A Retrospective Analysis of the Italian Medicines Agency's Evaluation of Real-World Evidence in Innovativeness Assessments

HTA69

Authors. Civitelli, D[1].; Mills, M.[1,2]; Kanavos, P.G[1]

1. *Health Policy Department, LSE Health, London School of Economics and Political Science, London, UK*
2. *HTA-Hive, London, UK*

Background and Objectives

Agenzia Italiana del Farmaco (AIFA) is the organization responsible for regulatory compliance and pharmaceutical pricing and reimbursement in Italy. The latter involves reimbursement negotiations, price discounts, managed access agreements, and innovativeness status appraisals.

The recognition of innovative status provides some key advantages to manufacturers, namely access to a 1.3 billion euro a year Fund for Innovative Drugs and the exclusion from mandatory price cuts. Additionally, drugs that receive full or conditional innovative status benefit from immediate inclusion in regional formularies, leading to faster patient access.

Innovative status can only be conferred to drugs used for the treatment of severe or debilitating illnesses. The decision is based on three criteria, namely *unmet therapeutic need, added clinical benefit* and *quality of evidence [1]*.

Unmet therapeutic need is scored as:
- *Maximum*: there are no alternatives approved in the same indication.
- *Important*: some alternatives available, but with no impact on clinically relevant endpoints.
- *Moderate*: some alternatives available, but with a limited impact on clinically relevant outcomes and with an uncertain safety profile.
- *Poor*: several alternatives available, with a large positive impact on clinical outcomes and with a satisfactory safety profile.
- *Absent*: several disease-modifying alternatives with satisfactory safety profile are available.

Added clinical benefit is scored as:
- *Maximum*: greater efficacy demonstrated on clinically relevant outcomes compared to available alternatives. The drug can cure the disease or significantly alter its natural history.
- *Important*: greater efficacy on clinically relevant outcomes, or the ability to reduce the risk of disabling or potentially fatal complications, or a better risk/benefit ratio compared to available alternatives, or the ability to prevent high-risk clinical procedures. The drug modifies the natural history of the diseases in a sub-population of patients, or provides a clinically meaningful improvement in terms of quality of life or progression free survival compared to available therapies.
- *Moderate*: moderate efficacy gains in subpopulations, with limited improvements in quality of life.
- *Poor*: small efficacy gain or proven gain on outcomes that are not clinically relevant. Minor advantages, such mode of administration, compared to available therapies.
- *Absent*: no added clinical benefit compared to available therapeutic alternatives.

The *quality of evidence* is evaluated according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method. Study design, risk of bias, inconsistency and imprecision are considered. GRADE rates the quality of evidence as *high, moderate, low or very low [2]*.

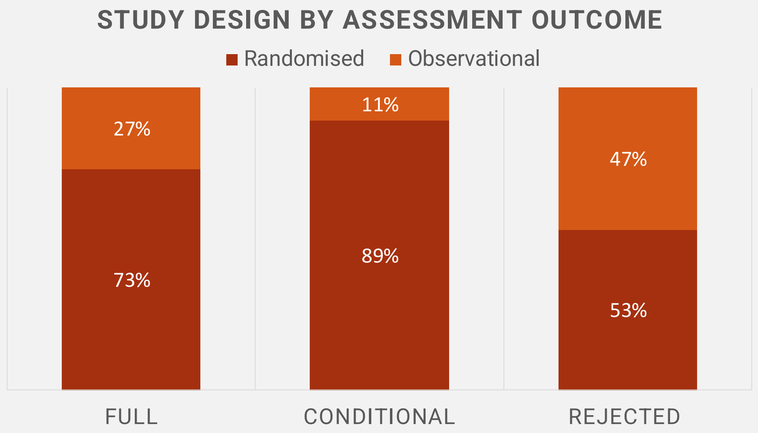
This study investigates whether the study design (randomised vs. observational) is correlated with different ratings of unmeet need, added clinical benefit, quality of evidence and overall Innovativeness Assessment outcomes.

Methodology

- Innovativeness Assessment reports published between 2020 and 2023 were downloaded from the AIFA website.
- The overall assessment outcome, unmet need, added clinical benefit, accuracy of the evidence, endpoint type and quality of evidence ratings were extracted from each report and coded in Excel.
- Descriptive statistics and chi-square tests of independence were carried out in STATA.

Results

- 110 reports were collected. 36% (n = 40) received full innovative status, 25% (n = 27) received conditional innovative status and 39% (n = 43) obtained a negative outcome.
- 73% (n = 29) of the drugs that were granted full innovative status were tested through a randomized control trial (RCT), and only 27% (n = 11) through an observational trial. 89% (n = 24) of the drugs that received conditional innovative status were evaluated through an RCT, whilst 11% (n = 3) through an observational trial. 53% (n= 23) of rejected drugs included an RCT in their submission, whereas 47% (n = 20) included an observational trial.



	N = 40	N = 27	N = 43
Randomised	29	24	23
Observational	11	3	20
			p-value = 0.006

- A large share of assessments that resulted in either full or conditional innovative status included RCTs (73% and 89%, respectively). Observational trials were much more commonly featured in rejected assessments. The p-value suggests that there is a statistically significant association between trial design and innovative status.
- *Unmet need* was always rated as *maximum, important or moderate*. RCTs were used in the majority of the submissions (approximately 70%) that received an important or moderate rating. However, no statistically significant correlation was detected (p = 0.142)
- *Quality of evidence* ratings differed significantly by trial design (p = 0.000). Specifically, only RCTs received a rating of *high* or *moderate*. Observational trials were exclusively rated as either low or very low. 59% (n = 19) and 83% (n = 15) of all trials that obtained a low and very low rating were observational, respectively.

	Randomised (%)	Observational (%)	p-value
Unmet need (Maximum)	1 (25%)	3 (75%)	
Unmet need (Important)	22 (73%)	8 (27%)	
Unmet need (Moderate)	53 (70%)	23 (30%)	p = 0.142
Unmet need (Low)	0	0	
Unmet need (Absent)	0	0	
Quality of evidence (High)	15 (100%)	0 (0%)	
Quality of evidence (Moderate)	44 (100%)	0 (0%)	
Quality of evidence (Low)	13 (41%)	19 (59%)	
Quality of evidence (Very Low)	3 (17%)	15 (83%)	**p = 0.000**
Surrogate endpoint	26 (34%)	50 (66%)	
Clinical endpoint	8 (24%)	26 (76%)	p = 0.263
Accuracy (High)	27 (100%)	0 (0%)	
Accuracy (Moderate)	36 (100%)	0 (0%)	
Accuracy (Low)	10 (33%)	20 (67%)	
Accuracy (Very Low)	3 (20%)	12 (80%)	
Accuracy (Non quantifiable)	0 (0%)	2 (100%)	**p = 0.000**
Added clinical benefit (Maximum)	0	0	
Added clinical benefit (Important)	24 (69%)	11 (31%)	
Added clinical benefit (Moderate)	37 (80%)	9 (20%)	
Added clinical benefit (Low)	8 (80%)	2 (20%)	
Added clinical benefit (Absent)	5 (71%)	2 (29%)	
Added clinical benefit (Non quant.)	2 (17%)	10 (83%)	**p = 0.001**

- The *accuracy* indicator, which refers to the level of certainty of the measured clinical benefit, follows a similar trend. Only RCTs receive scores of High or Moderate, whereas observational trials generate most Low and Very Low scores (67% and 80%, respectively). This correlation is statistically significant (p = 0.000).
- The share of drugs tested through RCTs and observational studies was roughly stable across added clinical benefit ratings. This ranged from 69% to 80% and 20% to 31% for RCTs and observational studies, respectively.

Conclusions

Assessments that received full or conditional innovative status feature randomized controlled trials more frequently. The use of observational studies was associated to a higher number of rejections. RCTs may be unfeasible or unethical for severe or rare diseases. Interestingly, there was no correlation between study design and unmet need ratings. AIFA scores observational trials significantly worse, both in terms of quality of the evidence and accuracy.

References
1. Italian Medicines Agency (AIFA). (n.d.). Innovative Medicines. Available at: https://www.aifa.gov.it/en/farmaci-innovativi (Accessed: 4 May 2025).
2. Fortinguerra, F., Tafuri, G., Trotta, F., and Addis, A. (2019). Using GRADE methodology to assess innovation of new medicinal products in Italy. British Journal of Clinical Pharmacology. Available at: https://bpspubs.onlinelibrary.wiley.com/doi/full/10.1111/bcp.14138 (Accessed: 4 May 2025).

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE