WORKING PAPERS

N° 1507

September 2025

"Economic Incentives to Develop and to Use Diagnostic Tests
A Literature Review"

David Bardey, Philippe De Donder and Vera Zaporozhets

Toulouse School of Economics

Economic Incentives to Develop and to Use Diagnostic Tests: A Literature Review

September 2025

David Bardey[a] Philippe De Donder[b] Vera Zaporozhets[c]

a Universidad de Los Andes, Bogotá, Colombia y Toulouse School of Economics, France. d.bardey@uniandes.edu.co
b Toulouse School of Economics, CNRS, University Toulouse Capitole, France. philippe.dedonder@tse-eu.fr
c Toulouse School of Economics, INRAe, University Toulouse Capitole, France. vera.zaporozhets@tse-eu.fr

ABSTRACT

This survey examines the economic literature on the incentives that shape both the use and the development of diagnostic tests, with a particular focus on companion (biomarker) tests central to precision medicine. Misdiagnosis, underdiagnosis, and overdiagnosis represent a substantial global burden, driving healthcare costs and adverse patient outcomes. The study synthesizes theoretical, empirical, and experimental evidence to assess how healthcare providers' decisions regarding diagnostic tests are influenced by payment schemes, altruism, and time constraints. Fee-for-service arrangements are shown to encourage excessive testing, while capitation and salary-based contracts help contain costs, though sometimes at the expense of quality. Physicians' non-monetary motivations, such as altruism and reputational concerns, interact with financial incentives in complex ways, occasionally leading to unintended consequences such as undertesting.

From a normative perspective, the literature highlights the trade-offs inherent in reimbursement design: mandating even costless diagnostic tests is not always optimal, and greater altruism does not necessarily enhance welfare. Current practices, such as reimbursing biomarker tests separately from associated treatments in the U.S., are criticized for discouraging their adoption. At the industry level, the survey explores incentives for developing innovative tests. Pre-approval companion tests can improve drug approval prospects and justify higher prices, whereas post-approval test development faces weaker incentives due to reduced market size. Competition among firms strengthens incentives relative to monopolistic settings, but test introduction may also dampen price competition.

The findings suggest that pay-for-performance schemes, procurement design, and value-based pricing can help better align private and social incentives for both test use and development. Overall, the survey underscores the importance of carefully designed reimbursement mechanisms and policy tools to promote the efficient integration of diagnostic innovations into healthcare systems.

KEYWORDS: Diagnostic tests, Healthcare systems, Incentives.
JEL Code: D86, H51 and I11.

1. Introduction

The burden of incorrect medical diagnoses is substantial, leading to significant harm, disability, and even death for millions of people worldwide. Estimates suggest that misdiagnosis, underdiagnosis, and overdiagnosis contribute significantly to healthcare costs and adverse patient outcomes. In the U.S. healthcare system, Shrank *et al.* (2019) estimate that such unnecessary expenditures account for approximately 25% of total healthcare spending. Globally, these errors are a major cause of preventable harm in healthcare. The direct financial burden of diagnostic errors is estimated to be 17.5% of total healthcare expenditure in OECD countries, or 1.8% of GDP, according to this institution. [1] This includes costs associated with additional treatment, prolonged hospital stays, and legal proceedings.

Understanding and addressing the economic drivers of diagnostic errors is crucial for improving patient outcomes and alleviating the financial burdens they impose. To shed light on these factors, this survey reviews the economic literature on the incentives shaping both the development and the use of diagnostic and prognostic tests. These tests encompass all procedures used to diagnose patients' conditions and/or determine the most suitable treatments.[2]

While we discuss diagnostic tests in general, we place particular emphasis on **companion tests**, which are diagnostic tools used alongside treatments to determine their compatibility with individual patients. These tests are particularly important in cancer care, where they are often referred to as biomarkers tests. Biomarker tests search for genes, proteins, and other substances (known as biomarkers or tumor markers) that provide insights into the specific type of cancer a patient has and the most effective treatments. As such, biomarker testing is the key component of precision (or personalized) medicine, the healthcare approach that tailors treatment to a patient's individual characteristics

Medical treatments are rarely safe and effective for everyone. In the case of cancer, for example, biomarker testing offers several advantages. First, it enhances treatment efficacy and minimizes adverse effects by identifying patients who are likely to respond to treatment and those who are not. Second, it enables better-informed medical decisions by helping health practitioners choose the most appropriate medication based on the patient's specific characteristics. Third, it helps prevent or limit unnecessary interventions and overtreatment by excluding patients deemed unlikely to benefit. According to D'Avó Luis and Seo (2021), the potential cost per life-year gained from biomarker-guided drugs (therapies that require biomarker testing before prescription) falls below the threshold value used in the literature to define a cost-effective intervention.

Companion tests are playing an increasingly important role, both in optimizing the use of existing treatments and in the approval of new ones. For example, a review by the European Medicines Agency reveals that approximately

[1] https://www.oecd.org/en/publications/the-economics-of-diagnostic-safety_fc61057a-en.html

[2] A prognostic test is a medical test used to predict the likely course or outcome of a disease or medical condition, essentially estimating a patient's future health trajectory. Unlike diagnostic tests, which determine if a patient currently has a disease, prognostic tests focus on predicting the potential progression, severity, or duration of an existing condition. Our survey concentrates on diagnostic tests for non-communicable diseases. The challenges associated with communicable diseases, such as those highlighted during the Covid-19 pandemic, are distinct: tests play a crucial role in infection control and public health surveillance, where rapid detection, large-scale deployment, and epidemiological externalities are central. In contrast, for non-communicable diseases, diagnostic tests mainly inform individual treatment decisions and patient management, with incentives shaped more by provider behavior, reimbursement rules, and test development costs.

half of the cancer drugs authorized between 2015 and 2018 required patients to undergo genetic testing before determining the appropriate treatment (Antoñanzas *et al*., 2019).

This survey examines the individual decisions of healthcare providers regarding the use of diagnostic tests, and later, the incentives for their development. As we will see, the use of diagnostic tests by healthcare providers is influenced by various factors, including their level of altruism, time constraint, and the payment schemes they operate under. For example, since diagnostic tests can affect doctors' time constraints —either positively or negatively—they also impact doctors' income. Depending on the prevailing payment scheme, this can, in turn, alter the trade-off between patient welfare and financial compensation. While updating their knowledge to effectively interpret diagnostic test results requires time and entails opportunity costs, once this investment is made, doctors can improve both the accuracy and speed of their diagnoses. This investment decision is shaped not only by the type of payment scheme offered by payers but also by doctors' intrinsic motivations such as altruism toward their patients.

Section 2 examines how responsive healthcare providers, primarily doctors, are to incentives, and to the various payment and reimbursement schemes, such as **fee-for-service** (payment per service), **capitation** (fixed amount per patient), **salary** (fixed wage), and **pay-for-performance** (payment tied to quality or outcomes). It reviews both theoretical and empirical studies, including several based on laboratory experiments, and concludes that payment schemes can effectively incentivize healthcare providers to modify their medical practices. Section 3 then focuses on the specific case of diagnostic tests, particularly companion tests. Fisher *et al.* (2003) and Brody (2010) highlight that a significant proportion of medical testing decisions are deemed inappropriate, leading to either overprovision or underprovision. While public attention often centers on overprovision, medical literature extensively documents cases of underprovision of diagnostic testing (see, for example, Newman-Toker *et al.* (2013), Singh *et al.* (2013), Zhi *et al.* (2013), O'Reilly (2014). Sollman (2015) estimates that the economic impact of undertesting could be as high as 38% of total healthcare expenditures.

The empirical studies reviewed here suggest that healthcare providers are driven by a mixture of monetary and altruistic motivations —for example, they tend to prescribe fewer tests when aware of their patients' out-of-pockets costs. Diagnostic tests are more frequently used when they are well-known and easy to interpret. Additionally, theoretical studies indicate that healthcare providers may exhibit overconfidence, relying too heavily on their own expertise and too little on diagnostic tests.

Sections 2 and 3 take a positive, or descriptive, approach examining how physicians respond to incentives, particularly in relation to the payments schemes they operate under. These sections help illuminate the trade-offs doctors face and the role of the payment schemes incentives in shaping their decisions. Section 4 then adopts a normative approach, exploring how health authorities can design optimal reimbursement schemes to achieve their objectives. Theoretical studies highlight that any reimbursement rule involves trade-offs between objectives; for instance, rewarding good health outcomes may lead to increased health expenditures. They also reveal counterintuitive findings, such as the idea that even costless diagnostic tests should not necessarily be made mandatory or that greater physician altruism does not always improve social welfare. Additionally, the U.S. practice of reimbursing biomarkers tests separately from their associated treatments has been criticized, as it is at least partially responsible for the limited use of biomarker tests in practice.

Finally, Section 5 takes a dynamic approach, reviewing the theoretical and empirical literature on the incentives for the industry to develop innovative diagnostic tests. A key distinction is made between tests developed alongside

treatments (companion tests) and those introduced after treatments are already available. Unsurprisingly, the outlook for the latter is often bleak, particularly when the reduction in market size caused by the test's introduction is not sufficiently offset by an increase in the (often regulated) price. Beyond higher pricing, proposed mechanisms to encourage the development of innovative tests include pay-for-performance schemes —where the reimbursement is tied to treatment success —and procurement design rules.

The academic literature we review employs three main methodologies: theoretical (the development and analysis of analytical models), empirical (the use of databases, primarily through regression analysis) and experimental (the design of laboratory experiments). For each study, we aim to clearly indicate the methodology employed, noting that some contributions combine multiple approaches—for example, developing an analytical model and complementing it with empirical or experimental analysis. We summarize the key findings at the end of each section and revisit them in the conclusion.

2. Healthcare providers and incentive schemes

In this section, we survey the main general results presented in the economic literature studying how healthcare providers' incentives, in particular doctors, are influenced by the payment schemes and reimbursement rules that they face. According to McGuire (2000), doctors can modify their medical practice through two types of behavior, both related to their time constraint. On the one hand, doctors may adjust the total time allotted to their medical practice in general, or to each patient on average. On the other hand, they can modify the volume of services provided to each patient or during each event.

Along the first dimension, Showalter and Thurston (1997) take advantage of a reform of the tax system to study physicians' labor supply in the USA. They find that self-employed physicians are sensitive to tax rates, suggesting that the physicians' labor supply depends on their income, and, consequently, on the payment scheme used. Batalgi *et al.* (2003) find similar results in the context of Norwegian hospitals. All in all, the total time worked by physicians seems to depend on the remuneration obtained. In other words, and as can be expected, physicians' labor supply shows a degree of elasticity.

As pointed out in Gosden *et al.* (1999), there seems to be a general agreement that fee-for-service (FFS henceforth) schemes favor supply-induced demand behaviors from healthcare providers. They show that these behaviors tend to generate inflation of heath care costs.[3] For instance, Brekke *et al.* (2017) study empirically how general practitioners (GP) respond to fee changes at the intensive margin. They use detailed administrative data covering all GPs in Norway during the period 2006–2011. Their results reveal that a higher consultation fee leads to more visits and lower treatment intensity. Fortin *et al.* (2021) obtain similar findings in Canada. Based on linked administrative and survey panel data, they study the labor supply behavior of physicians who could adopt either a standard fee-for-service contract or a mixed remuneration contract. Under the latter, physicians receive a *per diem* while the fee for services provided is reduced. These authors estimate a structural discrete choice model that incorporates service intensity (services provided per hour) and contract choice into a labor supply framework to control for the selection bias. Their results indicate that supply of services is reduced under a mixed payment contract. In particular, the number of hours spent seeing patients is less sensitive to incentives than the supply of services that include diagnostic tests.

[3] It is the case in a labor supply set-up including a time constraint, when the substitution effect dominates the income effect. In other words, when the FFS rate increases, doctors value more the time devoted to work and less the time devoted to leisure. See for instance Devlin and Sarma (2008).

By contrast, capitation payment (CAP hereafter) and salary are two payment schemes which seem to be relatively effective in controlling healthcare expenditures, sometimes at the cost of the quality or quantity of services provided to patients (Bardey *et al.*, 2012). This negative effect on health care quality is likely to be stronger with a salary than with a capitation payment. Indeed, with a capitation payment, a lower quality may generate more events per patient, and this increasing number of events may affect the number of patients who need to visit their doctors and then decrease doctors' income with this payment scheme. Moreover, if the demand is sufficiently sensitive to quality, *ceteris paribus*, physicians have more incentives to be cautious with the quality of services than when they receive a flat payment as a salary.

However, it is important to acknowledge that measuring the quality supplied is a complex issue since quality is not always observable, and consequently not always contractible. Moreover, even though healthcare providers may respond to financial incentives and their competitive environment, their decisions are also guided by intrinsic motivations and altruistic concerns. In such a context, as pointed out by Benabou and Tirole (2003), one must be cautious when introducing financial incentives since they may crowd out intrinsic motivation. Typically, these countervailing effects mean that remuneration schemes conveying explicit financial incentives may produce unintended quality outcomes.

Because it is generally difficult to account for healthcare providers' intrinsic motivation and actual quality of care in both theoretical models and empirical estimations, several studies address these issues through experiments designed to replicate physicians' trade-offs. Green (2014) designs an experiment to compare the outcome generated by several prominent physicians' payment schemes including fee-for-service, capitation, salary, and payment for performance (P4P henceforth). She shows that doctors' intrinsic motivations play a significant role in their decision-making. While retrospective payment schemes tend to 'crowd out' intrinsic motivations, under FFS or a blended scheme that combines FFS and P4P physicians provide a lower overall quality of services. Finally, the results of her experiments reveal that when physicians receive either a salary or a CAP they provide a higher overall quality of service.

This quality/quantity trade-off is analyzed by Lagarde and Blauuw (2021) in a real-effort experiment that replicates situations of multitasking environments where some of the outputs achieved are rewarded while others are not. More precisely, they design a laboratory experiment where they test the impact of physicians' financial incentives on quality and quantity outcomes according to the remuneration scheme used. In practice, the two activities are: a routine activity (medical data entry) and a cognitive activity that aims to capture the diagnosis elaboration. Subjects are randomly allocated to a control, gain or loss contract. Interestingly, the authors find that participants increase performance differently when patients are submitted to potential losses or when physicians are rewarded for the outcomes achieved. While patients' losses contribute to increasing participants' performance through a greater attention that reduces the number of mistakes, bonuses tend to increase the time spent on the rewarded activity. Contrary to the prediction obtained from theoretical models of multitasking, the authors do not observe externalities, either negative or positive on the non-incentivized activity (*i.e.*, the diagnosis task).

Green *et al.* (2017) find quite different results when studying other payment schemes often used in the healthcare sector. They design an experiment to focus on the difference between a flat rate and P4P on health outcomes. In line with Lagarde and Blauuw (2021), they find that compared to a flat rate, a payment for performance scheme increases the number of incentivized measures met, but this positive result comes at the expense of the quality of care through unintended effects on adherence to standards of care. Consistent with the general findings of Bénabou

and Tirole (2003), this suggests greater caution in implementing pay-for-performance schemes for physicians, as the negative outcome may be interpreted as evidence of a crowding-out effect.

Finally, Byambadalai *et al.* (2023) provide an interesting theoretical model in which doctors choose health care quality according to their altruism level and their competitive environment. In contrast to the rest of the literature, they abstract from the payments schemes' properties to rather focus on the consequences of the competitive environment. In other words, they study how altruistic preferences are modified by markets' incentives. In addition to their theoretical setting, they conduct a laboratory experiment using a within-subject design. Subjects are asked to choose health care qualities for hypothetical patients in different market structures, from monopoly to quadropoly. Prices, costs, and patients' benefits are experimental incentive parameters.

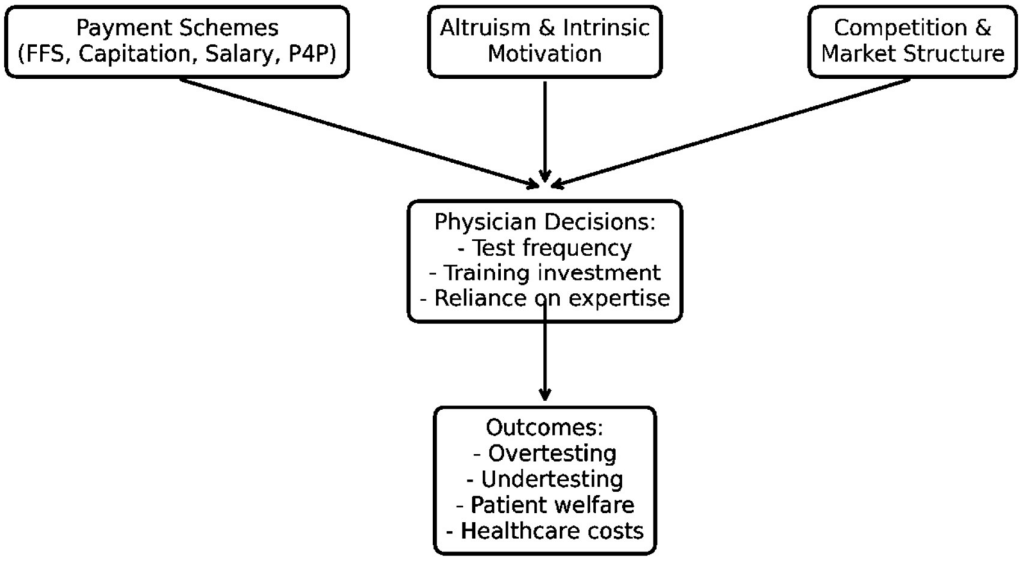
When healthcare services are provided by a physician in a monopoly position, the subjects choose quality by trading off profits and altruistic patients benefits. By contrast, when the experiment replicates the institutional setting of a competitive environment, due to the uncertainty toward their competitors' altruism, each subject competes for patients by choosing qualities. The authors compute the Bayes-Nash equilibrium that describes subjects' quality decisions as functions of altruism. Using a nonparametric method, they estimate the population altruism distributions from the quality observed in this Bayes-Nash equilibrium in different markets and incentive configurations. They conclude that competition tends to reduce altruism, although equilibrium quality levels under duopoly and quadrupoly are much higher than under monopoly. Although markets crowd out altruism, the disciplinary powers of market competition are stronger. Counterfactuals corroborate the hypothesis that physicians' preferences can change according to markets' competition degree.

Key Takeaways

Healthcare providers may respond to changes in monetary incentives either by adjusting the total time they devote to medical practice or by modifying the volume of services delivered per patient (see Figure 1). With respect to the former, empirical evidence from countries such as the United States and Norway shows that providers' overall labor supply is sensitive to monetary incentives. As for the latter, there is a broad consensus that fee-for-service (FFS) reimbursement rules tend to drive up healthcare costs, whereas capitation (CAP) and salary-based payments help contain costs —though sometimes at the expense of quality.

Theoretical research underscores the importance of intrinsic motivations, such as altruism, and their interaction with monetary incentives—which can sometimes generate crowding-out effects that unintentionally affect quality. Laboratory experiments confirm the presence of these effects.

Figure 1. Incentives, Physician Behavior, and Outcomes



We now turn to the literature that focuses more specifically on the use of diagnostic tests.

3. Physicians' incentives to use diagnostic test

A key distinction in this section is between ambulatory care and healthcare delivered in hospitals. In most healthcare systems, diagnostic tests performed in ambulatory settings are reimbursed by health insurance. Patients may face some out-of-pockets costs according to the generosity of their health insurance coverage, but physicians' decisions regarding diagnostic tests are not of a financial nature, as tests prices do not affect their income, at least directly, *i.e.*, through the payment scheme. Physicians' trade-offs in the use of diagnostic tests depend on: i) their medical practice and the expected patients' outcomes, ii) non-medical incentives such as profit or revenue incentives, medical liability fears, and patient demand (which are often cited as reasons for deviations from the clinical guidelines).[4] The first component of this trade-off reflects physicians' altruism and the value they place on patient welfare, which depends critically on the diagnostic gain in precision relative to the welfare cost of the test for the patient. Physicians' non-medical incentives are more complex: they encompass not only their own welfare—such as income and workload—but also aspects of their patients' non-medical welfare.

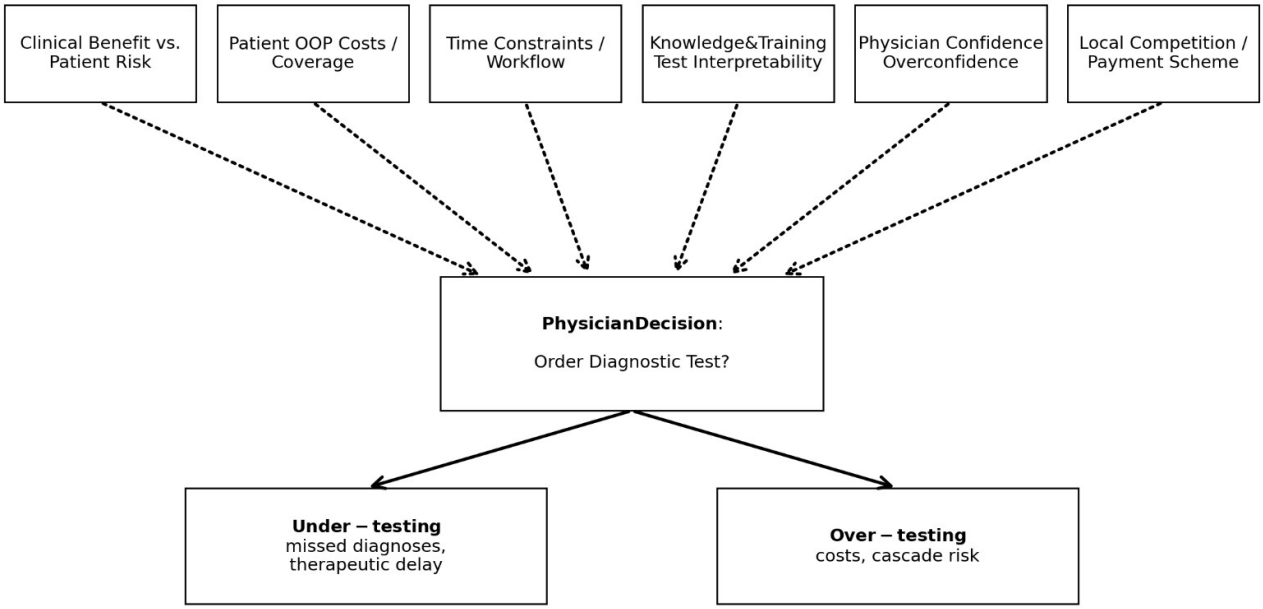
For instance, Tierney *et al.* (1990) study how physicians' decisions may vary according to the information they have about their patients' out-of-pockets diagnostic tests costs. In their study, 121 physicians were divided into two groups. Those who belonged to the intervention group knew the out-of-of-pocket amount paid by patients, while those in the control group did not have access to this information. While during the 14 weeks preceding the experiment there were no differences between the physicians of the two groups, during the 26-weeks intervention period, the authors find that physicians of the treatment group ordered 14 percent fewer diagnostic tests. The authors conclude that making patients' out-of-pocket costs visible influences physicians' incentives to prescribe diagnostic tests. This finding suggests that adjusting copayment levels for diagnostic tests, and ensuring that this

[4] See for instance Lee and Levy (2012) and Smith-Bindman *et al.* (2008).

information is visible to physicians, may help align prescribing behavior with the optimal number of tests.

As illustrated by Byambadalai *et al.* (2023), the relative importance of the various types of non-medical incentives depends on the physicians' institutional environment. For ambulatory care, the density of doctors affects the competition intensity which, in turn, alters physicians' incentives to induce their patients' demand (Léonard *et al*., 2009), with diagnostic tests being one way to implement this demand induction strategy. However, using tests beyond guidelines' recommendations not only increases the burden of health expenditures, but can also be detrimental to patients through the diagnostic therapeutic cascade (Deyo, 2002).[5]

Figure 2. Drivers of Diagnostic Testing



At first glance, the trade-off physicians face when deciding whether to prescribe diagnostic tests resembles that encountered with personalized medicine tools. In both cases, financial incentives and costs largely depend on the prevailing payment and reimbursement schemes. However, an important difference arises: most conventional diagnostic tests are straightforward for physicians to interpret, as they are well integrated into general medical practice and embedded in clinical guidelines. In contrast, few physicians feel comfortable interpreting the complex information provided by genetic or molecular tests. Thus, differences may emerge not only according to the type of test used, but also with respect to physicians' background and the time required to interpret the results. A key distinction, therefore, is that personalized medicine tools often require physicians to invest in professional training to handle the diagnostic information they generate—an investment that can be viewed as a sunk cost, which subsequently enables them to reduce the variable costs of treating patients with these tools (see Bardey *et al.*, 2021).

The fee-for-service (FFS) payment scheme is commonly associated with supply-induced demand (see Bardey and Lesur, 2006), a concern that also extends to diagnostic testing. Yip and Hsiao (2009), in a preliminary evaluation of

[5] This expression is mostly used for cardiovascular troubles. It corresponds to the tight linkage between the diagnostic tests and the therapeutic intervention. See for instance Lucas *et al.* (2008).

the Chinese health system, argue that inappropriate incentives under China's FFS model have contributed to rapidly rising costs alongside low-quality services. A key manifestation of these inefficiencies has been the excessive use of diagnostic tests, which the authors attribute primarily to the FFS structure. While they acknowledge that altering the payment system would not, by itself, resolve all inefficiencies, they contend that the current FFS scheme is largely responsible for the overuse of diagnostic testing. They also draw on international experiences to suggest reforms that are better suited to the Chinese context.

Anaya et al. (2016) argue that the dynamics may differ for certain innovative diagnostic tests. In particular, they note that such tests can often help physicians save time in establishing a diagnosis, which in some cases reduces the number of patient visits and, under an FFS scheme, leads to lower income. To encourage more effective use of personalized medicine tools, the authors suggest adopting alternative payment schemes, although they do not specify which ones.

Building on this premise, Bardey et al. (2021) combine a theoretical model with an experiment to examine how different payment schemes influence providers' incentives to adopt personalized medicine tests. They find that pay-for-performance (P4P) schemes create stronger incentives than fee-for-service (FFS) or capitation to encourage their use. Notably, after controlling for potential selection bias, their experimental results suggest that physicians who choose to invest in acquiring personalized medicine tools exhibit a form of commitment-device behavior: by dedicating time to update their skills, they become more devoted to their patients. Based on these findings, the authors argue that P4P or capitation payments are more conducive than FFS to promoting personalized medicine, and they further recommend partially subsidizing physicians' investment costs in these tools.

Physicians' willingness to use diagnostic tests depends not only on their ability to interpret the information these tests provide but also on their capacity to establish a diagnosis without them. To explore this issue, Dai and Singh (2020) analyze a setting in which a partially altruistic physician must diagnose a patient. Physicians can rely solely on their own diagnostic ability or supplement it with a perfect diagnostic test (i.e., one that reveals the patient's true condition). They are aware of whether their diagnostic ability is high or low, which constitutes private information and gives rise to adverse selection. In addition to altruism toward their patients, physicians also care about their reputational payoff, which depends on how peers perceive their diagnostic skill. Physicians then decide whether to use a test whose cost is borne by patients. The authors show the existence of a unique separating equilibrium: high-ability physicians rely only on their own judgment, whereas low-ability physicians use the test. This generates an inefficiency, as high-ability physicians may forego tests even when their use would be optimal. Interestingly, the authors find that greater altruism can exacerbate this underuse of diagnostic tests by high-ability physicians. They also note that financial incentives—such as concerns about malpractice liability—may further reinforce undertesting in equilibrium.

Healthcare providers' decisions may differ in the hospital setting, as the properties of payment schemes can vary from those in ambulatory care. Similar to outpatient medicine, both retrospective reimbursement and fee-for-service schemes tend to encourage therapeutic cascades, thereby contributing to rising healthcare expenditures. However, as Allen (2015) notes, alternative payment schemes designed to better align incentives between payers and providers can also be implemented to contain costs. These schemes will be discussed in the next section.

Key Takeaways

We summarize the main drivers of diagnostic testing decisions and the resulting outcomes in Figure 2. The general

findings outlined in Section 2 also apply to diagnostic tests. Empirical evidence from China, for instance, shows that the FFS reimbursement rule leads to an increase in the use of diagnostic tests. Other empirical studies suggest that healthcare providers balance both monetary and altruistic motivations —prescribing fewer tests when aware of their patients' out-of-pockets costs, for instance. The use of diagnostic tests is more common when they are well-known and easy to interpret. Laboratory experiments indicate that the time and effort invested in learning to use certain tests (such as personalized medicine tests) serve as a commitment device, leading to increased utilization. Finally, theoretical research suggests that healthcare providers may exhibit overconfidence, relying too heavily on their expertise while relying too little on diagnostic tests.

4. Reimbursement rule and payment schemes: diagnostic test and normative approach

This section complements the previous one by adopting a normative perspective, reviewing studies that examine the optimal reimbursement schemes for healthcare providers to induce socially optimal testing behavior.

Ghamat *et al.* (2018) study an issue similar to Dai and Singh (2020) (see the end of the previous section) in a slightly different set-up and with a more normative approach. In particular, the authors examine performance-based payment contracts to promote the optimal use of an optional diagnostic test for newly diagnosed cancer patients. As in Dai and Singh (2020), they model the interaction between two parties—a healthcare payer and a physician who also benefits from a private information. In Dai and Singh (2020), adverse selection arises from physicians' private information about their own ability, whereas in Ghamat *et al*. (2018) it stems from private information about patients' characteristics. Beyond this, Dai and Singh also highlight that physicians' diagnostic effort is non-contractible, thereby creating scope for moral hazard. Due to this information asymmetry, the authors show that it is not optimal to incentivize physicians to use a diagnostic test for all patients, even if the test is costless. The intuition is that universal testing would require higher payments to satisfy physicians' participation constraints. As a result, the contractual cost makes compulsory diagnostic testing an inefficient policy. Interestingly, the authors show that physicians are not always able to take advantage of their private information. As in Dai and Singh (2020), social welfare is not always increasing in the physician's degree of altruism.

Carroni et al. (2023) revisit the physician–patient agency problem within the game-theoretical framework of persuasion models. They study the case of a patient who experiences symptoms but is uncertain about being ill and consults a physician remunerated under a fee-for-service (FFS) scheme. In this setting, demand inducement arises when the physician persuades a healthy patient that she is ill. The patient incurs a health loss in two situations: when receiving unnecessary treatment despite being healthy, and when remaining untreated despite being ill.

The authors assume that sick patients are heterogeneous with respect to their willingness to receive treatment. Patients with low values of this parameter are reluctant to accept treatment, whereas those with high values are, ceteris paribus, more eager to do so. The physician observes each patient's parameter and strategically selects the type of diagnostic test to order—specifically, its precision, which determines the likelihood of type I and type II errors. A central feature of the model is that the physician has discretion both in recommending a test and in choosing its accuracy. Because patients ultimately decide whether to follow the physician's treatment recommendation, the diagnostic test serves as a signaling device to convey information and persuade them. In particular, patients with intermediate levels of willingness to undergo treatment typically require the additional

information provided by a diagnostic test before consenting to treatment.

Building on this framework, the authors examine the consequences of different regulatory policies. They first consider regulations that impose minimum standards on test sensitivity and specificity (i.e., the rates of false negatives and false positives, respectively). Their results show that limiting false negatives has no effect, since physicians already have incentives to recommend tests with low false-negative rates. By contrast, imposing limits on false positives is welfare-enhancing, as it curbs physician-induced demand and thereby reduces unnecessary treatments.

The authors also examine whether mandating the use of diagnostic tests improves welfare. Their findings are mixed and depend on the cost of the test. They then analyze a regulatory intervention that alters physicians' financial incentives, revealing countervailing effects: higher reimbursement per patient reduces the number of untreated patients in need of care, but also increases incidence of unnecessary and potentially harmful treatments. Since diagnostic tests provide information to patients, the authors find that the beneficial effect dominates, suggesting that adjusting physicians' financial incentives can enhance social welfare.

Brandt and Cassou (2023) develop a framework in which a social planner contracts with a profit-maximizing hospital to decentralize the provision of diagnostic tests and treatments. Patients present with primary symptoms of varying severity, which healthcare providers observe as a costless signal before deciding whether to administer an imperfect diagnostic test. If a test is conducted, providers update their beliefs in a Bayesian manner, and both the test results and subsequent treatments are assumed to be verifiable and contractible. However, when no test is performed, the social planner cannot observe patients' primary symptoms, leaving providers with private information that they can use to extract rents.

Brandt and Cassou (2023) analyze prospective payment schemes, which are commonly used in hospitals. They derive optimal contracts within this class, interpreted as Pathway-Related-Group (PRG) payments—that is, transfers defined for each possible situation, combining patients' primary symptoms with any diagnostic test results. They show that prospective payments based on average costs are not incentive-compatible and may lead to excessive testing or overtreatment. Their findings further indicate that incentive compatibility in a PRG system requires cross-subsidies. Alternatively, compatibility can be achieved by rewarding good health outcomes, though this comes at the expense of higher healthcare expenditures. Ultimately, the authors argue that, given the complexity of implementing incentive-compatible PRG payments with cross-subsidies, retrospective cost reimbursement may be more practical in real-world settings.

Mougeot and Naegelen (2022) address this issue by examining the allocative–efficiency trade-off in the presence of the shadow cost of public funds. They analyze a setting with two actors: a laboratory that produces either one or two drugs (one of which is associated with a companion test), and a Health Authority responsible for regulation. A population of patients is diagnosed by a physician who acts as their perfect agent. While the standard treatment is effective for some patients, another fraction of the population responds only partially to it. In the absence of the companion test, the physician cannot identify which patients will benefit fully from the drug.

The authors analyze a pricing policy that implements the second-best allocation, balancing allocative efficiency

against distributional effects in the presence of a shadow cost of public funds. [6] They first consider the case with a single drug, followed by the case with two drugs and a companion test. Their results show that optimal prices are higher when personalized medicine is prescribed—that is, when the new treatment's effectiveness exceeds the average effectiveness of the standard treatment. However, they also demonstrate that personalized medicine is welfare-enhancing only if the cost of the companion test is sufficiently low.

In a brief qualitative article, Allen (2015) notes that most U.S. payers reimburse biomarker tests and their associated treatments separately. He argues that this arrangement is suboptimal, as it results in relatively few biomarker tests being prescribed. Allen further suggests that the growing adoption of bundled payments represents a step in the right direction, since it transfers some financial risk from payers to providers, who then become residual claimants of the expenditure reductions enabled by diagnostic tests.

Allen (2015) highlights two main obstacles to the widespread use of biomarker tests: first, many tests identify biomarkers implicated in a patient's disease for which no targeted drug yet exists; and second, even when personalized treatments are available, biomarkers often fail to determine which option is most appropriate for a given patient.

Finally, we present case studies discussed in the OECD report (2025) that examines the properties of different reimbursement schemes. Across OECD health systems, the report emphasizes that fee-for-service (FFS) rewards outputs rather than outcomes and is associated with the use of low-value diagnostics. To address this, it recommends shifting toward bundled payments covering cycles of care and outcomes, as these can alter practice patterns and reduce low-value care. Australia's experience with inappropriate coronary angiography illustrates how an output-driven funding model can simultaneously foster overdiagnosis in some groups and underdiagnosis in others, while offering little accountability for decisions along the diagnostic pathway or for broader population health objectives. By contrast, Latvia combines capitation and FFS with pay-for-performance payments to general practitioners, tied to cancer detection stage and screening coverage, thereby introducing explicit incentives for earlier and guideline-concordant diagnoses. Taken together, these cases illustrate a continuum from volume-based reimbursement to pathway- and outcome-oriented models, complemented by targeted P4P mechanisms that better align diagnostic practices with public health objectives. These insights are summarized in the following table.

[6] The concept of shadow cost of public funds is used in public economics/regulation theories to capture, usually in partial equilibrium environments, the idea that taxes introduce distortions and inefficiencies (Laffont and Tirole, 1993). Roughly speaking, the collection of 1$ generates a social cost of X, dubbed the shadow cost of public funds.

Table 1: Incentives and Diagnostic Test Use

Payment Model	Incentive Target	Effect on Diagnostic Use	Risk / Accountability
Fee-for-Service (FFS)	Individual services / tests (volume)	Higher test volume; higher risk of low-value or duplicate diagnostics	Over-diagnosis (some groups); weak accountability
Bundled Payments	Episode or cycle of care (pathway + outcomes)	Encourages guideline-concordant, pathway-aware testing	Lower cascade risk; shared accountability for outcomes
Pay-for-Performance (P4P)	Quality Key Performance Indicators (e.g., staging, screening)	Boosts early detection and screening coverage	Depends on Key Performance Indicators design; complements base payment

Key Takeaways

The theoretical contributions reviewed in this section mainly address adverse selection, in which healthcare providers hold private information—either about their own ability or about patients' health status—that is not observable to other stakeholders. Some models also incorporate moral hazard, in which providers take hidden actions, such as exerting diagnostic effort. Collectively, these studies underscore the trade-offs inherent in reimbursement design—for example, rewarding good health outcomes may come at the cost of higher healthcare expenditures. They also offer counterintuitive insights, such as that even costless diagnostic tests should not always be mandatory, and that greater provider altruism does not necessarily lead to higher social welfare.

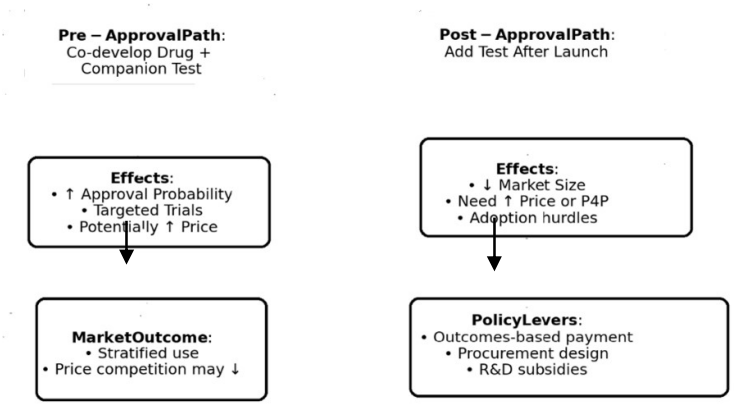
Beyond optimal reimbursement rules, this literature also examines the regulation of test characteristics—such as minimum specificity and sensitivity when chosen by healthcare providers—as well as the design of optimal contracts between hospitals and health authorities. Finally, the U.S. practice of reimbursing biomarker tests separately from their associated treatments has been criticized and is considered at least partly responsible for their limited use in practice.

Thus far, we have assumed the existence of a diagnostic test and examined the positive and normative properties of different payment schemes. In the next section, we take a more dynamic perspective, analyzing both the desirability of and incentives for developing new diagnostic tests.

5. Dynamic incentives and diagnostic tests

The decision to associate a diagnostic test with a drug can occur at two points of the drug's clinical development: either before/during clinical development (pre-approval case), or after marketing authorization (post-approval case). In the post-approval case, when the drug is already on the market, incentives for additional R&D are limited. However, in the pre-approval case, biomarker tests may increase the likelihood of drug approval. The most commonly cited example is the case of metastatic melanoma, for which the treatment vemurafenib (Zelboraf) was developed alongside the COBAS BRAF V600E test and received simultaneous FDA marketing approval. This marked the fastest FDA approval in history.

Figure 3. Companion Diagnostic Development: Pre- and Post-Approval



From the perspective of a pharmaceutical firm, introducing a companion diagnostic test post-approval reduces the number of eligible patients and, consequently, lowers revenue *per* drug. As a result, if drug prices do not increase to reflect testing, gross drug sales will drop. Given these considerations, there has been growing interest in combining drugs with companion diagnostic biomarker-based tests, both to increase the probability of approval of the bundle drug-companion test and to raise the price of this bundle.

Many cancer drugs were initially launched without an accompanying biomarker test. Regulatory authorities such as the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have actively encouraged the use of biomarker tests in the development and application of prescription drugs.[7] In this context, Antoñanzas *et al.* (2019) report that, in the three years preceding their study, approximately half of the cancer drugs authorized by the EMA required patients to undergo genetic testing before treatment decisions were made. Gromova *et al.* (2020) estimate that around 65% of drugs approved by the EMA and FDA between 2015 and 2019 were associated with at least one biomarker in their development program.[8] However, pharmacogenomic tests are not yet widely available (*e.g.*, Alcenat *et al.*, 2021). In addition to their costs, one potential reason for their limited use is the complexity and imprecision of biomarker predictions, which contributes to delays in physicians updating their medical knowledge and practice.

There are also situations in which, without diagnostic tests to identify treatment responders, a drug lacks sufficient value to payers. For example, Nebacumab, a treatment for sepsis, was shown not to be cost-effective in the absence of a companion test and was therefore withdrawn from the market (Danzon and Towse, 2002). Similarly, some off-patent drugs that have been replaced by newer generations may experience a "second life" through biomarker tests, if such tests identify patients for whom the older drugs are more effective than the newer alternatives. In other

[7] See Gromova *et al.* (2020). Note that the different regulatory requirements can also create challenges when it comes to the development of therapeutics and companion diagnostic tests. For instance, in the US, marketing approval for drugs and diagnostics is performed by a single agency, the FDA. However, this is not the case in the EU. The European Medicine Agency regulates the marketing approval for drugs, while each EU member state Notified Body monitors the performance standards of diagnostic tests.

[8] Gromova *et al.* (2020) point out that immunosuppressants, immunostimulants, drugs used in diabetes, antithrombotic drugs, antineoplastic agents and antivirals are the medical specialities which have developed most drugs that include one or several biomarkers.

words, biomarkers tests not only facilitate the development of new and more efficient treatments, but they can also improve the matching with all available treatments, including older drugs that may be particularly effective for patients with specific biomarkers.[9] In this context, it is worth pointing out that the use of biomarker tests is not always associated with costly medical treatments.

Health authorities are typically concerned with several aspects. In the absence of diagnostic tests, little or no patient stratification occurs, increasing the risk of side effects among non-responders. Conversely, when tests are available, responders benefit from improved health outcomes, but costs may increase due to higher drug prices and smaller patient populations. A salient feature of biomarker tests is their tendency to define small patient niches, often reclassified as rare conditions (with prevalence below 40 per 100,000), for which new orphan drugs are marketed at nearly unaffordable prices (Graf von der Schulenburg and Frank, 2015). This trend is expected to continue with the expansion of personalized medicine and the development of new diagnostic tests (Antoñanzas et al., 2019).

In a theoretical model, Brekke *et al.* (2024) study the incentives for pharmaceutical companies to develop biomarker tests that perfectly reveal a drug's suitability for a patient, and the impact of such tests on the equilibrium market allocations in a setting with two pharmaceutical firms and a single health plan. In their model, the two drug producers first choose whether to develop a test and then set the price of the drug they produce. Observing these decisions, the insurance plan decides which (if any) of the two drugs to include in its health plan and let physicians decide which (if any) drug to prescribe to each patient. The two drugs differ both vertically and horizontally, with each patient represented by a location in a two-dimensional Hotelling space. This information is not known to anyone, unless a biomarker test is developed. By solving this model, the authors obtain a rich set of insights. In case of a monopoly, the drug maker has an incentive to develop a test only if the market is not served otherwise--*i.e.*, if the quality of the treatment is low. In such cases, too few patients are treated due to the firm's pricing decision. It is important to note that the monopoly has no incentive to develop a test when its drug is anyway prescribed in the absence of testing, because the reduction in the market size caused by the test is not sufficiently offset by the higher price (even though the test increases the drug's value for those who receive it).

Competition drastically changes incentives and outcomes, as firms have more incentives to develop a test than they would under a monopoly. A low-quality drug producer has an incentive to develop a test, since it would otherwise be excluded from the market. This, in turn, induces the other firm to develop its own test. While developing a biomarker test for the low-quality drug improves welfare, this is not always the case for the high-quality drug, because the presence of tests can dampen competitive effects during price-setting. An extension studies the case in which the market is not fully covered when two tests are developed and shows how the results are affected as a function of the drug qualities.

In a similar vein, Antoñanzas *et al.* (2018) study how to incentivize pharmaceutical companies to develop a biomarker test for a drug already marketed without such a test. This scenario corresponds to the monopoly case studied by Brekke *et al.* (2024). They assume that the drug's price is exogenous and cannot be changed after the development of the biomarker test.[10] In this case, health authorities need an alternative instrument to incentivize the development of the test, since itwould reduce the market size of the prescribed drug. They show how health

[9] See Bardey *et al.* (2016) for a drugs' entry model where the drugs' price regulation must consider that more drugs available allow to reduce their adverse effect thanks to a better match.

[10] They claim that this the most common real-world situation and provide the example of the drug panitumumab for patients with metastatic colorectal cancer.

authorities may use pay-for-performance instruments in this context. In the absence of a test, the authorities should fully penalize the drug maker when the treatment fails for example, by requiring the firm to reimburse the entire price paid for the drug. Decreasing the size of this penalty when a test is used will then incentivize the development of such a test, even if this test is imperfect (reducing, but not eliminating, the fraction of patients who fail to respond fully to the drug). They study how this optimal penalty rate is affected by exogenous factors, such as the drug's health value, the severity of its side-effects or the cost of monitoring the fraction of well-treated patients.

Addressing the same issue, a companion article by Antoñanzas *et al.* (2019) analyzes the decision-making process of health authorities and pharmaceutical firms in the context of treatment personalization. While in their companion article they assume a fixed drug price, in this study they consider that the health authority sets the drugs prices to maximize net health benefits.[11] Laboratories take the drugs prices as given when deciding whether to invest in a test that identifies patients who will respond to the drugs. The authors conduct comparative statics to characterize the equilibrium outcomes based on the price level and the drug's response rate. They show that the decision to develop a test depends more on the drugs' response rate than on the price level. When the proportion of responders is relatively high, health authorities do not incentivize the firms to search for biomarkers, since the drug is generally effective. When the response rate is low, the drug is not adopted. Personalized medicine is most likely to occur at intermediate levels of the proportion of responders.

These authors also compare the two prices set by the Health Authority based on the firm's R&D decision. They point out that when the firm invests in R&D to stratify the patients' population, the price set by the Health Authority is not necessarily higher than the price when no such stratification occurs. More specifically, this comparison depends on the test price, the severity of adverse effects, the cost of producing the drug and the test precision. They show that when the test cost is relatively low, the price set by the health authority is higher when the firm invests in R&D than when the treatment is administered to all patients. Conversely, when the test is more expensive, the price set by the health authority is higher without R&D expenses. Finally, for intermediate test costs, the relationship between the drug price and the test price mainly depends on the treatment effectiveness.

Many researchers highlight the need for flexible and value-based pricing to reflect higher benefits of targeted treatment and to encourage pharmaceutical firms to develop drugs with associated biomarker tests (*e.g*., Danzon and Towse (2002) and Garrison and Austin (2007)). Another possibility is to support research through R&D subsidies (Hsu and Schwartz, 2008). Using a theoretical model calibrated for several diseases, Danzon and Towse (2002) conclude that testing is often socially optimal, particularly when the proportion of non-responders is high, when serious adverse reactions are possible, or when the test is inexpensive.

Alcenat *et al*. (2021) study laboratories' incentives to increase their drugs' effectiveness in a moral hazard framework (*i.e.,* where the effort undertaken by the laboratory to increase its effectiveness is not observable by a health authority). They analyze the drug reimbursement contract of a laboratory producing a new treatment that is associated with a genetic/biomarker test. In their model, the health authority can recommend either a standard treatment, or the use of a genetic/biomarker test to prescribe the most suitable treatment to each patient based on the test result. Their model reveals that the moral hazard informational structure impacts the optimal contract designed by the health authority when one of the two treatments dominates in the absence of genetic tests, provided the price of the new treatment is below a certain threshold. In these cases, moral hazard reduces the frequency of personalized medicine implementation compared to when effort is observable. This is since, *ceteris paribus*, the

[11] The authors consider a kind of payment for performance in the sense that the price is paid only when patients are well-treated.

laboratory cannot fully internalize the benefits of its effort. In contrast, when the new treatment is preferred without genetic information and its price exceeds the threshold, moral hazard does not impact the implementation of personalized medicine. However, the authors make two restrictive assumptions: they assume that the companion test has zero cost, and their effort variable only impacts the probability that the personalized treatment is preferred over the standard treatment, rather than considering that the benefit of personalized treatment increases with effort.

While the two previously discussed articles provide valuable insights into policies supporting the optimal use of personalized medicine in the presence of companion tests, neither addresses competition between laboratories, unlike Brekke *et al.* (2024). In their settings, a single laboratory supplies both standard and personalized treatments and decides whether to develop a companion test. In practice, however, many biomarker tests are developed by independent entities (Allen, 2015). Scott-Morton and Seabright (2013) highlight the pharmaceutical industry's limited incentives to develop companion tests alongside innovative drugs. As with many R&D issues, this stems from the divergence between the social and private value of biomarkers. Specifically, biomarkers generate social value by reducing unnecessary demand and avoiding costly expenditures associated with ineffective treatments, yet this demand-reduction effect lowers firms' private returns. Furthermore, laboratories may hold private information on drug efficacy across subgroups and withhold it, even when disclosure would substantially benefit patients and providers. Scott-Morton and Seabright argue that well-designed procurement mechanisms and price regulation could help bridge this gap by better aligning private and social incentives.

Key Takeaways

The incentives to develop a (companion) diagnostic test are weaker once the associated treatment has already been approved, as the resulting reduction in market share must be offset by a sufficient increase in the treatment price. Consequently, developing a diagnostic test can enhance the likelihood of the companion drug receiving approval from health authorities and, in some cases, increase its price. Competition among laboratories strengthens the incentives to develop companion tests compared to a monopoly setting, and these incentives are also greater when the average efficiency of the treatment is low. However, the development of companion diagnostic tests may reduce price competition between laboratories. If the treatment price remains fixed after the introduction of a companion test, pay-for-performance schemes —where the reimbursement depends on treatment success — are necessary to encourage test development. Alternatively, procurement design and price regulation can help align the private and social values generated by biomarkers, considering the private information laboratories hold about specific patient groups that would benefit from their new tests.

6. Conclusion

This survey examines both the financial incentives for developing innovative diagnostic tests and the decisions healthcare providers make regarding the use of existing tests. To address the latter, we begin by reviewing the literature on providers' responsiveness to incentives. Theoretical, empirical, and experimental evidence consistently shows that healthcare providers are sensitive to monetary incentives, particularly in relation to their labor supply. At the same time, this literature underscores the importance of non-monetary motivations, such as altruism, and warns that monetary incentives may crowd out these intrinsic values.

There is broad agreement that fee-for-service (FFS) payment schemes tend to increase healthcare costs, while capitation (CAP) and salary-based schemes are more effective at containing costs, albeit sometimes at the expense of quality. The case of diagnostic tests, however, is more nuanced. Some empirical studies do find

that FFS arrangements lead to excessive testing—as observed in China—but the evidence reviewed in the Introduction suggests that underutilization is often the more pressing concern. Diagnostic tests can save time by reducing the need for additional visits and therapeutic interventions, yet FFS schemes may still discourage their use.

Healthcare providers' mixed motivations also play a crucial role. Experimental studies show, for example, that altruistic considerations can lead physicians to prescribe fewer tests when they are aware of their patients' out-of-pocket costs. The care setting (ambulatory versus hospital) is also important, since different rules for cost sharing typically apply. Diagnostic tests are more commonly used when they are familiar and easy to interpret. Laboratory experiments further reveal that the sunk investment of time and effort required to adopt certain tests—such as personalized medicine tools—acts as a commitment device, encouraging their continued use. Finally, theoretical studies suggest that providers may sometimes display overconfidence, relying too heavily on their own expertise and not enough on diagnostic testing.

Designing optimal incentive schemes requires a clear understanding of the objectives and constraints faced by healthcare providers. The theoretical models reviewed here typically incorporate some degree of adverse selection, whereby providers hold private information, and in some cases moral hazard, as their actions are not fully observable. These models generally indicate that mandating the use of diagnostic tests is not optimal, even when such tests are costless. They also criticize the U.S. practice of reimbursing biomarker tests separately from their associated treatments, which contributes to their underuse. Finally, several contributions examine the regulation of diagnostic test characteristics, such as imposing minimum standards for specificity and sensitivity.

In the development of companion tests, incentives differ markedly before and after the approval of the associated treatment, with stronger incentives in the pre-approval phase. Before approval, a companion test increases the likelihood of drug approval by targeting use to patients with the highest probability of responding. After approval, however, developing such a test narrows the market size, often requiring substantial price increases to preserve overall profitability. The literature shows that this post-approval challenge is particularly acute for monopolists but less pronounced under competition between innovators. Incentives to develop tests are also stronger when the average effectiveness of the treatment is low. At the same time, test adoption can reduce price competition among innovators, resulting in higher equilibrium prices. Companion tests may also enhance the value of older, off-patent drugs, as these are less likely to be tied to costly treatments. Finally, policies such as pay-for-performance schemes, procurement design, and price regulation can be leveraged to encourage the development of companion tests.

References

Alcenat, S., Marechal F. and Naegelen, F. (2021). “Implementation of Personalized Medicine in a Context of Moral Hazard and Uncertainty About Treatment Efficacy,” *International Journal of Health Economics and Management* 21(1): 81–97. https://doi.org/10.1007/s10754-020-09290-2

Allen, T. (2015). “Payment for Cancer Biomarker Testing,” *Arch Pathol Lab Med* 39(3):300-4. https://doi.org/10.5858/arpa.2014-0190-ED

Anaya J-M, Bardey, D., Duarte-Rey, C., Sarmiento-Monroy, J-C, Castiblanco, J. and Rojas-Villarraga, A. (2016). “Personalized Medicine. Closing the Gap Between Knowledge and Clinical Practice,” *Autoimmunity Reviews* 15(8): 833-842. 10.1016/j.autrev.2016.06.005

Antoñanzas, F., Rodríguez-Ibeas, R., and C. Juárez-Castelló, (2018). “Personalized Medicine and Pay for Performance: Should Pharmaceutical Firms be Fully Penalized When Treatment Fails?”, *Pharmacoeconomics* 36: 733-743. 10.1186/s13561-019-0244-8

Antoñanzas, F., Rodríguez-Ibeas, R., and Juárez-Castelló, C. (2019). “Pre-approval Incentives to Promote Adoption of Personalized Medicine: A Theoretical Approach,” *Health Economics Review* 9:28, 2-10. https://doi.org/10.1186/s13561-019-0244-8

Bardey, D., Cremer, H. and Lozachmeur, J-M. (2012). “Doctors' Remuneration Schemes and Hospital Competition in a Two-Sided Market,” *The B.E. Journal of Economics and Analysis & Policy* 1(13). https://doi.org/10.1515/1935-1682.3254

Bardey D. and Lesur, R. (2006). “Optimal Regulation of Health System with Induced Demand and Ex Post Moral Hazard", *Annals of Economics and Statistics* 83/84, 279-293. https://www.jstor.org/stable/20079171

Bardey D., Lozachmeur, J-M. and Jullien, B. (2016). "Health Insurance and Diversity of Treatment: A Policy Mix Perspective", *Journal of Health Economics* 47, 40-53, 2016. https://doi.org/10.1016/j.jhealeco.2016.01.003

Bardey, D, Kembou Nzalé, S. and Ventelou, B. (2021). “Physicians’ Incentives to Adopt Personalised Medicine: Experimental Evidence,” *Journal of Economic Behavior and Organization* 191, 686–713. https://doi.org/10.1016/j.jebo.2021.09.022

Batalgi, B., Bratberg, B, and Holmas, T. (2003). “A Panel Data Study of Physicians' Labor Supply: The Case of Norway,” Bergen University *Working Paper*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=388623

Benabou, R. and Tirole, J. (2003). “Intrinsic and Extrinsic Motivation”, *The Review of Economic Studies* 70(3): 489-520. https://doi.org/10.1111/1467-937X.00253

Brandt NM, Cassou M. (2024). “Care Protocols Adherence and Contracts: Managing the Diagnostic-Therapeutic Cascade,” Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4774356

Brekke K., Holmås H., Monstad K. and Straume, O. (2017). "Do Treatment Decisions Depend on Physicians' Financial Incentives?," *Journal of Public Economics* 155(C): 74-92. https://doi.org/10.1016/j.jpubeco.2017.09.012

Brekke, K., Dalen D.M. and Straume, O. (2024). “Competing with Precision: Incentives for Developing Predictive Biomarker Tests,” *Scandinavian Journal of Economics* 126(1): 60–97. https://doi.org/10.1111/sjoe.12543

Brody, H. (2010). “Medicine’s Ethical Responsibility for Healthcare Reform—The Top Five List,” *New England J. Medicine* 362(4):283–285. https://doi.org/10.1056/NEJMp0911423

Byambadalai, U., Ma, A. and Wiesen, D. (2023). “Changing Preferences: An Experiment and Estimation of Market-Incentive Effects on Altruism,” *Journal of Health Economics* 92:102808. https://doi.org/10.1016/j.jhealeco.2023.102808

Carroni E, Pignataro G. and Siciliani, L. (2023). “Persuasion In Physician Agency,” *Discussion Papers in Economics*, University of York. Available at: https://ideas.repec.org/p/yor/yorken/23-01.html

Dai, T. and Singh, S. (2020). “Conspicuous by Its Absence: Diagnostic Expert Testing Under Uncertainty,” *Marketing Science* 39(3): 540–563. https://doi.org/10.1287/mksc.2019.1201

Danzon P. and Towse, A. (2002). “The Economics of Gene Therapy and of Pharmacogenetics,” *Value in Health* 5(1): 5-13. https://doi.org/10.1046/j.1524-4733.2002.51081.x

D′Avó Luís, A.B. and Seo, M.K. (2021). “Has the development of cancer biomarkers to guide treatment improved health outcomes?”, *The European Journal of Health Economics* 22:789–810. https://doi.org/10.1007/s10198-021-01290-4

Devlin, R.A. and Sarma, S. (2008). “Do Physician Remuneration Schemes Matter? The Case of Canadian Family Physicians,” *Journal of health economics 27*(5): 1168-1181. https://doi.org/10.1016/j.jhealeco.2008.05.006

Deyo, R. (2002). “Cascade Effects of Medical Technology,” *Annual Review of Public Health* 23:23-44. https://doi.org/10.1146/annurev.publhealth.23.092101.134534

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL and Pinder, EL. (2003). “The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care”, *Ann. Internal Medicine* 138(4):273–287. https://doi.org/10.7326/0003-4819-138-4-200302180-00006

Fortin, B., Jacquemet, N., and Shearer, B. (2021). “Labour Supply, Service Intensity, and Contracts: Theory and

Evidence on Physicians," *Journal of Applied Econometrics* 36(6): 686-702. https://doi.org/10.1002/jae.2840

Ghamat, S., Zaric, G. and Pun, H. (2018). "Contracts to Promote Optimal Use of Optional Diagnostic Tests in Cancer Treatment," *Production and Operations Management* 27(12): 2184–2200. https://doi.org/10.1111/poms.12780

Garrison, L. and Austin, M. (2007). "The Economics of Personalized Medicine: A Model of Incentives for Value Creation and Capture," *Drug Information Journal* 41: 501-509. https://doi.org/10.1177/009286150704100408

Gosden, T., Pedersen, L., and Torgerson, D. (1999). "How Should We Pay Doctors? A Systematic Review of Salary Payments and Their Effect on Doctor Behaviour," *QJM: An International Journal of Medicine* 92:47–55. https://doi.org/10.1093/qjmed/92.1.47

Graf von der Schulenburg, J. and Frank, M. (2015). "Rare is Frequent and Frequent is Costly: Rare Diseases as a Challenge for Health Care Systems," *European Journal of Health Economics* 16(2):113-8. https://doi.org/10.1007/s10198-014-0639-8

Green, E. (2014). "Payment Systems in the Healthcare Industry: An Experimental Study of Physician Incentives," *Journal of Economic Behavior & Organization* 106: 367–378. https://doi.org/10.1016/j.jebo.2014.05.009

Green, E., Peterson, K. Markiewicz, K., O'Brienc, J and Arring, N. (2017). "The Impact of a Merit-Based Incentive Payment System on Quality of Healthcare: A Framed Field Experiment," *Working Paper*. http://dx.doi.org/10.2139/ssrn.3066257

Gromova, M., Vaggelas, A., Dallmann, G. and Seimetz, D. (2020). "Biomarkers: Opportunities and Challenges for Drug Development in the Current Regulatory Landscape," *Biomark Insights* 15: 1-13. https://doi.org/10.1177/1177271920974652

Hsu, J. and Schwartz, E. (2008). "A Model of R&D Valuation and the Design of Research Incentives," *Insurance: Mathematics and Economics* 43(3): 350-367. https://doi.org/10.1016/j.insmatheco.2008.05.003

Laffont, J-J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*. MIT Press.

Lagarde, M. and Blauuw, D. (2021). "Effects of Incentive Framing on Performance and Effort: Evidence from a Medically Framed Experiment," *Journal of the Economic Science Association* 7: 33–48. https://doi.org/10.1007/s40881-021-00100-0

Lee, D. and Levy, F. (2012). "The Sharp Slowdown in Growth of Medical Imaging: An Early Analysis Suggests Combination of Policies Was the Cause," *Health Affairs* 31(8), 1876–1884. https://doi.org/10.1377/hlthaff.2011.1034

Léonard, C., Stordeur, S. and Roberfroid, D. (2009). "Association between Physician Density and Health Care Consumption: A Systematic Review of the Evidence," *Health Policy* 91(2): 121-134. https://doi.org/10.1016/j.healthpol.2008.11.013

Lucas, F.L., Siewers, A.E., Malenka, D.J. and Wennberg, D.E. (2008). "Diagnostic-Therapeutic Cascade Revisited: Coronary Angiography, Coronary Artery Bypass Graft Surgery, and Percutaneous Coronary Intervention in the Modern Era," *Circulation 118*(25): 2797-2802. https://doi.org/10.1161/CIRCULATIONAHA.108.789446

McGuire, T. (2000). Physician Agency, In *Handbook of Health Economics* (Vol. 1, Chap. 9, pp. 461–536). https://doi.org/10.1016/S1574-0064(00)80168-7

Mougeot M and Naegelen, F. (2022). "Incentives to Implement Personalized Medicine Under Second-Best Pricing," *Health Economics* 31: 2411-2424. https://doi.org/10.1002/hec.4588

Newman-Toker DE, McDonald KM, Meltzer, DO. (2013). "How Much Diagnostic Safety Can We Afford, and How Should We Decide? A Health Economics Perspective," *BMJ Quality Safety* 22(S2): ii11–ii20. https://doi.org/10.1136/bmjqs-2012-001616

O'Reilly, KB. (2014). "Sin of Omissions: When Tests Fly Under the Radar," *CAP Today* 28(5): 60–66. Available at: https://www.captodayonline.com/sin-of-omissions-when-tests-fly-under-the-radar-2/

Scott-Morton, F and Seabright, P. (2013). "Research into Biomarkers: How Does Drug Procurement Affect the Design of Clinical Trials?", *Health Management, Policy and Innovation* 1(3): 1-15. Available at: https://www.tse-fr.eu/articles/research-biomarkers-how-does-drug-procurement-affect-design-clinical-trials

Showalter M. H. and Thurston, N. K. (1997). "Taxes and Labor Supply of High-Income Physicians," *Journal of Public Economics* 66(1): 73–97. https://doi.org/10.1016/S0047-2727(97)00021-2

Shrank, W., Rogstad, T. and Parekh, N. (2019). "Waste in the US Health Care System: Estimated Costs and Potential for Savings," *JAMA* 322(15):1501-1509. https://doi.org/10.1001/jama.2019.13978

Singh H, Giardina TD, Meyer AND, Forjuoh SN, Reis MD, Thomas EJ. (2013). "Types and Origins of Diagnostic Errors in Primary Care Settings," *JAMA Internal Medicine* 173(6):418–425. https://doi.org/10.1001/jamainternmed.2013.2777

Smith-Bindman, R., Miglioretti, D. and Larson, E. (2008). "Rising Use of Diagnostic Medical Imaging in a Large Integrated Health System," *Health Affair* 27(6): 1491–1502. https://doi.org/10.1377/hlthaff.27.6.1491

Sollman, P. (2015), “Achieving Value: Community Outreach and Other Strategies for Labs,” *Medical Laboratory Observer* 47(12): 22–24. Available at: https://openurl.ebsco.com/EPDB%3Agcd%3A9%3A6188521/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A111219016&crl=c&link_origin=scholar.google.com

Tierney, W., Miller, M. and McDonald, C. (1990). “The Effect on Test Ordering of Informing Physicians of the Charges for Outpatient Diagnostic Tests,” *The New England Journal of Medicine* 322: 1499-1504. https://doi.org/10.1056/NEJM199005243222105

Yip, W. and Hsao, W. (2009). “China's Health Care Reform: A Tentative Assessment,” *China Economic Review* 20(4): 613–619. https://doi.org/10.1016/j.chieco.2009.08.003

Zhi M, Ding EL, Theisen-Toupal J, Whelan J, and Arnaout R. (2013). “The Landscape of Inappropriate Laboratory Testing: A 15-year Meta-Analysis,” *PLoS One* 8(11): e78962. https://doi.org/10.1371/journal.pone.0078962