Paper to be presented at the DRUID Academy conference in Rebild, Aalborg, Denmark on January 15-17, 2014

The Interplay of Standardization and Research: Evidence from the Biotechnology Industry

Michael Raven
Technical University Berlin / ESMT European School of Management and Technology
Technology and Innovation Management
michael.raven@esmt.org

Abstract

State-of-the-art:
Standards can limit creativity, but recent studies have established that potential drawbacks are outweighed by the benefits along the innovation process (Tassey, 2000). In the past, research in standardization has often focused on compatibility of new products from a market perspective (Farrell and Saloner, 1985). In particular, attention has been paid to formal standards by standard setting organizations (SSO) as well as information and communications technology markets (Simcoe et al., 2009). Standards consortia and company standards have been investigated as additional types of standards (Leiponen, 2008). With regard to basic research, standardization is regarded as a catalyst which facilitates technology transfers (Bozeman, 2000). However, few studies have examined the interdependencies between standardization and research (Blind and Gauch, 2009).

Research gap:
The scientific research community has recognized the need for standards in biotechnology for several years (Almeida et al., 2006). Nevertheless, until today there are no formal standards as traditional patterns of standardization do not work in the research-driven environment of biotechnology (Lonien, 2012). It remains to be resolved what drives standardization and who is involved in standardization along the research process. Thus, this paper investigates the characteristics of standards which assist technology transfers in biotechnology.

Theoretical arguments:
Deriving from contingency theory, an efficient standardization process ought to adapt to environmental conditions. Therefore, we firstly hypothesize that the biotechnology industry demands evolving standards in early research stages and continuous adoption to the technological progress. Furthermore, research findings suggest that standards play an increasing role as source of information, the more R&D activities are market-oriented (Zi and Blind, 2013). Thus, we secondly hypothesize that researchers involved in standardization publish less or in more applied journals. However, one might also argue that standardization activities circulate relevant knowledge and are beneficial for those who seek knowledge. Therefore, the higher the R&D intensity, the higher the knowledge sourcing benefits from standardization

activities. A similar pattern holds for patenting researchers which are also more successful in publishing (Agrawal and Henderson, 2002).

Method:
The implementation of our study is based on a database of informal standards assembled from BioSharing and Web of Knowledge (Thomson Reuters). BioSharing is a community-based online platform which registers standards for describing and sharing biosciences. In analogy to Wikipedia, the usage of BioSharing is free of charge. Everybody can register and contribute. Our database consists of 515 standards categorized according to different types of standards. Thereof, 91 standards are linked to scientific publications. The publications relate to 56 different biotechnology journals in the time period between 1996 and 2011.

Results:
In agreement with our first hypothesis, we retain a new kind of standardization, i.e. community-driven standards. They combine benefits of formal and informal standards: Community-driven standards allow for an open and transparent, but still adjustable and quickly developed standardization process. In analogy to the formation of consortia, community-driven standards can be ascribed to coalition theory.

With regard to the second hypothesis, we find that standards are linked to journals with an aggregated one-year impact factor of 10.43. In comparison, a control group of biotechnology journals reported by the Journal Citation Reports shows an aggregated one-year impact factor of 3.78. In addition, each publication linked to a standard is matched to the most-related article in the same volume of the same journal using a search algorithm developed by the National Library of Medicine. A negative binomial model confirms that standards are more valuable than their closest peer group in terms of forward citations. A Wilcoxon signed-rank test verifies robustness of the results even under non-parametric assumptions at the 5% significance level.

The results are further refined by investigating the type of standards important to research activities. Our database contains 327 semantic standards, 57 measurement and testing standards, 127 interface standards, as well as 4 compatibility, quality and variety reducing standards. An extension of the model reveals that especially measurement and testing standards are most valuable to future research in terms of forward citations.

Overall, our findings alter the understanding of different categories of standards and their respective usage. We conclude that standardization is an essential building block in the research process driven by highly skilled scientists.

Jelcodes:O32,O31

Working paper

DRUID Academy Conference 2014

The Interplay of Standardization and Research:

Evidence from the Biotechnology Industry

JEL code at level 3: *O32 – Management of Technological Innovation and R&D*
Additional JEL code: *O31 – Innovation and Invention: Processes and Incentives*
Paper profile: *Innovation Management*
Main type of paper: *Econometric Evidence*

Michael Raven

michael.raven@esmt.org

Technical University Berlin
Straße des 17. Juni 135, 10623 Berlin, Germany

European School of Management and Technology
Schloßplatz 1, 10178 Berlin, Germany

Berlin, December 16, 2013

ABSTRACT

In the past, research in standardization has often focused on compatibility of new products from a market perspective. With regard to basic research, standardization is regarded as a catalyst which facilitates technology transfers, but few studies have examined the interdependencies between standardization and research. It remains to be resolved what drives standardization and who is involved in standardization along the research process. This paper investigates the characteristics of standards which assist technology transfers in biotechnology. We point out that the biotechnology industry relies on evolving standards in early research stages and continuous adoption to the technological progress. Despite recent findings with regard to formal standards linked to commercialization we show that community-driven standards are a complementary standardization tool which is mainly used in basic research. Furthermore, we find that within basic research articles related to standards are used more often than their closest peer group in terms of forward citations. In addition, the negative binomial model reveals that especially interface standards are most valuable to future research.

1. Introduction

The fundamental purpose of standards is to enable interoperability and coordination. Standards arguably limit creativity in the innovation process, but recent studies have established that potential drawbacks of standards are outweighed by the benefits (Allen & Sriram, 2000; Tassey, 2000; Temple, Blind, Jungmittag, Spencer, & Witt, 2005). Standards reduce the costs of innovating by narrowing the set of product differentiations while promoting interdependent research tasks (Baldwin & Clark, 2000). With regard to the innovation process as a whole, standardization is regarded as a catalyst which facilitates technology transfers (Bozeman, 2000).

In the past, research in standardization has often focused on compatibility of new products from a market perspective (Farrell & Saloner, 1985). Most attention has been paid to formal standards by standard setting organizations (SSO) as well as information and communications technology markets (Simcoe, Graham, & Feldman, 2009). However, few studies have investigated the interdependencies between standardization and research (Blind & Gauch, 2009; Zi & Blind, 2013).

Interoperability and coordination of research activities is of particular importance in industries which rely on different disciplines, technologies and skills such as the biotechnology industry (Gillis, 2003). Especially due to the vastly increasing amount of data the research community has recognized the need for efficient standards for several years (Almeida et al., 2006; Quackenbush, 2006; Wang, Gorlitsky, & Almeida, 2005). However, until today there exist no international standards by standard-setting organizations (SSO) in biotechnology. The German Institute for Standardization (DIN) recently proposed the establishment of a technical committee (TC) at the International Organization for Standardization (ISO) with respect to biotechnology (Lonien, 2012), but, so far, no progress could be accomplished. In 1995 the DIN has already released 166 standards in affiliation with the European Committee for Standardization (CEN) in three particular subfields of Biotechnology. However, these standards are mostly generic and limited to large-scale production, performance criteria and criteria for reaction vessels. Therefore, they are of little relevance to the industry, e.g. *DIN EN 1619* states "before putting to use, they [working cell banks] should be characterized with an appropriate degree of accuracy" without further specifying the appropriate degree of accuracy.

Industry experts have pointed out that informal standards might be more suited to the needs of biotechnology as traditional patterns of standardization do not work (Rai, 2010). Especially in medical biotechnology, several de-facto standards have evolved. Regulatory authorities such as the *US Food and Drug Administration* (FDA) and the *European Medicines Agency* (EMA) impose increasing requirements on the market entry of new products. Concepts of the regulatory authorities such as *Good Manufacturing Practices* (GMP) and *Quality by Design* (QbD) diffuse into the whole industry, even outside of the regulatory framework. The *World Health Organization* (WHO) affects standardization in medical biotechnology by publishing the *WHO Technical Report Series* (TRS) as well as providing reference preparations which serve as measurement standards. Furthermore, few consortia have evolved in medical biotechnology supporting the process of drug developments. Two examples are the *CMC-Biotech Working Group* (CMC-BWG) and the *Predictive Safety Testing Consortium* (PSTC). CMC-BWG publishes practitioner guidelines which support the standardization of quality requirements and PSTC assists the standardization of biomarkers. Noticeable across the whole biotechnology industry is the establishment of certain technological platforms which can also be seen as de-facto standards, e.g. host organisms such as *Escherichia coli* and *Chinese Hamster Ovary Cells.*

The question arises, which necessities are current efforts of standardization in biotechnology confronted with. As standardization in the beginning of the innovation process has received little attention, this paper investigates the properties of standards at the research stage of the innovation process as well as how these standards assist technology transfers at the example of the biotechnology industry. Furthermore, we examine who is involved in standardization along the research process; whether standardization is driven by top researchers or rather by practitioners. The aim of the study is to gain a better understanding of the role of standardization along the research process.

2. PROPERTIES OF STANDARDS IN RESEARCH

Standards procure the diffusion of technology, as part of the innovation system (Besen & Farrell, 1994). This matter nourishes and creates value for the R&D process as well as other investments in knowledge creation (Temple et al., 2005). Overall, standards are a source of relevant information to opinion leaders in the innovation system. This implies that the research community as well as the standards community constantly monitor, alert, and match standardization efforts. On the one hand, the research community pulls information for research and pushes information on standardization. On the other hand, the same holds true vice versa for the standards community. In order to understand the interdependencies we need to define the properties of standards in research as well as the relation between research and standards communities.

So far, the literature differentiates between three categories of standards. First of all, formal standards are established by SSOs such as ISO or CEN. Formal standardization follows a strict procedure which is transparent to the stake holders and guarantees a high level of consensus, but can also be tedious and costly. Furthermore, consortia standards and company standards have been specified as additional categories of standards (Blind & Gauch, 2008; Leiponen, 2008). In contrast to formal standards, they evolve from an exclusive group or arrangement. Therefore, not all interests of all stake holders are always considered and lower levels of consensus are achieved. The main advantages are the faster development cycles as well as the general flexibility of these types of standards.

Deriving from contingency theory, an efficient standardization process adopts to the environmental conditions. If we combine what we know about biotechnology and standards, we can hypothesize that early research stages in biotechnology benefit from evolving standards with continuous adoption to the technological progress, which is neither served by SSOs nor by standards consortia.

In our analysis, we define an alternative category of standards which has not been described in the literature beforehand and can be found in an online-platform of biotech standards named *BioSharing*. The origin of *BioSharing* lies in the United Kingdom in 2009. In total, 35 community members, mainly research communities, but also the DIN, contribute to the platform. *BioSharing* is free of charge. Everybody can register and contribute to the platform. The mission of *BioSharing* is to "serve those seeking information on existing standards, but also to […] promote harmonization to stop wasteful reinvention." Thereby, the platform provides community-driven standards which combine benefits of formal standards as well as company and consortia standards. In analogy to formal standards, community-driven standards are transparent and easily accessible, but at the same time they are flexible and require only short development times similar to company and consortia standards. In analogy to the formation of consortia, community standards are based on coalition theory (Olson, 1971).

Blind and Gauch (2007) have pointed out the obstacles for researchers to engage in standardization, although standards ought to evolve in parallel to the research process. With regard to the content of standards in the research process they differentiate between three types of standards: Firstly, after pure basic research semantic standards have to be established in order to define a clear terminology. The definition of terms is necessary in order to enable the communication between researchers, product and process developers. Secondly, after oriented basic research, measurement and testing standards are a prerequisite for the verification of product development. For example, agreements about minimum information guidelines ensure that cartilage cells are identified across project and country boundaries. Thirdly, interface standards mediate the progress from applied research to experimental development. For example, research projects in biotechnology often deal with interfaces when partial research results ought to be integrated into an all-embracing database or a complex overall model. Afterwards, compatibility and quality standards facilitate the diffusion of new products, but will not be discussed in greater detail in this paper.

Within the *BioSharing* platform standards are classified according to “terminology artifact”, “reporting guidelines”, and “exchange format”. In combination with key words such as “ontology” for semantic standards, “minimum information” for measurement and testing standards, and “File XXX” for interface standards an accurate matching with regard to the content of the standards was ensured. Overall, 64.5% of the standards in *BioSharing* refer to semantic standards, 11.6% of the standards deal with measurement and testing standards, and 24.9% of the standards define interface standards. Biotechnology being a rather new scientific discipline it comes as no surprise that the majority of standards address terminology issues. Furthermore, the development of *high-throughput-screening* and *next-generation-sequencing* have increased the need for interface standards dealing with complex databases.

3. The Link between Standardization and Research

Research findings about formal standards suggest that standards play an increasing role as source of information the more R&D activities are market-oriented (INTegrating REsearch and Standardisation / INTEREST). Past research has argued that the use of formal standards as inventive input is positively linked to the degree of commercialization (Working Paper, Timothy Simcoe). Furthermore, researchers involved in formal standardization publish less or in lower ranked journals (Zi & Blind, 2013). In line with these findings, we derive the following first hypothesis:

Hypothesis 1: *Standards are mainly used in applied research and less used in basic research.*

In spite of the above mentioned research findings, there is also a line of arguments in favor of rejecting the first hypothesis. In a related research field, it is established that patenting researchers are more successful in publishing (Agrawal & Henderson, 2002; Czarnitzki, Glänzel, & Hussinger, 2007, 2009; Stephan, Gurmu, Sumell, & Black, 2007; Van Looy, Callaert, & Debackere, 2006). In analogy to the field of patents, standardization activities arguably circulate relevant knowledge and are beneficial for those who seek knowledge. Therefore, one can also argue the higher the R&D intensity, the higher the knowledge sourcing benefits from standardization.

From the general definition of standards we can derive that standards avoid misconceptions in joint research projects. Through the unification of methods in subsequent research efforts scientists can be more productive. Consequently the online-platform *BioSharing* has pointed out in its mission to stop wasteful reinventions. Therefore, we argue in our second hypothesis that standards are an essential element for research progress. Thus, articles with regard to standards are used more often than comparable articles.

Hypothesis 2: *In basic research, articles concerned with standards are crucial building blocks and are used more often than comparable articles.*

4. METHODS

4.1. DATA AND MEASUREMENT

The implementation of our study is based on two databases: a catalogue of informal standards provided by *BioSharing* and the `Web of Knowledge´ provided by *Thomson Reuters* (Web of Science, Journal Citation Reports). As already pointed out *BioSharing* is a community-based online platform which started in 2009. It registers standards for describing and sharing biosciences experiments and promotes harmonization in the biological sciences. The operating mode of *BioSharing* can be described in analogy to the *Wikipedia* encyclopedia: It is free of charge, everybody can register and contribute.

As of May 2013, *BioSharing* consists of 518 standards which are categorized according to different types of standards. 95 standards are linked to 98 scientific publications which allow an indepth analysis of the relationship between research and standardization. The publications relate to 56 different biotechnology journals in the time period between 1996 and 2013.

As an operationalization of the first hypothesis we compare the arithmetic mean of the impact factors of the journals which publish standards to the arithmetic mean of journals in biotechnology where the `Journal Citation Reports´ by *Thomson Reuters* provide a peer group of biotechnology journals. Thereby, we assume that journals in basic research have higher impact factors than more applied journals which relate to more market-oriented research (Garfield, 1972, 2006).

With regard to the operationalization of the second hypothesis we differentiate between publications directly linked to standards and publications using these standards publications: Each publication linked to a standard in *BioSharing* is matched to the most-related article in the same volume of the same journal using a search algorithm developed by the National Library of Medicine (NLM) in order to constitute a meaningful control sample (Furman &

Stern, 2011). As a robustness check, a second control sample is constituted on the basis of the most-related article without taking the volume and the journal into account.

Either forward citations in the first year after publication, forward citations in year t, or cumulative forward citations since publication serve as dependent variables (Garfield, 1979). Table 1 provides an overview of the information for the dependent variables.

Table 1 | Overview of dependent variables

Variable	Definition	Source
Journal	Name of the publishing journal	Science Citation Index (SCI)
Impact factor	One-year and five-year impact factor of publishing journal	Journal Citation Reports
Forward Citations$_{jt}$	Number of Forward Citations to article j in year t	SCI
Cumlulative Citations$_{jt}$	Number of Forward Citations from publication date to year t	SCI
Age	Year t – article publication year	SCI
Average Forward Citations$_{j1}$	Average Forward Citations$_{j1}$ to articles in the same journal with t=1	Journal Citation Reports

With regard to the independent variables we introduce a dummy variable `standard article´ which differentiates between articles related to standards and the control group. In a more sophisticated model we further differentiate between the three different types of standards. The article title is used as an identifier across the different databases. Furthermore, we stepwise extend the model by controlling for the publication year, the number of authors, the top 50 universities, and the journal category. Table 2 provides an overview of the information for the independent variables.

Table 2 | Overview of independent variables

Variable	Definition	Source
Standard article	Dummy variable equal to 1 if article is associated with a standard and equal to 0 if article belongs to the control group	BioSharing / PubMed
Type of standard	Dummy variables equal to 1 for *semantic standards, measurement and testing standards* as well as *interface standards* respectively and equal to 0 otherwise	BioSharing / PubMed
Article title	Name of article	BioSharing / PubMed / SCI
Names of authors	Names of publishing authors	SCI
Number of authors	Number of publishing authors	SCI
Country of Origin	Origin of lead author of the publication	SCI
Top 50 University	Dummy variable equal to 1 if corresponding author is associated with an institution that appears in the top 50 according to FT ranking	
Journal category	Dummy variable equal to 1 if the publishing journal has an impact factor above 10	Journal Citation Reports
Publication year	Year in which article j is published	SCI

4.2. Summary Statistics

Our data consists of 95 standards linked to 98 articles on the *BioSharing* platform. For the second hypothesis the data processing results in 196 articles, i.e. 50 percent of the articles relate to standards and 50 percent relate to the control group (Table 3). On average the articles are published in 2008 and on average approximately 11 authors participate in one article. The forward citations are collected for each year after the publication. Therefore, a sample of 1,160 citations results with a mean of 13.23 forward citations per article per year. On a cumulative basis each article receives on average 78.33 forward citations.

Table 3 | Means and standard deviations

Variable	Mean	SD	Min	Max
Article characteristics (n = 196)				
Standard article$_j$	0.50	0.50	0	1
Publication year	2008.01	3.17	1996	2013
Authors	10.86	13.99	1	98
Citation characteristics (n = 1,160)				
Forward Citations$_{jt}$	13.23	41.28	0	682
Cumulative Citations$_j$	78.33	224.5976	0	1962
Age	4.99	3.17	0	17

Figure 1 provides an overview of the distribution with regard to the publication year. The majority of standards have been recently published. 2013 being the end of the observation period, the articles are on average five years old. Only for few standards forward citations for more than six years are available.

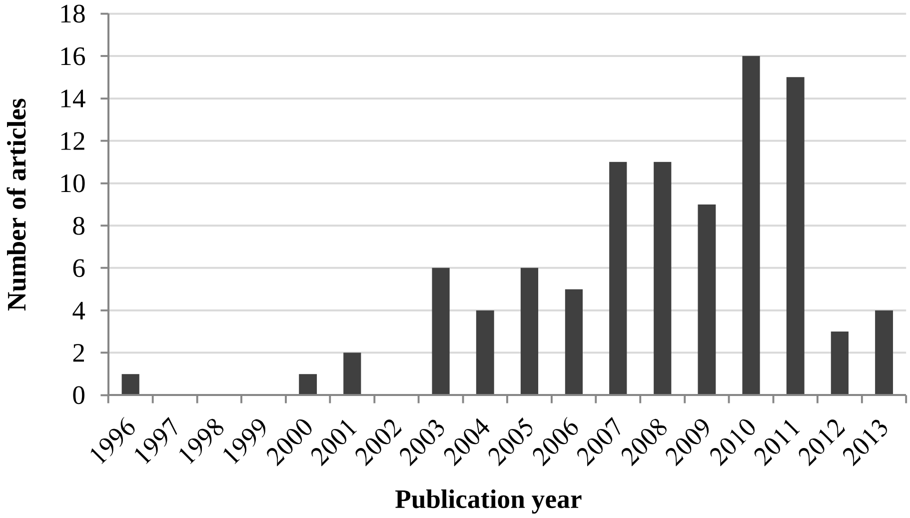

Figure 1 | Distribution of publication year

In terms of the country of origin, the sample is representative of the biotechnology research landscape. The articles relating to standards are partially biased towards the United Kingdom as the origin of the *BioSharing* database is also in the United Kingdom. Figure 2 provides an overview of the distribution with regard to the country of origin.

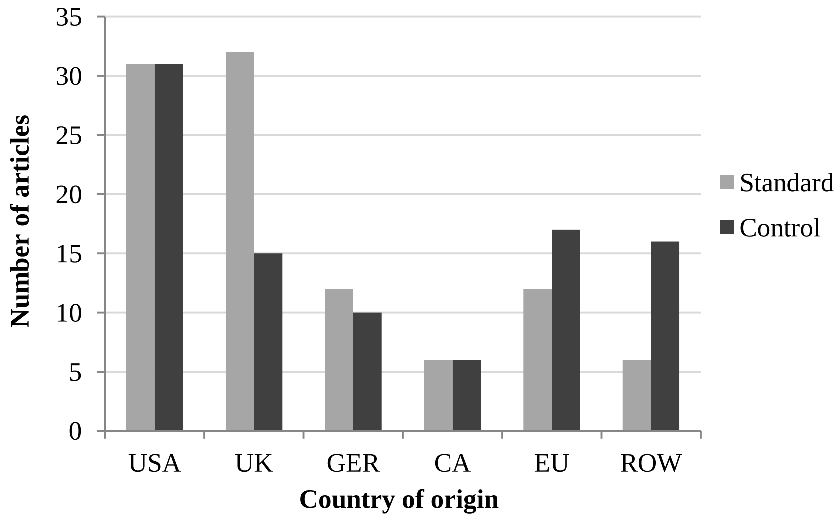

Figure 2 | Distribution of publication year

One particular characteristic of the forward citations is the skewed distribution. Forward citations are count data, i.e. greater or equal to zero. Furthermore, most articles are cited five times or less per year. Very few articles receive hundreds of citations in one year (Figure 3).

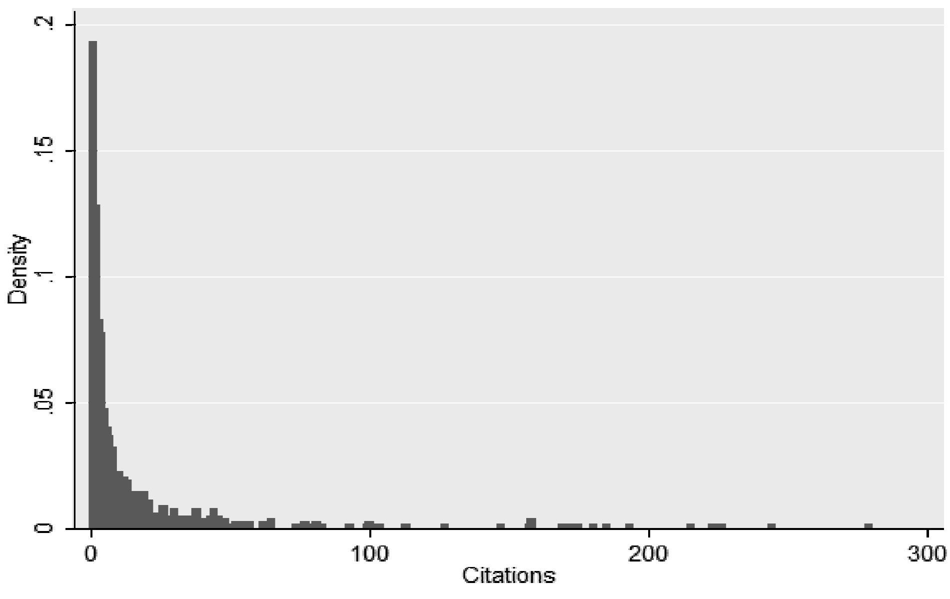

Figure 3 | Distribution of citations

An investigation of the descriptive statistics differentiated by the control group already provides a good overview about the empirical findings. On average, standard articles are cited 2.8 times more often than control articles. Noticeable is that also 2.9 times more authors per publication are involved with standard articles in comparison to control articles (Table 4).

Table 4 | Means and standard deviations, by control group

	Standard articles: Articles associated with BioSharing	Control articles: Most-related article control
Number of papers	98	98
Forward citations	19.45 (53.69)	6.99 (21.13)
Cumulative Citations	115.33 (292.12)	41.33 (115.36)
Authors	16.09 (13.99)	5.52 (4.75)
Publication Year	2008 (3.17)	2008 (3.17)

As previously done for patents (Mehta, Rysman, & Simcoe, 2010), the citation age profile for standards illustrates the diffusion of standards compared to control articles over time (Figure 4). In general, standard articles are cited significantly more often than control articles. Noteworthy is that standard articles also pursue a different time trend than the control articles. In the first two years, forward citations to both groups rise, while standard articles receive more forward citations. However, in the second year after publication, the control articles reach a maximum and subsequently decline. In contrast, standard articles reach their maximum in the third year after publication and continue to stay at a relative high level, although we have to point out that after six years the confidence interval dramatically increases due to the lack of available data. The intervals around the median show the 90 percent confidence intervals and are constructed according to Conover (1980).

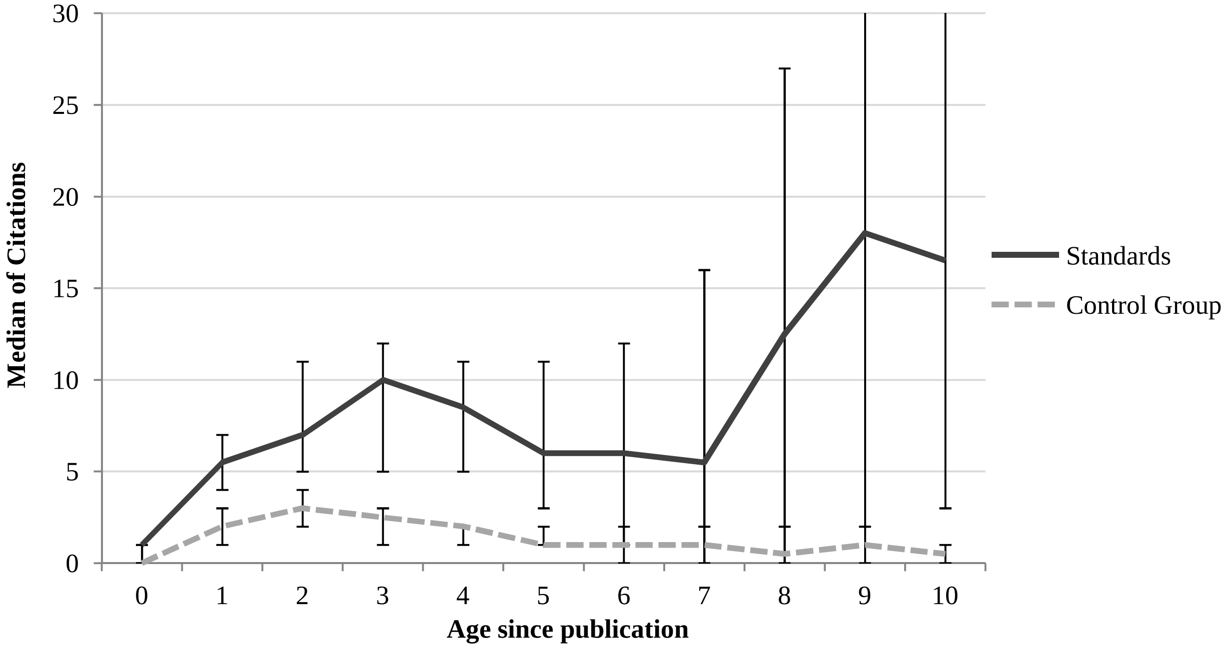

Figure 4 | Citation age profile

An overview of the type of standards shows that the different content of the standards is not proportionally linked to articles within the online-platform (Figure 5). As of July 2013, the online-platform contains 329 semantic standards, 60 measurement and testing standards, and 129 interface standards. While the majority of standards in *BioSharing* consist of semantic standards, only few semantic standards are linked to publications in journals, i.e. only 6.1 percent of the semantic standards. The opposite holds true for measurement and testing standards, i.e. 73.3 percent of the measurement and testing standards are linked to publications in journals. Therefore, even within the community standards different types of standards might diffuse differently or have diverse relevance to the biotechnology field.

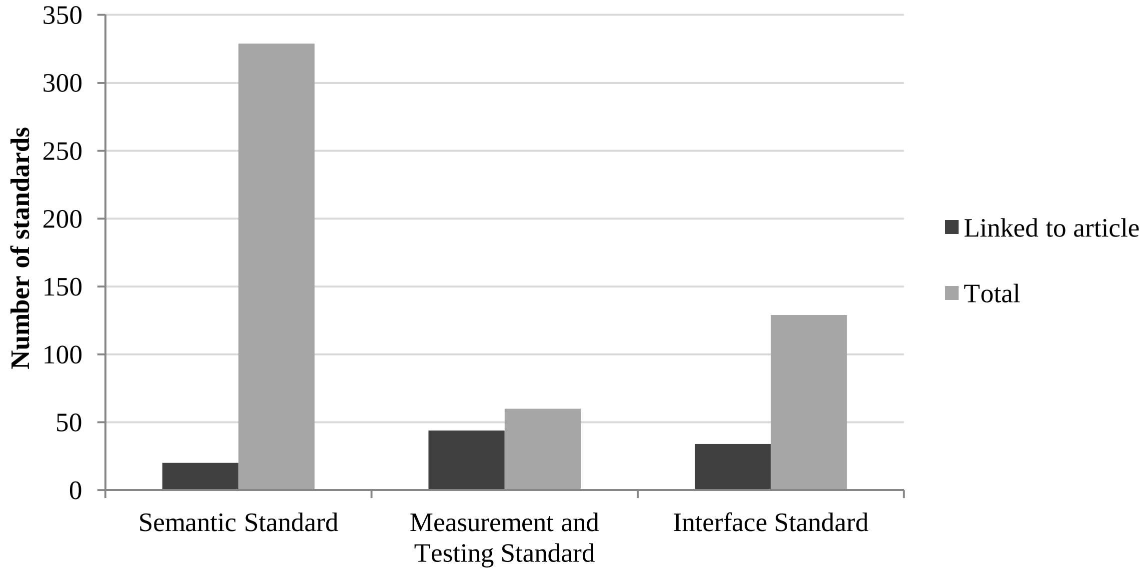

Figure 5 | Type of standards

5. RESULTS

With regard to the first hypothesis we find that standards in our database are linked to journals with an aggregated one-year impact factor of 10.07 (median impact factor: 5.32). In comparison, a control group of biotech journals reported by the `Journal Citation Reports´ (JCR) shows an aggregated one-year impact factor of 3.78 (median impact factor: 2.47). Given the 95 percent confidence intervals, we can conclude that standards are significantly more likely to publish in biotechnology journals with higher impact factors (Table 5). Assuming that higher impact factors relate to basic research (Garfield, 1972, 2006), we retain that community-driven standards are crucial in early research activities. Therefore, community standards can be regarded as complementary standardization tools in view of the mainly market-driven usage of formal standards and hypothesis 1 must be disregarded.

Table 5 | Mean and median of the impact factor, by biotechnology journals

	Impact factor of journals publishing standards	Peer group of biotech journals according to JCR
Arithmetic mean	10.07	3.78
Confidence interval to the mean (95% significance level)	7.95 – 12.20	--
Median	5.32	2.47
Confidence interval to the median (95% significance level)	4.20 – 6.53	--

In order to investigate the role of standards within basic research, we examine publications linked to standards in comparison to their peer group in the same journal. In a first step, we compare forward citations in the first year after publication of standard articles with average forward citations to articles in the same journal. A paired t-test shows at the 1 percent significance level that standard articles receive on average one and a half times more forward citations in the first year than articles in the same journal. Since the standard articles are also included in the average forward citations to the journal and we only consider the first year after publication, this test can be regarded as a very conservative test. In a second step, we expand the paired t-test by comparing cumulative forward citations of standard articles with most-related articles published in the same journal in the same volume. The paired t-test reveals at the 1 percent significance level that standard articles receive on average 2.8 times more forward citations than the control group over the entire lifetime of a publication. As the paired t-test relies on a normal distribution we have to review the robustness of the above mentioned results. Therefore, we use a Wilcoxon signed rank sum test as non-parametric robustness check (Hollander & Wolfe, 1973). The Wilcoxon test for the difference in means without any assumption about the distribution remains significant at the 5% significance level.

Although one can argue that normal distribution is a reasonable assumption for cumulative forward citations, figure 3 in the previous chapter clearly demonstrates that the assumption does not hold true for forward citations on a yearly basis. Therefore, we cannot use an ordinary least square regression model, but have to account for the characteristics of count data by applying a poisson model or a negative binomial model. Since we are confronted with over-dispersed count data, i.e. the conditional variance is larger than the conditional mean, we have chosen a negative binomial model over a poisson model in analogy to Furman & Stern (2011). However, all reported results also hold true in a poisson model at the same level of significance or even higher.

Our baseline model identifies the effect of a standardization article on yearly forward citations. Therefore, the first model estimates the following equation:

$$(1) \qquad \textit{Forward Citations}_{jt} = f\ (\textit{Standard article}_j)$$

The first column of table 6 reports the results for the baseline model. A significant positive coefficient implies a positive relation between independent variable and forward citations. A positive coefficient translates into an incidence-rate ratio greater than one while a negative coefficient translates into an incidence-rate ratio smaller than one. Specifically, standard articles receive 2.781 more forward citations than most-related articles. The intersection with the y-axis is statistically significant accounting for the fact that also the control articles receive forward citations. However, the pseudo R^2 value shows that the baseline model only has limited explanatory power.

Table 6 | Negative Binomial Model

	Negative Binomial Model [Incidence-rate ratios in brackets in top line] Estimated coefficients in 2nd line (SE in parantheses in bottom line)	**Negative Binomial Model with Conditional Fixed Effects (1)** [Incidence-rate ratios in brackets in top line] Estimated coefficients in 2nd line (SE in parantheses in bottom line)	**Negative Binomial Model with Conditional Fixed Effects (2)** [Incidence-rate ratios in brackets in top line] Estimated coefficients in 2nd line (SE in parantheses in bottom line)
Standard article	[2.781] 1.023 (0.104)***	[2.052] 0.719 (0.087)***	
Semantic standard			[1.356] 0.304 (0.171)*
Measurement and testing standard			[1.686] 0.522 (0.120)***
Interface standard			[3.613] 1.285 (0.138)***
Age		[1.059] 0.058 (0.016)***	[1.059] 0.058 (0.016)***
Authors		[1.028] 0.028 (0.004)***	[1.025] 0.025 (0.004)***
Number of article pairs		98	98
Constant	1.945 (0.074)***	0.013 (0.329)	0.278 (0.327)
Observations	1,160	1,160	1,160
Pseudo R^2	0.012	0.173	0.177
Log likelihood	-3,656.24	-3,044.90	-3,030.84
Test of alpha=0	3.01 (0.128)***	0.858 (0.049)***	0.830 (0.048)***

*** Significance at the 1 percent level
** Significance at the 5 percent level
* Significance at the 1 percent level

As an extension of the baseline model we account for conditional fixed effects over time. Namely, dummy variables are included for each article pair (Allison & Waterman, 2002; Hausman, Hall, & Griliches, 1984). Furthermore, we have already reported in figure 4 that at least in the first years after publication citations rise over time. Therefore, we also include the age of the publication in the model. In addition, the summary statistics have shown that on average more authors are involved in standard articles than in control articles. This comes as no surprise as standards require a level of consensus which is more likely to be achieved if many authors are involved in the publication process. However, researchers are also more likely to cite their own publications in future research than random articles. Therefore, we have to disentangle the effect of the standardization from the self-citing effect, i.e. more authors involved in standard articles are more likely to self-cite the standard articles. The following equation follows for the second model:

$$(2) \qquad \textit{Forward Citations}_{jt} = f\ (\textit{Standard article}_j + \textit{Age}_t + \textit{Authors}_j + \textit{Article pairs}_j)$$

The second column of table 6 reports the results for the negative binomial model with conditional fixed effects. Standard articles are still positively correlated to forward citations at the 1 percent significance level, although the coefficient slightly decreased. Standards receive approximately two times more forward citations than control articles. As expected also the age of the publication is positively linked to forward citations. On average one additional year since the publication year accounts for approximately six percent more forward citations. Furthermore, also the number of authors has a statistically significant influence on the number of forward citations. Therefore, we can conclude that the self-citation effect of one additional author on average leads to an increase of approximately three percent of forward citations. However, the disproportionately high number of authors in standard articles are not accountable for the overall standardization effect. In contrast to the first model, the intersection with the y-axis is not significant anymore, but the explanatory power of the extended model has increased up to 17.3 percent.

These results are further refined by investigating the type of standards important to research activities. Instead of relying on a dummy variable for the standard articles, we introduce three new dummy variables for each type of standard:

$$(3) \qquad \textit{Forward Citations}_{jt} = f\ (\textit{Type of standard}_j + \textit{Age}_t + \textit{Authors}_j + \textit{Article pairs}_j)$$

The new model reported in the third column of table 6 illustrates the diverse relevance of different type of standards. All types of standards are positively linked to forward citations, but the respective size of the effect is different. Interface standards are significantly correlated at the one percent level and receive 3.6 times more forward citations than the control articles. Therefore, interface standards are most valuable to future research in terms of forward citations. Measurement and testing standards are also significant at the one percent level, but already have much smaller relevance in terms of forward citations. Measurement and testing standards receive on average approximately 69 percent more forward citations than control articles. Semantic standards are only linked to forward citations at the ten percent significance level. Furthermore, semantic standards have a relatively small effect on forward citations as forward citations increase by 36 percent when referring to semantic standards.

Overall, the alpha test for all three model specifications is significant at the one percent level. Thus, the overdispersion of the data is confirmed and the application of a negative binomial model is reinforced.

6. Discussion

In a first step we have recognized a new kind of standardization process in our database. In contrast to SSOs, standards consortia, and corporate standards, our database contains of community driven standards. They combine benefits of formal and informal standards since standardization in its traditional form cannot properly address the current needs of the biotechnology industry. Community-driven standards allow for an open and transparent, but still adjustable and quickly developed standardization process.

Based on the above mentioned findings we reject the first hypothesis. On the contrary, the relative meaning of standards in publications suggest that standardization is highly relevant to basic research in biotechnology. This finding supports the line of argument that standards serve as a crucial source of information for basic researchers. Instead of limiting standardization to market-oriented challenges our paper points out a complementary standardization tool which serves the needs of basic research.

Our empirical results confirm the argument of the second hypothesis. Standards are essential elements of basic research. Thus, they are cited more often than their peer group. Furthermore, the development of citations over time has revealed that standards provide a different path of diffusing knowledge than the control articles. While the lifecycle of the control articles is approximately five years with a maximum in the second year, standard articles experience a longer increase of forward citations and do not show a definite decline of forward citations afterwards. The lifecycle of standard articles is longer than their peer group.

With regard to the content of standards, interface standards are most important for the research process. This finding underlines the necessity of standards especially in dispersed and interdisciplinary research fields such as biotechnology. However, these findings are most likely transferable to similar fields such as nanotechnology, medical engineering, or neurosciences.

Overall, these findings enhance the understanding of the interplay between standards and research and alter the concept of different categories of standards. The empirical results clarify the meaning of standards for the research progress.

6.1. Limitations and Further Research

The main rationale for this quantitative study is an empirical characterization of the complex relation between research and standardization. While our study investigates the issue on the basis of research articles as unit of analysis, the investigation still needs to be expanded to the individual level of researchers where authors serve as a unit of analysis. Although some of the findings already enhance the understanding of the conceptual mechanisms, detailed qualitative studies are still needed to investigate incentives, personal preconditions, and systemic drivers for standardization in research. Our studies have not yet investigated the implications of research standards for the transfer of R&D results through standardization. While we differentiate between different types of standards, the quantitative effects on the technology transfer model still need to be resolved. It remains for further research to explore the complementarity of community-driven standards and the three previously explored categaries of standards, i.e. consortia standards, company standards, and formal standards. In general, it is left to future research to quantify the economical benefits of standardization along the research process.

The focus of our study on the biotechnology industry allows us an in-depth analysis of the research questions without any confusion due to different industry backgrounds. As a shortcoming of such a specific data sample generalizability of the results might be questionable. However, standardization in the biological sciences requires flexible, easily accessible, and quickly developed standards due to a rapidly changing technology basis and multidisciplinary challenges. We assume that many other high-tech industries have to comply with these requirements and our results can be transferred to other industries.

Further potential limitations of our study arise with regard to our data sample. First of all, approximately 20 percent of the standards in the *BioSharing* database are linked to research articles while - potentially due to a bias - 80 percent of the standards are not linked to a particular publication source. The data sample does not contain an announcement date of the standards within the *BioSharing* database. Therefore, we cannot disentangle whether the inclusion in the online-platform has a significant effect on the citation age profile.

6.2. Implications

Our research findings have various implications for future research, SSOs, as well as policy makers. With regard to future research we have to expand the tool box of different standard categories. Furthermore, our findings enable the augmentation of ones horizon in view of the challenges where standardization facilitates progress. Based on our findings SSOs can adjust their focus of standardization efforts and further engage in early innovation stages of the product lifecycle. We already point out that interface standards are of particular importance. Last but not least, policy makers can optimize their sponsorships in order to overcome market failures especially with regard to basic research. Thus, policy makers should also emphasize standardization in basic research.

References

Agrawal, A., & Henderson, R. 2002. Putting Patents in Context: Exploring Knowledge Transfer from MIT. ***Management Science***, 48(1): 44–60.

Allen, R. H., & Sriram, R. D. 2000. The role of standards in innovation. ***Technological Forecasting and Social Change***, 64(2-3): 171–181.

Allison, P. D., & Waterman, R. P. 2002. Fixed–Effects Negative Binomial Regression Models. ***Sociological Methodology***, 32(1): 247–265.

Almeida, J. S., Chen, C., Gorlitsky, R., Stanislaus, R., Aires-de-Sousa, M., Eleuterio, P., et al. 2006. Data integration gets "Sloppy." ***Nature Biotechnology***, 24(9): 1070–1071.

Baldwin, C. Y., & Clark, K. B. 2000. ***Design Rules: The power of modularity***. MIT Press.

Besen, S., & Farrell, J. 1994. Choosing How to Compete - Strategies and Tactics in Standardization. ***Journal of Economic Perspectives***, 8(2): 117–131.

Blind, K., & Gauch, S. 2008. Trends in ICT standards: The relationship between European standardisation bodies and standards consortia. ***Telecommunications Policy***, 32(7): 503–513.

Blind, K., & Gauch, S. 2009. Research and standardisation in nanotechnology: evidence from Germany. ***The Journal of Technology Transfer***, 34(3): 320–342.

Bozeman, B. 2000. Technology transfer and public policy: a review of research and theory. ***Research Policy***, 29(4-5): 627–655.

Conover, W. J. 1980. ***Practical nonparametric statistics*** (2nd ed.). New York: John Wiley and Sons.

Czarnitzki, D., Glänzel, W., & Hussinger, K. 2007. Patent and publication activities of German professors: An empirical assessment of their co-activity. ***Research Evaluation***, 16(4): 311–319.

Czarnitzki, D., Glänzel, W., & Hussinger, K. 2009. Heterogeneity of patenting activity and its implications for scientific research. ***Research Policy***, 38(1): 26–34.

Farrell, J., & Saloner, G. 1985. Standardization, Compatibility, and Innovation. ***The RAND Journal of Economics***, 16(1): 70–83.

Furman, J. L., & Stern, S. 2011. Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research. ***American Economic Review***, 101(5): 1933–1963.

Garfield, E. 1972. Citation Analysis as a Tool in Journal Evaluation - Journals Can Be Ranked by Frequency and Impact of Citations for Science Policy Studies. ***Science***, 178(4060): 471–+.

Garfield, E. 1979. Is Citation Analysis a Legitimate Evaluation Tool. ***Scientometrics***, 1(4): 359–375.

Garfield, E. 2006. The history and meaning of the journal impact factor. ***Jama-Journal of the American Medical Association***, 295(1): 90–93.

Gillis, M. 2003. Harnessing new technologies for the 21st century. ***Proceedings***, (Sep): 63–75.

Hausman, J., Hall, B., & Griliches, Z. 1984. Econometric-Models for Count Data with an Application to the Patents R and D Relationship. ***Econometrica***, 52(4): 909–938.

Hollander, M., & Wolfe, D. A. 1973. ***Nonparametric statistical methods*** (2nd ed.). New York: Wiley.

Leiponen, A. E. 2008. Competing Through Cooperation: The Organization of Standard Setting in Wireless Telecommunications. ***Management Science***, 54(11): 1904–1919.

Lonien, J. 2012. A blessing in disguise: International standards for the biotechnology industry. ***Biotechnology Journal***, 7(11): 1321–1323.

Mehta, A., Rysman, M., & Simcoe, T. 2010. Identifying the Age Profile of Patent Citations: New Estimates of Knowledge Diffusion. ***Journal of Applied Econometrics***, 25(7): 1179–1204.

Olson, M. 1971. ***The Logic of collective action: public goods and the theory of groups***. Cambridge, Mass: Harvard University Press.

Quackenbush, J. 2006. Standardizing the standards. ***Molecular Systems Biology***, 2.

Rai, A. 2010. Law and technology: Unstandard standardization: the case of biology. ***Commun. ACM***, 53(1): 37–39.

Simcoe, T. S., Graham, S. J. H., & Feldman, M. P. 2009. Competing on Standards? Entrepreneurship, Intellectual Property, and Platform Technologies. ***Journal of Economics & Management Strategy***, 18(3): 775–816.

Stephan, P. E., Gurmu, S., Sumell, A. J., & Black, G. 2007. Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients. ***Economics of Innovation and New Technology***, 16(2): 71–99.

Tassey, G. 2000. Standardization in technology-based markets. ***Research Policy***, 29(4-5): 587–602.

Temple, P., Blind, K., Jungmittag, A., Spencer, C., & Witt, R. 2005. The Empirical Economics of Standards. ***DTI economics paper, No. 12, June 2005***.

Van Looy, B., Callaert, J., & Debackere, K. 2006. Publication and patent behavior of academic researchers: Conflicting, reinforcing or merely co-existing? ***Research Policy***, 35(4): 596–608.

Wang, X., Gorlitsky, R., & Almeida, J. S. 2005. From XML to RDF: how semantic web technologies will change the design of "omic" standards. ***Nature Biotechnology***, 23(9): 1099–1103.

Zi, A., & Blind, K. 2013. Researchers' participation in standardisation in relation to their scientific publishing and patenting activities: a case study from a public research organisation in Germany. Presented at the EU-SPRI Early Career Researcher Conference (ECRC), Madrid.