

# Hodnotenie kvality vedeckého výskumu

*Martin Plesch, Vladimír Bužek*

# Prečo robiť výskum?

- Hľadanie pravdy
- Úspech a sláva
- Krása a dokonalosť
- Intelektuálne dobrodružstvo a zvedavosť
- Učiť sa
- Učiť
- ...

# Ako zistiť, či to robíme dobre?

- Diskusie s kolegami
- Prednášky a postery na konferenciách
- Publikácie/citácie
- Patenty
- Aplikácie
- Uznanie komunitou
- Uznanie verejnosťou
- Money, money, money...

# Prečo hodnotiť kvalitu

- Súťaž o zdroje – nedá sa financovať každá vedecká aktivita
- Rozhodovanie vyžaduje kritériá a/alebo priority
- Označiť niekoho za lepšieho nevyhnutne vedie k označeniu iného za horšieho (vedec, skupina, oblasť, priorita)

# Veda, výskum a vývoj

- Čím bližšie sme k aplikácii, tým ľahšie sa hodnotí výstup
- Čím je práca abstraktnejšia, tým je kvalita ťažšie hodnotiteľná a vyžaduje dlhšie obdobie pre porovnávanie
- Základný výskum predstavuje high risk – high gain stratégiu na dlhé časové obdobie

# Publikácie

- Reflektuje priamu produktivitu vedca
- Recenzované a nerecenzované
- CC a non-CC
- K téme a mimo témy
  
- Nehovorí o kvalite
- Nehovorí o obsahu
- Kumuluje sa s vekom
  
- Úpravy: publikácie za posledných 5 rokov, publikácie k téme, publikácie na autora atď.

# Citácie

- Reflektujú ohlas výsledkov v komunite
- CC a non-CC
- Samocitácie
- Viacnásobné citácie
- Nestabilný systém pre články s veľkým počtom citácií
- Riziko dohody v úzkej komunite

# Impact factor časopisu

- Priemerný počet citácií na článok
- Hovorí o kvalite časopisu, v ktorom bol publikovaný článok (?)
- Umožňuje rýchlejšie posúdiť vedecký výstup (vyjadruje očakávaný počet citácií na publikáciu)
- Dá sa mierne manipulovať zo strany editorov
- Neúmerne zvýhodňuje časopisy s prehľadovými článkami



# Konferencie

- Zmysel má len sledovanie pozvaných prednášok
- Vypovedá o postavení v komunite
- Pozvanie sa dá (bez problémov) kúpiť
- Neexistuje rating konferencií
- Závisí od schopností vedca dobre prednášať (súčasť vedeckých schopností?)

# Rôzne scientometrické kritéria I

- V princípe sa dá analyzovať kompletný scientometrický profil vedca
  - + *všetky publikácie (prezentácie, etc.)*
  - + *všetky citácie na tieto publikácie*
  - + *impakt faktory časopisov s týmito publikáciami*
  - *takýto popis je ťažkopádny a neprehľadný*
- Celkový počet publikácií daného vedca
  - + *vystihuje produktivitu*
  - *nereflektuje impakt výsledkov*
- Celkový počet citácií
  - + *kumulatívne reflektuje impakt*
  - *nezohľadňuje produktivitu vedca (skreslenie reviews)*

# Rôzne scientometrické kritéria II

- Relatívny počet citácií
  - + *umožňuje porovnať autorov rôzneho veku*
  - *preferuje malú produktivitu*
- Počet významných článkov – publikácie citované viac ako  $y$  krát
  - + *dobrá prehľad impaktu výsledkov*
  - *$y$  je náhodne určené číslo*
  - *starší autori sú preferovaní (publikácie sú dlho na trhu)*
- Počet citácií  $q$  najvýznamnejších prác
  - + *vystihuje impakt*
  - *$q$  je náhodne určené číslo*

# Relevantné parametre

$N_p$  Celkový počet publikácií

$N_c^j$  Počet citácií na  $j$ -tu publikáciu

$I$  Impakt faktor časopisu

Počet spoluautorov ...

Hľadáme jedno číslo

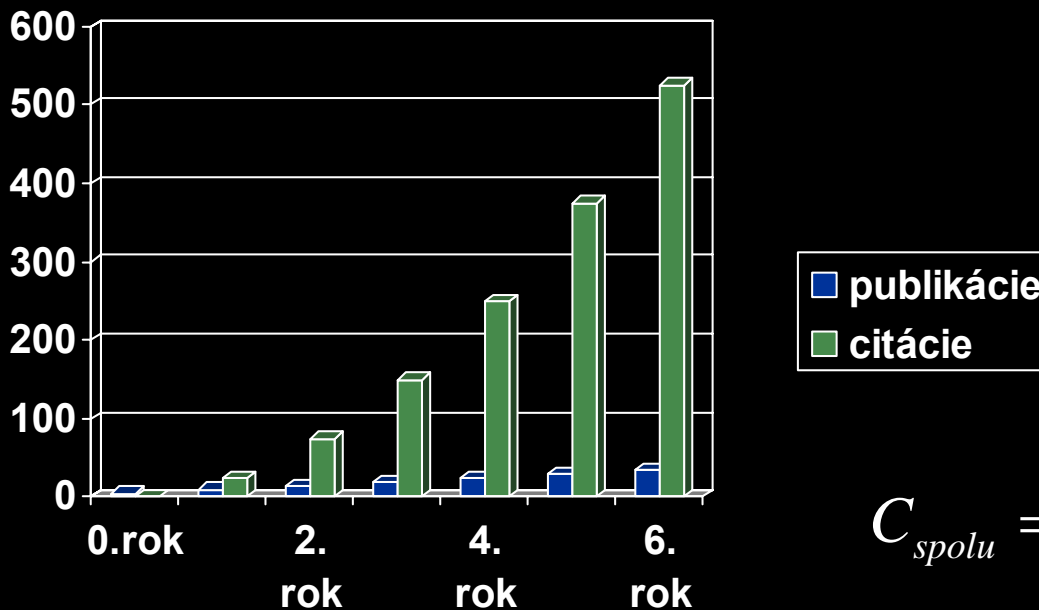
# Hirschov index

***Hirschov index*** je definovaný ako počet  $h$  publikácií daného autora, ktoré sú citované prinajmenšom  $h$  krát, pričom všetky ostatné publikácie autora sú citované menej ako  $h$  krát.

**Autori s rovnakým *h-indexom*** majú rovnakú-porovnateľnú vedeckú výkonnosť hoci počet ich publikácií a citácií sa môže výrazne odlišovať.

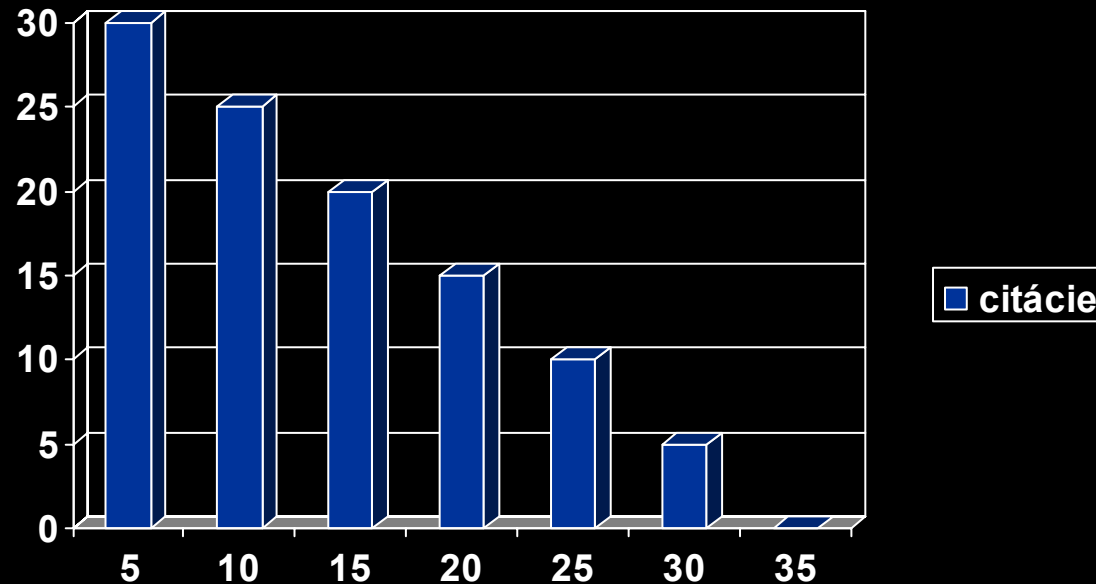
# Jednoduchý (lineárny) model I.

- $p$  publikácií každý rok ( $p=5$ )
- Každá publikácia má počínajúc prvým rokom po jej publikovaní  $c$  citácií ročne ( $c=5$ )
- Otázka: Aký je celkový počet citácií po  $(n+1)$  rokoch? ( $n=6$ )



$$C_{spolu} = \sum_{j=1}^n p \cdot c \cdot j = \frac{n(n+1)}{2} p \cdot c$$

# Jednoduchý lineárny model II.



**h=15 (po 7 rokoch)**

# Vlastnosti *h-indexu*

- *Hirschov index* je lineárnou funkciou počtu rokov  $n$ ; parameter výkonnosti  $m$

$$h = m \cdot n$$

- Celkový počet citácií je kvadratickou funkciou *h-indexu* fenomenologická konštanta  $a \sim 4$

$$N_{c-total} = a h^2$$



# H-index: 20 rokoch vo výskume

m~1 úspešný vedec

m~2 dá sa nájsť na špičkových univerzitách

m~3 výnimočná vedecká osobnosť

h~10-12      associate professor;      N~ 500

h~18      full professor;      N~1300

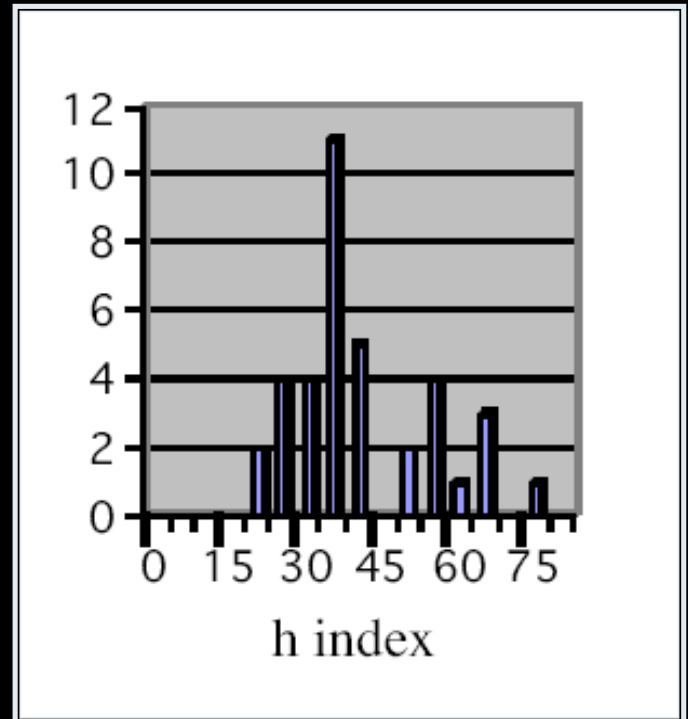
h~15-20      fellow of the APS;      N~1600

h~45      fellow of the NAS;      N~8100

Tieto hodnoty korešpondujú m~1

# H-index a Nobelove ceny

- Fyzici za ostatných 20 rokov
- Publikácie od roku 1955
- $22 < h < 79$
- $\langle h \rangle = 35$
- $0.47 < m < 2.19$
- $\langle m \rangle = 0.47$



Nobelova cena nie je výsledkom jednej originálnej myšlienky, či šťastia, ale dlhého celoživotného snaženia (malé  $m$ )

# Hirschov index tímu

***Hirschov index*** je definovaný ako počet  $h$  publikácií všetkých členov tímu, ktoré sú citované prinajmenšom  $h$  krát, pričom všetky ostatné publikácie členov tímu sú citované menej ako  $h$  krát.

Tímy s rovnakým *h-indexom* majú rovnakú-porovnateľnú vedeckú výkonnosť hoci počet ich publikácií a citácií sa môže výrazne odlišovať.

# H-index tímu

- *Hirschov index* skupiny výskumníkov, individuálne hodnoty h-indexu sú menšie ako kolektívna hodnota

$$\{h_j; j = 1 \dots N; h_j \leq h^m\}$$

$$h^m \leq h_{\text{group}} \leq h_{\text{sum}} = \sum_j h_j$$

- *Príklad:* v skupine je 10 výskumníkov. Každý má  $h=30$ , ale súčasne každý má 10 prác so 100 citáciami. Spolu majú 100 prác so 100 citáciami, teda kolektívne  $h=100$ .

# H-index školiteľa

***Hirschov index* školiteľa je definovaný ako počet  $h$  publikácií všetkých študentov školiteľa, ktoré sú citované prinajmenšom  $h$  krát, pričom všetky ostatné publikácie členov tímu sú citované menej ako  $h$  krát.**

- *Hirschov index* školiteľa môže byť väčší ako jeho individuálny h-index.
- *Príklad:* školiteľ má 10 študentov. Každý má  $h=30$ , ale súčasne každý má 10 prác so 100 citáciami. Spolu majú 100 prác so 100 citáciami, teda h-index školiteľa je  $h=100$ .

# Ponaučenia

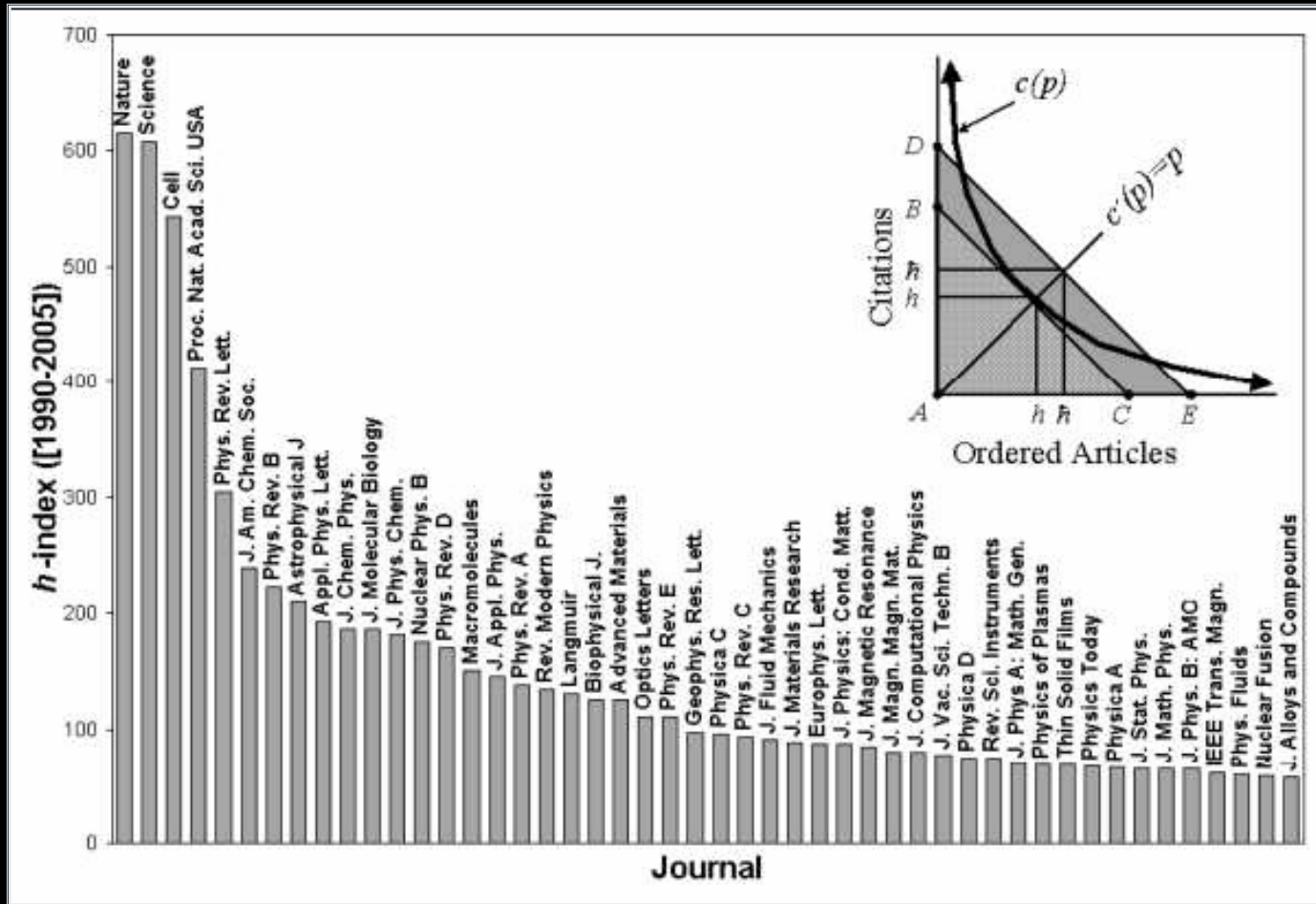
- Ak je h-index tímu rovný h-indexu jedného z riešiteľov (vedúceho tímu) potom
  - v tíme sa všetci okrem jedného ulievajú
  - vedúci tímu má okolo seba skupinu mladých ľudí
  - všetci robia na vedúceho tímu
- Kolektívny h-index by sa mal zohľadňovať pri posudzovaní výkonnosti skupín, napr. pri rozhodovaní o financovaní. Malo by sa posudzovať h v porovnaní s inými skupinami pracujúcimi v danom odbore. Resp. s h-indexom výskumnej tematiky.
- PhD študenti a postdokovia by mali hľadať tímy s vysokým h-indexom.

# Hirschov index časopisu [8]

*Hirschov index* časopisu je definovaný ako počet  $h$  článkov v časopise publikovaných od jeho vzniku (resp. za definované obdobie), ktoré sú citované prinajmenšom  $h$  krát, pričom všetky ostatné články sú citované menej ako  $h$  krát.

Časopisy s rovnakým *h-indexom* majú rovnaký-porovnateľný impakt hoci počet článkov v nich publikovaných ako i celkový počet citácií na tieto články sa môže výrazne odlišovať.

# H-index časopisov





# Hirschov index výskumnej témy

*Hirschov index* výskumnej témy je definovaný ako počet  $h$  článkov publikovaných na danú tému (na celom svete), ktoré sú citované prinajmenšom  $h$  krát, pričom všetky ostatné články na danú tému sú citované menej ako  $h$  krát.

# H-index výskumných tém

	<b>h</b>	<b>m</b>
<b>optical lattices</b>	<b>43</b>	<b>0.90</b>
<b>magnetoresistance</b>	<b>172</b>	<b>2.39</b>
<b>q-critical phenomena</b>	<b>42</b>	<b>2.63</b>
<b>spin glasses</b>	<b>108</b>	<b>3.38</b>
<b>heavy fermions</b>	<b>97</b>	<b>3.73</b>
<b>superstrings</b>	<b>99</b>	<b>3.96</b>
<b>quantum computation</b>	<b>73</b>	<b>5.21</b>
<b>M-theory</b>	<b>79</b>	<b>6.58</b>
<b>quantum dots</b>	<b>149</b>	<b>7.89</b>
<b>nanowires</b>	<b>105</b>	<b>8.75</b>
<b>carbon nanotubes</b>	<b>167</b>	<b>12.85</b>

# Ponaučenia

- Ako si mladí môžu voliť tému výskumu

$h < 100$	$m > 3$ hot topic – perspektíva pre mladých
$h > 100$	$m > 3$ dobre zavedená téma, silná konkurencia
$h > 100$	$m < 2$ etablovaná „stará“ téma, silná konkurencia

- Veľké  $h$  korešponduje veľkej komunite
- Je zaujímavé vedieť koľko ma daný vedec prác s viac ako  $h$  citáciami jeho výskumnej témy, t.j. ako sa podieľa na formovaní výskumnej oblasti v globálnom kontexte. Toto je parameter, ktorý by sa mal zohľadňovať pri posudzovaní grantov.



**Bez ružových okuliarov**

# H-index a SKVH kritéria

	Počet
<b>Publikácie CC</b>	<b>30+10</b>
<b>Citácie</b>	<b>60+20</b>
<b>H-index</b>	<b><math>h \sim 5</math></b>
<b>m (20 rokov)</b>	<b><math>m \sim 0.25</math></b>

# Ponaučenia

- Publikácie sú dôležité, pretože ich prostredníctvom informujeme o nami dosiahnutých výsledkoch
- Citácie sú ako spätná väzba – vyjadrujú relevantnosť našich výsledkov v kontexte všeobecného úsilia o získanie nových poznatkov.
- Jedným z kvantifikátorov kvality, inovatívnosti a impaktu vedeckej práce je *h-index*. Preto by sa mal zohľadňovať pri posudzovaní výkonnosti a efektívnosti výskumu. Následne pri rozhodovaní o financovaní.

**Ďakujem za pozornosť**

# Použitá literatura

1. J.E.Hirsch: “*An index to quantify and individual’s scientific research output*”, *arXiv:physics/0508025* v5 (2005); *Proc. Nat. Acad. Sci. USA* **102**, 16569 (2005); *Nature* **441**, 265 (2006)
2. M.G.Banks: “*An extension of the Hirsch index: Indexing scientific topics and compounds*” *arXiv:physics/0604216* v2 (2006)
3. C.W.Miller: “*Superiority of the h-index over the Impact Factor for physics*” *arXiv:physics/0608183* v1 (2006)
4. P.Ball: “*Index aims for fair ranking of scientists*” *Nature* **436**, 900 (2005)
5. *Rating games (editorial)* *Nature* **436**, 889 (2005)
6. P.D.Batista, M.G.Campiteli, O.Kinouchi, A.S.Martinez: “*An index to quantify an individual’s scientific research valid across disciplines*” *arXiv:physics/0509048* v4 (2005);
7. L.Bornmann & H.-D.Daniel: “*Does the h-index for ranking scientists really work?*” *Scientometrics* **65**, 391 (2005).
8. T.Braun, W.Glanzel, & A.Schubert: “*A Hirsch-type index for journals*” *The Scientist* **19**, 8 (2005)
9. S.B.Popov: “*A parameter to quantify dynamics of a researcher’s scientific activity*”, *arXiv:physics/0508113* v1 (2005);
10. I.Faktor: “*Hirschův index a česká věda aneb Domácí realita*”, *Vesmír* **85**, 555 (2006). <http://www.vesmir.cz/clanek.php3?CID=6889>



# Literatúra pre fajšmekrov

1. P.Pyykko: “*Power-law distribution of individual Hirsch indices, the comparison of merits in different fields and the relation to a Pareto distribution*”, *arXiv:physics/0608282 v1* (2006)
2. S.Lehmann, A.D.Jackson, & B.Lautrup: “*Measures and mismeasures of scientific quality*” *arXiv:physics/0512238 v2* (2005)
3. A.F.J.vanRaaij: “*Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgement for 147 chemistry research groups*” *arXiv:physics/0511206 v1* (2005)