30.2
Johnson & Johnson	Sylvant	2014	D36.0
Bausch Health Companies	Jublia	2014	B35.1
Pfizer	Kerydin	2014	B35.1
Salix Pharmaceuticals	Ruconest	2014	D84.1
Merck & Co	Belsomra	2014	F51.0, G47.0

Sanofi	Cerdelga	2014	E75.2
AstraZeneca	Movantik	2014	K59.0
Gilead Sciences	Harvoni	2014	B17.1
Boehringer Ingelheim	Ofev	2014	J84.1
Roche	Esbriet	2014	J84.1
Amgen	Blincyto	2014	C91.0
Abbvie	Viekira Pak	2014	B17.1
Merck & Co	Zerbaxa	2014	V13.02
Daiichi Sankyo	Lixiana	2015	I80.2
Novartis	Farydak	2015	C90.0
Allergan	Avycaz	2015	V13.02
United Therapeutics	Unituxin	2015	M95.00, M95.03
Amgen	Corlanor	2015	I50.0
Novartis	Entresto	2015	I50.0
Novartis	Odomzo	2015	C44.9
Bristol-Myers Squibb	Daklinza	2015	B17.1
Allergan	Vraylar	2015	F84.0, F84.1, F84.2, F84.3, F84.5, F84.8, F84.9
Relypsa	Veltassa	2015	E87.5
Alexion Pharmaceuticals	Strensiq	2015	E83.39
Johnson & Johnson	Darzalex	2015	C90.0
Takeda	Ninlaro	2015	C90.0
Bristol-Myers Squibb	Empliciti	2015	C90.0
Alexion Pharmaceuticals	Kanuma	2015	E75.5
Johnson & Johnson	Uptravi	2015	I27.0, I27.2
Pfizer	Eucrisa	2016	L23, L24, L20.8, L21, L20, L71.0, L30.1, I83.1
Merck & Co	Zinplava	2016	A04.7
Merck & Co	Zepatier	2016	B17.1
Jazz Pharmaceuticals	Defitelio	2016	K76.5
AbbVie	Venclexta	2016	C91.1
Intercept Pharmaceuticals	Ocaliva	2016	K74.3
Gilead Sciences	Epclusa	2016	B17.1
Shire	Xiidra	2016	H04.1, H19.3
Sarepta Therapeutics	Exondys 51	2016	G71.0
Spark Therapeutics	Luxturna	2017	H35.5
Novartis	Kymriah	2017	C91.0
Ultragenyx Pharmaceutical	Mepsevii	2017	E76.2
Melinta Therapeutics	Vabomere	2017	V13.02
Pfizer	Besponsa	2017	C91.0
AbbVie	Mavyret	2017	B17.1
Celgene	IDHIFA	2017	C92.0
Gilead Sciences	Vosevi	2017	B17.1
Portola Pharmaceuticals	Bevyxxa	2017	I80.2
Synergy Pharmaceuticals	Trulance	2017	K59.0
PTC Therapeutics	Emflaza	2017	G71.0
Shionogi	Symproic	2017	K59.0
Novartis	Rydapt	2017	C92.0
Amgen	Parsabiv	2017	E21.0, E21.1, E21.2, E21.3
Sanofi	Dupixent	2017	L23, L24, L20.8, L21, L20, L71.0, L30.1, I83.1
Lexicon Pharmaceuticals	Xermelo	2017	E34.0
Alexion Pharmaceuticals	Ultomiris	2018	D59.5
Stemline Therapeutics	Elzonris	2018	C92.0
Les Laboratoires Servier	Asparlas	2018	C91.0

Shire	Motegrity	2018	K59.0
Astellas Pharma	Xospata	2018	C92.0
Pfizer	Daurismo	2018	C92.0
Verastem	Copiktra	2018	C91.1
AstraZeneca	Lumoxiti	2018	C91.4
Shire	Takhzyro	2018	D84.1
Dompé	Oxervate	2018	H16.0
Amicus Therapeutics	Galafold	2018	E75.2
Shionogi	Mulpleta	2018	D69.6
Agios Pharmaceuticals	Tibsovo	2018	C92.0
Achaogen	Zemdri	2018	V13.02
BioMarin Pharmaceutical	Palynziq	2018	E70.0
Dova Pharmaceuticals	Doptelet	2018	D69.6
AstraZeneca	Lokelma	2018	E87.5
US WorldMeds	Lucemyra	2018	F11.2

Our analysis focused on whether there is a trend towards increasing usage of disease definitions at these lower levels in the approval of NMEs. We constructed a trend line of the percentage of NME approvals granted with an indication definition below ICD-10 level 3, using seven-year moving averages to smooth out the high volatility of residuals. We again found that the residuals are not normally distributed around the trend line (data not shown), and so again found that a bootstrapping method would be required rather than a *t* test.

The null hypothesis is that there is no slope to the trend line; that is, there is no change in the share of approvals occurring at lower than ICD-10 level 3. By permuting the observed data points 1 million times, we asked how frequently the observed upward trend would occur assuming no actual underlying slope. The observed mean difference was not exceeded even once in 1 million trials, and hence we found that the null hypothesis could be rejected at the $p < 1 \times 10^{-6}$ level, as shown in Supplementary Figure 3b.



Supplementary Figure 3 | **Test of narrowing indications**. **a** | The ICD-10 hierarchy has a maximum of seven levels, with the first three comprising the disease category, and subsequent levels providing additional detail. An example is provided for myeloid leukaemia. **b** | A bootstrapping test indicates that the trend to smaller indications shown in Figure 2b is unlikely to be due to chance.

Percentage of NMEs designated as treatments for rare diseases

Orphan drug designations were taken from FDA Center for Drug Evaluation and Research (CDER)[10] as well as Seoane-Vazquez, Rodriguez-Monguio, Szeinbach, Visaria.[11]

References

1. Baedeker M, Ringel M & Schulze U. Value of 2019 FDA approvals: back to the recent average. *Nat. Rev. Drug. Discov.* **19**, 85 (2020).
2. U.S. Food and Drug Administration. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER)
3. EvaluatePharma® January 2020, Evaluate Ltd.
4. The Economist Intelligence Unit (2019), EIU Country data.
5. Schulze U, Baedeker M, Chen YT & Greber D. R&D productivity: On the comeback trail. *Nat. Rev. Drug. Discov.* **13**, 331-332 (2014).
6. DiMasi JA, Grabowski HG & Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **47**, 20-33 (2016).
7. Pammolli F, Magazzini L & Riccaboni M. The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug. Discov.* **10**, 428-38 (2011).
8. Sanford C. Bernstein research. Anderson, Wagner, Man. Global pharmaceuticals: R&D productivity continues to see upswing—good news for drug investors. (17 October 2016).
9. Lanthier M, Miller KL, Nardinelli C. & Woodcock J. An improving approach to measuring drug innovation finds steady rates of first-in-class pharmaceuticals, 1987-2011. *Health Aff.* **32**, 1433-1439 (2013).
10. FDA research. Woodcock, J, CDER: Innovation in New Drug Approvals of 2019 Advances Patient Care Across a Broad Range of Diseases (2020).
11. Seoane-Vazquez E, Rodriguez-Monguio R, Szeinbach SL &Visaria J. Incentives for orphan drug research and development in the United States. *Orphanet. J. Rare Dis.* **3**, 33 (2008).