RESEARCH ARTICLE

Analysis of diagnostic product portfolios using the Portfolio-To-Impact modelling tool [version 1; peer review: 2 approved with reservations]

Maël Redard-Jacot ¹, Devy M. Emperador ¹, Eva Junyent¹, Mickey Urdea², Rich Thayer², Rangarajan Sampath ¹

¹Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland

²Halteres Associates, LLC, San Ramon, CA, USA

V1 First published: 16 Feb 2021, 10:116
<https://doi.org/10.12688/f1000research.29057.1>
 Latest published: 16 Feb 2021, 10:116
<https://doi.org/10.12688/f1000research.29057.1>

Abstract

Background: The Portfolio-To-Impact version 2 (P2I v.2) financial forecasting tool estimates funding requirements for development of portfolios of candidate health products (drugs, biologics, vaccines or diagnostics). The assumptions and archetypes relating to diagnostics in P2I v.2 are based on limited data and may not accurately describe research and development costs, timelines and probability of success. This study aimed to revise the P2I v.2 tool by modifying the diagnostic assumptions to improve accuracy of predictions for diagnostic portfolios.

Methods: Data from expert interviews and historical information on development of 26 existing diagnostics were used to determine approximate research and development costs, timelines and probability of success for development of diagnostics, and to revise diagnostic archetypes and development phases. To compare the revised tool with P2I v.2, data on 27 candidates from the Foundation for Innovative New Diagnostics (FIND) tuberculosis and pandemic preparedness portfolios were input into both versions.

Results: The number of diagnostic archetypes increased from two in P2I v.2 to three in the revised tool. Total estimated costs to move the 27 candidates along the pipeline to launch were US\$641.62 million with P2I v.2 and US\$274.00 million with the revised model. The number of expected launches was 21.65 over five years with P2I v.2 and 11.48 over eight years with the revised model. Development timelines were extended and probability of success was lower with the revised model compared with P2I v.2.

Conclusions: Outputs from the revised tool were in line with expert experience, suggesting that the proposed revisions improve the accuracy of the tool for estimating research and development costs, timelines and probability of success relating to diagnostic portfolios. Additional improvements to the tool could include further refinement

Open Peer Review

Approval Status ? ?

	1	2
version 1
16 Feb 2021	view	view
1. Alec Morton , University of Strathclyde, Rottenrow, UK		
2. Célia Lemaire , Université de Strasbourg, Strasbourg, France		

Any reports and responses or comments on the article can be found at the end of the article.

diagnostics, forecasting, archetypes, costs, probability of success, portfolio, P2I

First published: 16 Feb 2021, 10:116 <https://doi.org/10.12688/f1000research.29057.1>

Introduction

The Portfolio-To-Impact (P2I) model is a financial forecasting tool developed by the Special Programme for Research and Training in Tropical Diseases (TDR)¹. The model enables users to estimate funding requirements for development of a portfolio of candidate health products, from late stage preclinical studies to phase III clinical trials, and estimates the number of potential product launches over time, as well as the probability of success of health products. The P2I v.2 model allows a range of different types of medical products to be assessed, including drugs, biologics, vaccines, and diagnostics. These are described as ‘archetypes’ and have further subdivisions within the model depending on the complexity of the product^{1,2}. The model is intended to be continually refined and improved, and users are encouraged to provide inputs to update key parameters and assumptions. In line with this, a second iteration of the model (P2I v.2) that includes additional archetypes and has modified cost assumptions was subsequently developed and used to estimate research and development costs associated with a pipeline of candidates for 35 neglected diseases². The model has also been used to assess the vaccine portfolio of the European Vaccine Initiative (EVI), which includes vaccine candidates for various diseases of poverty and emerging infectious diseases³.

The development pathway for diagnostics differs considerably from that of pharmaceuticals and vaccines, as they do not follow the traditional pathway from preclinical to phase III studies. Rather, after initial concept and research and feasibility planning, diagnostics move from the design and development phase directly to clinical validation, approval and launch^{4,5}. However, there is a lack of academic literature available on development characteristics for diagnostics, thus the archetypes and assumptions used within the P2I v.2 model were based solely on interviews with diagnostics developers and other diagnostic experts. Based on the experience of the Foundation for Innovative New Diagnostics (FIND), a product development partnership (PDP) focusing on diagnostics for diseases of poverty⁶, the diagnostic assumptions in the P2I v.2 model tend to underestimate the research and development costs and timelines

for diagnostic discovery and early stage development. Additionally, due to the complex nature of the diagnostic landscape, the archetypes and assumptions in P2I v.2 may not accurately describe all diagnostics. A range of factors can impact diagnostic research and development costs, timelines and probability of success, including the underlying technology of an assay, the maturity of the manufacturing company and their expertise and experience (in clinical, regulatory, quality management and infectious disease areas, as well as in assay development), the disease area, ease of obtaining samples for validation, and the Stringent Regulatory Authority (SRA) being pursued.

FIND has a large diagnostics pipeline across six disease areas (antimicrobial resistance, hepatitis C and human immunodeficiency virus [HIV], malaria and fever, neglected tropical diseases, tuberculosis [TB] and pandemic preparedness), and plans to use the P2I v.2 model as a complementary tool to existing portfolio management methods for prospective scenario planning, to inform funding decisions, refine technology assessments, and predict product launches and estimated research and development costs. The objective of this study was to propose revisions to the P2I v.2 diagnostic archetypes and assumptions, to improve accuracy of predictions for diagnostic portfolios, and to compare outputs of the revised model with the original P2I v.2 tool.

Methods

Revision of the P2I v.2 model

This work was carried out between April and December 2019. Data to inform the refinement of the P2I v.2 tool were collected through two methods: a) expert interviews and b) collection of historical data on development of existing diagnostics. The expert interviews were performed by the FIND project team. Seven experts with experience in diagnostic development in both high-income countries (HICs) and in low- and middle-income countries (LMICs) were asked if they were willing to be interviewed; all experts were either employees of FIND, consultants, or FIND external partners (Table 1). We obtained consent orally before the beginning of the interviews. Interviewees were provided with background information

Table 1. Expert interviewees and diagnostic tests supported for development.

Expert	Diagnostic	Archetype
Employee	Laboratory-based molecular platform	Simple technical platform development
	Assay development for menu expansion	Assay development
Employee	Point of care immunoassay for Ebola, HCV, HIV	Assay development
Employee	Point of care immunoassay for TB	Assay development
Employee	Point of care immunoassay for malaria	Assay development
Consultant	Point of care molecular assay to detect TB	Simple technical platform development
Consultant	Point of care molecular assay to detect TB	Simple technical platform development
External expert	Point of care molecular platform and assays to detect TB, HCV	Simple technical platform development

HCV, hepatitis C virus; HIV, human immunodeficiency virus; TB, tuberculosis.

about the P2I v.2 tool, and were asked to give feedback on the tool's diagnostic assumptions in relation to at least one product that they had supported through development. Data ranges for research and development costs, development timelines, and probability of success per product development phase were collected.

The historical data for existing diagnostics (ideally commercially available tests) were retrieved from the database housed by Halteres Associates (San Ramon, CA, USA), a consulting company with extensive experience in diagnostic development. A set of 26 diagnostic products for which sufficient data were available were selected from a broad group of product types and product development stages (Table 2). Research and development costs were approximate, and were based either on information from the companies themselves or from persons with knowledge of the companies; additionally, a number of the products were not yet marketed, thus final numbers for research and development costs and timelines were not available. As such, ranges for research and development costs and timelines were used except where otherwise stated.

Diagnostic archetypes were determined by identifying clusters among the 26 products with similar estimates for research and development costs, timelines and probability of success. This was done by calculating a composite 'level of risk' score that comprised factors thought to have most impact on these three assumptions. Initially, the following factors were included: target product profile, biomarker status, assay technology, instrumentation, per assay development costs, system development costs, regulatory requirements, preclinical/clinical studies, team competence, financial strength, payment/reimbursement model, product line extension, voice of customer, partner options, market development and commercialization plan/budget, entity stability, number of tests needed, and market demand/development. From this list, 15 factors were extracted that were believed to have the most impact (resulting in the removal of entity stability, number of tests needed, and market demand/development from the list), and the list was then further reduced for practicality of scoring, to obtain a set of eight criteria. Each of the 26 products was then assessed using the eight criteria, scored on a scale of 0 to 3 (Table 3). Approximate groupings were then determined based on the level of risk score.

Comparison of P2I v.2 and revised models

In order to compare the two models, data on the products from two of FIND's portfolios, TB and pandemic preparedness, were input into the P2I v.2 model, using the original diagnostic assumptions and archetypes (Table 4 and Table 5) and using the revised diagnostic assumptions and archetypes collected from the expert interviews and the historical database of 26 existing diagnostics. The research and development cost and timelines to move the candidates from the two FIND portfolios through the pipeline, and to estimate the likely number of launches per disease, was estimated. The model assumed a 2019 start, with all diagnostic candidates inputted based on their development phase at the time of review.

Potential impact

It can be challenging to compare diagnostic opportunities when they have comparable levels of risk. In these cases, the potential public health impact of a diagnostic can be a valuable differentiating factor. Factors that can influence the potential public health impact of a diagnostic product were identified from the historical database of 26 diagnostic products. Ten factors were identified; these were: increase in number of patients/year receiving appropriate treatment, cost per result, improvement in guideline adherence, time to intervention, improved access to critical data, support for diagnostic, surveillance or outbreak programmes (as intended), impact metrics, change in disease incidence or prevalence, reduction in loss to follow-up, and change in morbidity or mortality. The list was reduced to seven factors with the most influence, and a potential impact score for ranking was developed, scored on a scale of 0 to 3 (Table 6). Each of the 26 products from the historical database was scored according to this system.

Results

Refining the P2I v.2 model

Diagnostic archetypes. To determine the optimal diagnostic archetypes, clusters of diagnostic products that depicted similar estimates for research and development cost, timelines and probability of success were identified from the historical dataset of 26 diagnostic products (Figure 1). Approximate groupings of products that aligned with three archetypes emerged:

1. "New or existing entity, discovery" archetype: a research or commercial stage entity in the discovery phase of a new technology
2. "Existing entity, new product/technology" archetype: an established commercial company that is creating a new product that is not a product line extension
3. "Existing entity, product extension" archetype: an established commercial company that is creating a new product as an extension of an existing product line

Diagnostic assumptions. The expert interviewees proposed three development phases for diagnostics: "discovery", "design and development" and "clinical validation and launch readiness". These were fitted to the phases in the P2I v.2 model² by encompassing both "concept and research" and "feasibility and planning" from the original P2I v.2 tool into the "discovery" phase. The remaining two development phases were the same as those in the P2I v.2 tool ("design and development" and "clinical validation and launch readiness"). A fourth phase, "commercialization", was recommended by the interviewees, as this can be costly for diagnostics. Data for this phase are presented, but it was not included in the comparison analysis with the P2I v.2 tool.

Data from the expert interviews and the historical database of 26 existing diagnostics were used to identify the revised research and development cost, timeline and probability of success assumptions per product phase. Where the database

Table 2. Data for 26 diagnostic products used to inform revised P2I model assumptions.

Diagnostic test description	Discovery		Design and development		Clinical validation and launch readiness		Commercialization (up to 3 years) cost (US\$ million)
	Alpha trial cost (US\$ million)	Alpha trial time (years)	Beta trial cost (US\$ million)	Beta trial time (years)	Cost (US\$ million)	Time (years)	
IA hs-Tni	5–10 ^a	2–4	5–10 ^a	2–5 ^a	2–5	2–3	2–5
hs-HCV Ag	1–2	1–2	1–2	1–2	2–5	2–3	1–2
Lat TB	2–5	2–3	1–2	1–2	No data	No data	2–5
CMOS HIV VL	5–10 ^b	2–3	2–5 ^c	1–2	5–10	2–3	2–5
RDT HIV-ST 1	1–2	2–3	2–3	1–2	1–2	2–3	0.5–1
NAT PC	5–10	3–5	5–10	2–5	10–25	2–3	5–10
HCV VL	1–2	1	2–5	1–2	5–10	2–3	2–5
IA hs-TSH	20–50 ^d	3–5	2–5 ^e	1–2	2–5 ^e	1–2	5–10
RDT CAA	0.5–1	1–2	1–2	1–2	1–2 ^f	1–2	0.5–1
RDT Ov16	0.5–1	1–2	1–2	1–2	1–2 ^f	1–2	0.5–1
PCR BC Recurrence	150	3–5	10–25	2–5	10–25	1–2 ^g	10–25
Central IA RA	10–25	3–5	5–10	1–2	10–25	1–2 ^g	5–10
NAT PCR HAI	10–25	8–12	5–10 ^a	2–5	10–25	3–5 ^h	5–10
Cell CD4	2–5	3–5	5–10 ^a	1–2	2–5	1–2	0.5–1
RDT HIV-ST 2	0.5–1	1–2	2–3	1–2	1–2	2–3	0.5–1
NAT HPV	0.5–1	0.5–1	6.5	2–5	29	2–3	5–10
NAT HIV VL 1	0.5–1	0.5–1	4.5	2–5	9	2–3	5–10
NAT TB AMR	10–25 ^d	3–5	2–3	1–2	5–10	1–2	0.5–1
IA CrAg	0.5–1	1–2	2–3	1–2	2–5	1–2	2–5
NAT HIV VL 2	10–25 ^d	3–5	5–10	2–5	5–10	2–3	2–5
TB LAM	2	1	1	0.5	0.1	0.5–1	0.1
Cell CBC	10–25	8–12	5–10 ^a	2–5	10–25	3–5 ^h	5–10
RDT HIV	0.5–1	1–2	1–2	1–2	1–2	1–2	2–5
NAT TB	2–5	1–2	5–10		2–5	2–3	5–10
NAT Malaria	2–5	3–5	2–5	1–2	5–10	2–3	2–5
IA Malaria	2–5	3–5	5–10	2–3	5–10	2–3	2–5

^aIncluding instrument; ^bincluding manufacturing system; ^cwith new organization; ^dincluding instrument and disposable or cartridge; ^eassay only; ^fno regulatory; ^gCLIA LDT; ^hFDA. Products are not ascribed to a specific company due to confidentiality agreements. Ag, antigen; AMR, antimicrobial resistance; BC, breast cancer; CAA, circulating anodic antigen; CBC, complete blood count; CLIA, Clinical Laboratory Improvement Amendments; CMOS, complementary metal-oxide semiconductor; CrAg, Cryptococcal Antigen; FDA, United States Food and Drug Administration; HAI, haemagglutination inhibition assay; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; hs, high-sensitivity; IA, immunoassay; LAM, lipoarabinomannan; Lat, latent; PC, pancreatic cancer; RA, rheumatoid arthritis; RDT, rapid diagnostic test; ST, self-test; TB, tuberculosis; Tni, troponin I; TSH, thyroid stimulating hormone; VL, viral load.

Table 3. Level of risk criteria scoring and definitions.

Criterion	Score	Definition	Comments
TPP	0	The TPP is missing or incomplete	A complete, properly prepared and vetted TPP is essential to ensure the unmet needs and intended use for the product are properly understood and that the specified product requirements will meet the unmet needs
	1	A TPP for the product in development has been drafted but not properly vetted with experts and/or customers	
	2	A properly vetted TPP exists to guide development	
Assay technology	0	Product technology is in research phase; new technology is still under development	Assay technology includes all reagents and consumables. Development of an entirely new assay technology carries significant inherent risk until the evidence is generated supporting implementation of the proposed complete assay and consumable, including approval by a major regulatory agency (probably for a different application, but the same target molecule type, e.g. nucleic acid, antigen)
	1	Product technology is in early development phase; feasibility has been shown for new technology	
	2	Product technology has been approved by a major regulatory authority; this project is for a new application of the technology	
Per assay development costs	0	Estimated development costs through preclinical studies >US\$10 million	These costs are the estimated development costs per assay. They do not include costs for basic assay technology or instrument system development (if either is required). Costs should include all reagent and consumable-related costs necessary to perform the assay
	1	Estimated development costs through preclinical studies US\$2–10 million	
	2	Estimated development costs through preclinical studies <US\$2 million	
System development costs	0	Estimated development costs through clinical studies >US\$50 million	These costs are for the instrument system, including all necessary pre-analytical, analytical and reporting steps. It includes all hardware, software, reagents, consumables and information and communications technologies
	1	Estimated development costs through clinical studies US\$5–50 million	
	2	Estimated development costs through clinical studies <US\$5 million or no system	
Preclinical and clinical studies	0	The clinical studies are complicated. For example: sample/panel sources unknown or not available; and/or few/no appropriate study sites are known; and/or number of samples and/or sites must be large to obtain statistically significant data; and/or sites are widely geographically spread	Identification and procurement of necessary samples and panels, identification of appropriate study sites, number of study sites, location of sites and overall study length necessary to obtain statistically meaningful data are significant contributors to preclinical and clinical study risk, timeline and cost
	1	Study design, implementation and/or hurdles exist but reasonable mitigations have been identified; samples are available or reasonably simple to obtain	
	2	No identified hurdles to study design, implementation or execution exist; sufficient samples exist and are available	

Criterion	Score	Definition	Comments
Financial strength	0	The entity lacks a sustainable financial model	A key indicator of sustainable success is the ability to maintain sufficient gross and operating margins to support continued product/market and commercial development programmes. A financial plan alone is generally not sufficient if the entity has not also prepared an associated strategic operating plan or business model (ideally both)
	1	The entity has a financial model pertinent to the product in development but has not completed a strategic operating plan or business plan that should lead to a sustainable entity	
	2	The entity has both a financial model pertinent to the product in development and a strategic operating plan or business model that should lead to sustainability	
Voice of customer	0	The entity has no expertise with the intended customer base and/or in the required (initial) market channels; they have limited understanding of use cases or work flows	Access to internal or external resources to properly inform the TPP requirements, the market development plans, the competition, the key stakeholders and the customer access and support strategies is essential for success
	1	The entity has limited expertise with the intended customer and most of the stakeholders in the target ecosystem, e.g. the required market channels, the payers or funders, the competitors and other key market influencers in the appropriate countries. The entity understands the needs for use cases and work flows, but may not have completed the analyses	
	2	The entity has prior experience with the target customer base and has sufficiently defined all of the key stakeholders in the ecosystem. The have reasonably sufficient internal expertise to be successful in navigating the key interactions and deliverables	
Market development and commercialization plan and budget	0	The entity has been primarily focused on technology development and appears to lack sufficient understanding of, plans for and/or sufficient budget for market development and commercialization activities	Once technology and product development activities have been successfully navigated, many companies face a second chasm related to market development and commercialization. Often these activities are too challenging, specialized and/or resource demanding for an individual entity to manage on their own. Recognizing the necessary activities required for market development and sustainable commercialization takes knowledge, planning and networking to be successful. An entity that scores high here will have thought through these challenges
	1	The entity has prepared an initial plan and budget for market development and commercialization activities, but significant gaps remain and/or budget is insufficient	
	2	The entity has developed a comprehensive set of market development and commercialization plans with appropriate budget	

TPP, target product profile.

Table 4. Diagnostic archetypes from the P2I v.2 model.

Archetype	Description from P2I v.2 model	Examples	Further description
Assay development	Development of a diagnostic assay	Lateral flow tests, qualitative/quantitative molecular tests	Any new diagnostic that represents menu extension on an existing platform with an assay targeting a neglected disease.
Simple technical platform development	Development of a technical platform that enhances current technology	Ultrasensitive malaria rapid diagnostic test	Any new diagnostic that relies on a novel approach to sample processing or target detection.

This table is adapted from references 1 and 2 under a CC-BY 4.0 license.

Table 5. Diagnostic assumptions from the P2I v.2 model.

		Concept and research	Feasibility and planning	Design and development	Clinical validation and launch readiness	TOTAL
Cost per phase (US\$ million)	Assay development	1.5	1.5	2.0	3.5	8.5
	Simple technical platform development	1.5	1.5	100.0	3.5	106.5
Length of phase (years)	Assay development	0.5	0.5	2.0	1.3	4.3
	Simple technical platform development	0.5	0.5	2.5	2.0	5.5
Probability of success (%)	Assay development	71	71	100	100	—
	Simple technical platform development	71	71	75	100	—

This table is adapted from reference 2 under a CC-BY 4.0 license.

Table 6. Potential impact criteria scoring and definitions.

Criterion	Score	Definition	Comments
Cost per result	0	>US\$100 cost per reportable result	Product cost (not necessarily the true cost to serve) is a key decision driver in public tenders and other purchase decisions. Impact and cost effectiveness modelling are frequently insufficient to justify higher pricing than is available for other similar products. This includes the cost per result, which may be a new case identified, a diagnostic test result, a surveillance measure or other similar individual or population-based metric
	1	US\$10–\$100 cost per reportable result	
	2	<US\$10 cost per reportable result	
Improvement in guideline adherence (e.g. WHO guidelines)	0	Low probability; unlikely to impact adherence to guidelines	Adherence to guidelines is a major contributor to success in public health and individual healthcare, along with adoption and access measures
	1	Medium probability; improved adherence to guidelines possible with some interventions required (e.g. task shifting)	
	2	High probability; evidence exists that improved adherence is highly likely	

Criterion	Score	Definition	Comments
Time to intervention	0	Time to intervention/decision likely to be >1 week following deployment of the new test/technology	This can be an important metric in both individual morbidity and mortality measures and in overall public health (e.g. reduced transmission of disease)
	1	Time to intervention/decision could potentially be reduced to <1 day following deployment of the new test/technology	
	2	Time to intervention/decision could potentially be reduced to <1 hour following deployment of the new test/technology	
Improved access to critical data	0	No provisions made for data and information reporting beyond what is locally generated by the test or instrument system	Data and information reporting refer to test results, other test data and meta data and information that is reported in a timely manner to the proper recipients to inform a treatment decision or other healthcare-related intervention. Knowing when and where diagnostics have been deployed and/or reporting of appropriate impact measures are increasingly positioned as requirements by procurers and funders for continued programme support, and by governments and institutions for appropriate resource allocations and understanding of evidence base for positive impact. This applies to individual test results, to MDA decisions and other applications
	1	Data and information reporting needs are understood but gaps remain	
	2	Data and information reporting needs are understood and fully addressed	
Impact metrics	0	The entity has not considered what impact measures will be necessary to satisfactorily demonstrate the product can meet TPP requirements and deliver the intended measurable impact	Having a properly designed TPP is essential for achieving intended impact and long-term sustainability (but not necessarily sufficient). It is incumbent on the entity to produce data showing that it can address the unmet need in the target population with accepted and understood strong impact measures
	1	The entity has reasonably described the impact measures necessary to support the implementation of their product and has a reasonable plan for generating and reporting the required metrics	
	2	The entity has already generated strong data in support of the desired impact measures and has reasonable plans for completing any remaining gaps	
Change in disease incidence or prevalence	0	<5% change in target disease incidence or prevalence is anticipated following introduction of the product or technology	All products should have a demonstratable impact, either directly or indirectly, on the incidence or prevalence of disease. This may be one of the impact measures highlighted in the impact metrics criterion, but is sufficiently important to highlight separately
	1	5–25% change in target disease incidence or prevalence is anticipated following introduction of the product or technology	
	2	>25% change in target disease incidence or prevalence is anticipated following introduction of the product or technology	
Change in morbidity or mortality	0	<5% change in target disease morbidity or mortality per year is anticipated following introduction of the product or technology	All products should have a demonstratable impact on the morbidity (e.g. River Blindness) or mortality (e.g. Cryptococcal meningitis in Advanced HIV Disease) of disease. This may be one of the impact measures highlighted in the impact metrics criterion, but is sufficiently important to highlight separately
	1	5–25% change in target disease morbidity or mortality per year is anticipated following introduction of the product or technology	
	2	>25% change in target disease morbidity or mortality per year is anticipated following introduction of the product or technology	

MDA, Medical Devices Agency; TPP, target product profile.

Figure 1. Optimal diagnostic archetypes based on clustering of 26 diagnostic products. Products are not ascribed to a specific company due to confidentiality agreements. Ag, antigen; AMR, antimicrobial resistance; BC, breast cancer; CAA, circulating anodic antigen; CBC, complete blood count; CMOS, complementary metal-oxide semiconductor; CrAg, Cryptococcal Antigen; HAI, haemagglutination inhibition assay; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; hs, high-sensitivity; IA, immunoassay; LAM, lipoarabinomannan; Lat, latent; PC, pancreatic cancer; RA, rheumatoid arthritis; RDT, rapid diagnostic test; ST, self-test; TB, tuberculosis; Tni, troponin I; TSH, thyroid stimulating hormone; VL, viral load.

included data ranges, averages were calculated for the purposes of inputting into the P2I model. Data were weighted at 0.67 for “concept and research” and 0.33 for “feasibility and planning”, based on expert experience suggesting that greater time and cost is placed at the concept stage.

Table 7 lists the revised diagnostic assumptions. The average costs per phase were higher in the updated assumptions compared with the original P2I v.2 tool assumptions (“discovery”: US\$1.1–24.6 million *versus* 3.0 million (combined “concept and research” and “feasibility and planning”); “clinical validation and launch readiness”: US\$6.9–10.0 million *versus* 3.5 million), with the exception of the cost of the “design and development” phase, which was significantly lower than the initial assumptions (US\$3.3–6.9 million *versus* US\$2.0 to 100.0 million). As it was not possible to calculate specific probability of success scores for each archetype with the available data, an average across all archetypes was used for each phase. The probability of success was significantly lower in the updated diagnostic assumptions compared with the initial assumptions.

Comparison of P2I v.2 and revised models

FIND’s current TB and pandemic preparedness portfolios contained a total of 27 diagnostic candidates under development (**Table 8**). This included three products for Lassa fever, two on multi-disease fever-causing pathogens, one on yellow

fever, and 12 products for TB. Using the P2I v.2 model, 12 products were classified as the “simple technical platform development” archetype and 15 were categorized as the “assay development” archetype. All four development phases were covered. Using the revised model, nine products were classified as the “existing entity, new product/technology” archetype; 10 as the “new or existing entity, discovery” archetype; and eight as the “existing entity, product extension” archetype.

Using the P2I v.2 model, the total estimated cost to move the 27 candidates along the pipeline to launch was US\$641.62 million, with the majority of costs (77%) incurred by the development of the TB diagnostic candidates, followed by the multi-disease fever candidates (20%) (**Figure 2A**). Using the revised model, total estimated costs to move the 27 candidates along the pipeline to expected launch were around US\$274.00 million, substantially lower than the cost estimate from the initial model. However, the costs of developing the yellow fever and Lassa fever diagnostics were higher compared with the initial model due to increased costs during early development (**Figure 2B**).

The total expected number of launches across all 27 candidates using the P2I v.2 model was 21.65, with the majority of launches for the TB candidates (17.82), followed by the Lassa fever candidates (2.42) (**Figure 3A**). Using the revised model, the total launch probability for all 27 candidates was reduced

Table 7. Revised diagnostic assumptions applied to P2I v.2 model.

	Diagnostic archetypes	Discovery ^a	Design and development	Clinical validation and launch readiness	TOTAL ^b	Commercialization (up to 3 years)
Cost per phase (US\$ million)	New or existing entity, discovery	24.6 (2.0–150.0)	6.9 (2.0–25.0)	10.0 (2.0–25.0)	41.5	5.6 (0.5–25.0)
	Existing entity, new product/technology	8.8 (2.0–25.0)	5.0 (1.0–10.0)	7.2 (0.1–25.0)	21.0	4.4 (0.1–10.0)
	Existing entity, product extension	1.1 (0.5–2.0)	3.3 (1.0–6.5)	6.9 (1.0–29.0)	11.3	3.8 (0.5–10.0)
Length of phase (years)	New or existing entity, discovery	4.3 (2.0–12.0)	2.7 (1.0–10.0)	2.2 (1.0–5.0)	9.2	—
	Existing entity, new product/technology	3.8 (1.0–12.0)	2.3 (0.5–5.0)	2.4 (0.5–5.0)	8.5	—
	Existing entity, product extension	1.5 (0.5–3.0)	2.3 (0.5–5.0)	2.4 (0.5–5.0)	6.2	—
Probability of success (%)	All archetypes	18 ^c	80	60	—	45–95

^aPhase based on recommendation from experts and combines “concept and research” and “feasibility and planning” in P2I v2 model. ^bTotal up to “clinical validation and launch readiness”. ^cProbability of success in “concept and research”: 30%; probability of success in “feasibility and planning”: 60%.

Table 8. Candidates in the FIND TB and pandemic preparedness portfolios by archetype and development phase.

Disease (project)	N	P2I v.2 archetype	Revised model archetype	P2I v.2 development phase	Revised development phase
Lassa fever	2	Assay development	Existing entity, new product/technology	Concept and research	Discovery
Lassa fever	1	Assay development	Existing entity, product extension	Clinical validation and launch readiness	Clinical validation and launch readiness
Multi-disease fever	1	Simple technical platform development	New or existing entity, discovery	Concept and research	Discovery
Multi-disease fever	1	Simple technical platform development	New or existing entity, discovery	Concept and research	Discovery
Yellow fever	1	Assay development	New or existing entity, discovery	Concept and research	Discovery
TB	1	Assay development	Existing entity, product extension	Concept and research	Discovery
TB	1	Assay development	Existing entity, new product/technology	Concept and research	Discovery
TB	1	Assay development	Existing entity, new product/technology	Concept and research	Discovery
TB	1	Simple technical platform development	Existing entity, new product/technology	Feasibility and planning	Discovery
TB	1	Simple technical platform development	New or existing entity, discovery	Feasibility and planning	Discovery
TB	1	Simple technical platform development	New or existing entity, discovery	Feasibility and planning	Discovery
TB	1	Assay development	Existing entity, new product/technology	Design and development	Design and development

Disease (project)	N	P2I v.2 archetype	Revised model archetype	P2I v.2 development phase	Revised development phase
TB	2	Assay development	Existing entity, product extension	Design and development	Design and development
TB	1	Simple technical platform development	Existing entity, new product/technology	Design and development	Design and development
TB	1	Assay development	Existing entity, new product/technology	Clinical validation and launch readiness	Clinical validation and launch readiness
TB	4	Assay development	Existing entity, product extension	Clinical validation and launch readiness	Clinical validation and launch readiness
TB	1	Simple technical platform development	Existing entity, new product/technology	Clinical validation and launch readiness	Clinical validation and launch readiness
TB	2	Simple technical platform development	New or existing entity, discovery	Clinical validation and launch readiness	Clinical validation and launch readiness
TB	2	Simple technical platform development	New or existing entity, discovery	Clinical validation and launch readiness	Clinical validation and launch readiness
TB	1	Simple technical platform development	New or existing entity, discovery	Clinical validation and launch readiness	Clinical validation and launch readiness

TB, tuberculosis.

Figure 2. Cost by disease area using **A)** P2I v.2 model and **B)** revised model.

to 11.48 (Figure 3B), a little over 50% lower than the P2I v.2 model projected probability, most likely due to the lower probability of success rates.

The P2I v.2 model predicted that the majority of costs would fall in the first three years of development; development time for all diagnostic candidates was predicted to be completed by 2024 (Figures 4A and 4B). Using the revised model, development timelines were extended, with development time for all diagnostic candidates predicted to be completed by 2028 (Figures 4C and 4D).

Potential impact criteria

Level of risk scores versus potential impact scores for each product are shown in Figure 5. A number of products had the

same level of risk score but substantially different potential impact scores.

Discussion

The P2I v.2 tool predicted that we would require US\$642 million to move the 27 products from the FIND TB and pandemic preparedness portfolios forward through the pipeline, with a launch probability of 21.65 launches over five years. Using our revised P2I model with updated diagnostic archetypes and assumptions, costs were reduced to US\$274 million, with a launch probability of 11.48 launches over eight years. While the diagnostic archetypes and assumptions in P2I v.2 were solely based on interview feedback, our revised archetypes and assumptions were partially based on an analysis of the development history for 26 existing diagnostics, thus our revised tool

Figure 3. Number of launches by disease area using **A)** P2I v.2 model and **B)** revised model.

Figure 4. Total portfolio cost per year (**A** and **B**: P2I v.2; **C** and **D**: revised model).

incorporated actual data as well as expert feedback. The revised outputs are more in line with FIND's experience compared with the outputs of the P2I v.2 model, especially for early stage products. The considerable differences between the outputs of the two models suggest that the current P2I v.2 tool may not accurately capture the research and development costs, timelines, and probability of success for diagnostics.

We found that the previous P2I diagnostic archetypes, "assay development" and "simple technical platform development",

were too confining given the breadth of assay and platform technologies used for the 26 products we analysed from the historical database. For example, some products were being developed on existing platforms and others on new platforms, but both would have been categorized under the "assay development" archetype even though the costs and development times differed significantly. While some products that improve upon current technologies with simple modifications, such as readers for lateral flow assays, fit the "simple technical platform" archetype, others were new systems with novel

Figure 5. Level of risk score versus potential impact score for 26 diagnostic products. Products are not ascribed to a specific company due to confidentiality agreements. Ag, antigen; AMR, antimicrobial resistance; BC, breast cancer; CAA, circulating anodic antigen; CBC, complete blood count; CMOS, complementary metal-oxide semiconductor; CrAg, Cryptococcal Antigen; HAI, haemagglutination inhibition assay; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HPV, human papillomavirus; hs, high-sensitivity; IA, immunoassay; LAM, lipoarabinomannan; Lat, latent; PC, pancreatic cancer; RA, rheumatoid arthritis; RDT, rapid diagnostic test; ST, self-test; TB, tuberculosis; Tni, troponin I; TSH, thyroid stimulating hormone; VL, viral load.

disposables, chemistry and instruments, which were too complex in terms of the technology used to fit under this heading. However, while the updated diagnostic archetypes may better capture the complexity of diagnostics product development, they do not necessarily represent the optimal and complete view of the reality. We recommend continued collection of additional data in order to identify more precisely the archetypes and factors that have the most impact on the assumptions examined, and allow more effective prioritization of investment in diagnostic companies and products.

The reduction in cost using the revised model compared with the P2I v.2 tool is most likely due to the reduced cost for clinical validation in the updated assumptions, as well as the lower probability of success rates compared with the original assumptions. Notably, the lower probability of success

assumptions are in line with previous work on this subject⁷. As expected, our analysis and the feedback from the expert interviewees showed that developing newer products tends to require more investment and development time compared with menu expansion and product extension. This reflects FIND's experience in diagnostic development for LMIC settings. In addition, there is a significant development burden in the discovery stage, particularly under the "concept and research" phase, as successful products must not only be technically sound, but must reflect their intended use.

Evaluating the public health impact of a product in addition to the probability of success is valuable, as products with a high public health impact but low probability of success may still be worth investing in; although riskier than products with higher success probability, they are likely to improve

the health of a greater number of people if successful. We assessed the potential impact of a new diagnostic product at the same time as assessing its level of risk, as a way to prioritize investment in diagnostic products or technologies that, on the surface, display similar cost requirements. We believe that it is important to incorporate this notion into decision-making when estimating funding needs to move candidate health products through the pipeline. Understanding the potential impact of a product can inform decisions on whether to fund or include a high-risk product in a portfolio, as well as supporting advocacy and fundraising efforts.

The interviewees recommended the addition of average research and development costs, timelines and probability of success of commercialization to the model, by inserting a new development phase. Once a product has launched, understanding the costs associated with access and sustainability, especially in LMIC markets, is important, as the probability of success of a new diagnostic test depends not only on the development of the product itself but also on the way it is delivered to the target patient populations. Even when they are available, diagnostic tests are not necessarily affordable, appropriate for use in LMICs, and/or adopted in these settings. To maximize the access to a diagnostic product, a sizeable level of human and financial resources needs to be invested in marketing, sales, manufacturing, distribution and in-country registration, which can lead to higher costs. For more accurate estimates of overall funding needs for diagnostics, future iterations of the tool could incorporate the commercialization phase.

In addition to the limitations of the P2I model already described^{1,2}, there are some further limitations specific to our revised model that should be noted. First, the research and development costs and timelines for product development phase used in the revised P2I model correspond to averages of the historical data collected, as we were only able to obtain ranges of data rather than unique data points, from which averages were calculated to allow for inclusion into the tool. The averages may not effectively capture the large differences that were observed between products within the same archetype, which stem from a complex and multifaceted

environment of the diagnostics industry. Secondly, like the developers of the original P2I tool, we faced challenges in data collection for diagnostic development, as the information needed was either not publicly available or not systematically recorded. For FIND and other PDPs with interest in diagnostics, we recommend improving organizational processes and systems for tracking, storing and sharing this type of information and/or dedicating human resources, such as a Portfolio Manager, to this task. Similarly, the number of diagnostic products used to generate new archetypes and development assumptions was limited, thus more data are required to further validate the proposed archetypes and assumptions.

In summary, we believe that our proposed revisions to the archetypes and assumptions for diagnostic products of the P2I model improve the accuracy of the tool for estimating costs and timelines relating to diagnostic portfolios. During the preparation of this article, a corrected version of the P2I v.2. was published⁸. We hope the data and recommendations presented here are considered for inclusion in a future iteration of this tool, particularly with regards to the inclusion of commercialization and public health impact, which may be applicable to other health products.

Data availability

All data underlying the results are available as part of the article and no additional source data are required. Feedback from expert interviews is not provided due to data protection/confidentiality considerations. The diagnostic products used to inform revised assumptions are not ascribed to a specific company due to confidentiality agreements.

Acknowledgments

The authors would like to thank the experts who were interviewed as part of this study, and Madeleine Werner, Jon Bastow, Samuel Schumacher and Kavi Ramjeet for their contributions to the planning of the project. Medical writing and editorial assistance were provided by Rachel Wright, PhD, funded by FIND, according to Good Publication Practice guidelines.

References

1. Terry RF, Yamey G, Miyazaki-Krause R, *et al.*: **Funding global health product R&D: the Portfolio-To-Impact Model (P2I), a new tool for modelling the impact of different research portfolios** [version 2; peer review: 2 approved]. *Gates Open Res.* 2018; 2: 24.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Young R, Bekele T, Gunn A, *et al.*: **Developing new health technologies for neglected diseases: a pipeline portfolio review and cost model** [version 2; peer review: 3 approved]. *Gates Open Res.* 2018; 2: 23.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Gunn A, Bandara S, Yamey G, *et al.*: **Pipeline analysis of a vaccine candidate portfolio for diseases of poverty using the Portfolio-To-Impact modelling tool** [version 2; peer review: 3 approved]. *F1000Res.* 2019; 8: 1066.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Leeflang MMG, Allerberger F: **How to: evaluate a diagnostic test.** *Clin Microbiol Infect.* 2019; 25(1): 54–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. **TB diagnostics critical pathway.** 2017.
[Reference Source](#)
6. Foundation for Innovative New Diagnostics (FIND): **Vision and Mission.** 2019.
[Reference Source](#)
7. Urdea M: **Critical success factors for diagnostics companies competing in low- and middle income markets 2018** 29 May 2020.
[Reference Source](#)
8. Young R, Bekele T, Gunn A, *et al.*: **Developing new health technologies for neglected diseases: a pipeline portfolio review and cost model** [version 3; peer review: 3 approved]. *Gates Open Res.* 2020; 2: 23.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 27 April 2022

<https://doi.org/10.5256/f1000research.32103.r125480>

© 2022 Lemaire C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Célia Lemaire

EM Strasbourg Business School, Université de Strasbourg, Strasbourg, France

Thank you for the opportunity to review the article entitled "Analysis of diagnostic product portfolios using the Portfolio-To-Impact modelling tool".

In this article, the authors intend to revise an existing tool: The Portfolio-To-Impact version 2 (P2I v.2) financial forecasting tool in order to make the results more accurate and useful for prospective scenario planning (for the initial model, cf. Terry RF, Yamey G, Miyazaki-Krause R, et al.: *Funding global health product R&D: the Portfolio-To-Impact Model (P2I), a new tool for modelling the impact of different research portfolios [version 2; peer review: 2 approved]. Gates Open Res. 2018; 2: 24.*).

The target readers are heads of programs such as FIND how try to "inform funding decisions, refine technology assessments, and predict product launches and estimated research and development costs". The topic is important and the design approach is original. Considering the article as an improving proposition in progress, it is indexable as such. My comments aim to provide for ideas and warnings to strengthen the analysis for the future.

The purpose of the article:

The article argues that the initial model lacks precision, notably due to the complexity of the diagnoses studied. However, if the study carried out presents a notable progress, some uncertainty remains, notably because of the newness of the purpose and the questions related to the methods. This is probably due to a lack of explanation rather than to the inadequacy of the study, but it would be relevant to underline more precisely the advances made in the model and the methodological choices that underlie them.

The methods:

The approach combining the interviews and the retrospective data is fruitful and appropriate for the study. The use of feedback from the 26 products makes the analysis of the tool more refined. My suggestions concerning this part of the study first regards the improvement of the traceability of the results by giving more details on the data collected (1), and more details on the choices made (2). I also propose to deepen the elements concerning the cost structure (3) and to envisage a collection of data as the project progresses (4).

1. To improve traceability of results, provide more detail on the data collected:

Details about the seven experts interviewed should be provided: if the type of diagnosis they are experts in and the archetype in the diagnosis is specified, the reader lacks an understanding of their more precise field of expertise, their function, it should be understood from what precise point of view these experts are speaking, as they are making recommendations that are decisive in the revision of the model (e.g., the allocation of costs between the development phases).

Beyond that, the authors state that the experts provided information on research and development costs, development times and probabilities of success per development phase. I do not question the expertise or the interpretation, but more details would have been appreciated: at this stage, the content of the information and the data analysis process needs to be detailed in order to understand the logic of the demonstration in the article.

2. More details on the choices made:

Regarding the choice of the 26 products studied, it would be interesting to specify the reasons for abandoning the other products for the study.

Concerning the criteria for the level of risk, the authors propose 15 and then 8 criteria. Here, the reasons for the choices are not substantiated, and more clarity is needed on this point. In addition, some criteria were rejected even though they seemed central, for example, the team's competences. It should be specified why they were removed from the analysis.

Similarly, the impact criteria are reduced from 10 to 7 without explanation. This should be corrected to make the study more robust.

3. Scope of costs:

This topic is central to the article, and thus deserves special attention. Five difficulties are identified:

1. Regarding the impact analysis, costs per result are calculated. These are in fact production costs, which are partial costs. While they are useful for comparisons between several organizations, these costs are insufficient for internal analysis in terms of impact. The authors make this point when they suggest that it would have been better to include the commercialization costs.
2. There is little support for the 67%/33% choice in the allocation of costs to phases. The authors suggest that the interviewees indicated that a large proportion was dedicated to design, but this is not sufficient to understand the choice of allocation.
3. The authors show a much lower valuation with their revised model, but the traceability of the process leading to the determination of the costs lacks precision to be totally credible. In particular, the extension in time of the development necessarily induces higher indirect costs for the organizations, but these costs do not seem to

have been taken into account in the model (figure 4 p12).

4. No mention is made of the notion of opportunity costs, which correspond to the cost of abandoning certain products. However, determining these costs is essential in a management decision, because the abandonment of certain products, even expensive ones, can destabilize the overall cost structure.
5. Finally, as the authors point out in the discussion "The averages may not effectively capture the large differences that were observed between products within the same archetype", averages are too encompassing measures, they smooth out the differences. It would be interesting to use medians and to have more precise data, based on a collection over time.

4. Limitations of retrospective data without measurement over time:

The data on research and development costs are based on declarative data. This seems to be problematic because it is a difficult item to evaluate, even from within the organization and retrospectively, if data has not been collected as the project progresses. In addition, different organizations do not have the same cost structures, which can skew the results.

Other recommendations:

First, in a portfolio analysis, costs and probabilities of success are important, but marketability and associated prices are also central. Further studies on these points could be considered.

Secondly, the analysis in terms of impact deserves more investigation, and it would be interesting to specify, here again, the perspective from which the authors are working (public authorities, patients, investors, etc.). It would also be appropriate to consider the competition of other organizations in the model. Indeed, in a competitive dynamic, it is sometimes interesting to propose a 'niche' product in order to differentiate oneself and have a greater impact on a given pathology.

Thirdly, the authors propose that time and resources be devoted to improving organizational processes. This proposal is very relevant, and should be incorporated for analysis. In the same way, the lack of consideration of marketing costs is a pitfall to be overcome in future studies. Fourthly, since the study was conducted in 2019, does the pandemic change the findings?

Minor Notes:

1. The paragraph that states 'developing newer products tends to require more investment and development time compared with menu expansion and product extension' seems unhelpful.
2. To improve clarity, I suggest that different terms be used for different underlying realities. In particular, the use of the word archetype poses a problem for the reader since it refers to different realities as the text progresses; a specific term should be used for each element described. For example, in the introduction, the archetypes refer to "drugs, biologics, vaccines, and diagnostics". Here, the diagnostic part is thus one of the archetypes but is itself composed of archetypes. Then in table 4, the 'archetypes' refer to "Assay development" and "Simple technical platform development". Then, in the results, they are

"New or existing entity, discovery", "Existing entity, new product/technology" and "Existing entity, product extension". The perspectives adopted to distinguish these three kinds of archetypes are very different and should be made explicit in the text, so different terms should be used for each and the term archetype should be dedicated to the authors' contribution (currently in the results).

3. The initial model could be explained with more details. For this, the recently published article with a revision of the model is very clear about the expected inputs and outputs of the model. It would be interesting to draw on it to make the tool's approach more explicit. (Young R, Bekele T, Gunn A, *et al*: Developing new health technologies for neglected diseases: a pipeline portfolio review and cost model [version 3; peer review: 3 approved]. *Gates Open Res.* 2020; 2: 23.).
4. Regarding the notion of risk, it would be necessary to specify which risk is involved (financial, reputational, organizational, etc.) and from which point of view (program coordinator), as this would allow for a better understanding of why a higher development cost leads to a higher risk (Table 3).
5. Figure 5 should specify the reading direction to be clearer (where are the products with the highest risk?)

Thank you for your interesting work, I am looking forward to reading the next version of the paper.

References

1. Terry RF, Yamey G, Miyazaki-Krause R, Gunn A, *et al*.: Funding global health product R&D: the Portfolio-To-Impact Model (P2I), a new tool for modelling the impact of different research portfolios. *Gates Open Res.* 2018; 2: 24 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Young R, Bekele T, Gunn A, Chapman N, *et al*.: Developing new health technologies for neglected diseases: a pipeline portfolio review and cost model. *Gates Open Research.* 2020; 2. [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Areas of expertise: costing in health care organizations, key performance indicators and management control systems in health and social care sectors.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 28 February 2022

<https://doi.org/10.5256/f1000research.32103.r124697>

© 2022 Morton A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alec Morton

Department of Management Science, University of Strathclyde, Rottenrow, Glasgow, G4 0GE, UK

Thanks for the opportunity to review this interesting paper. The research reported here represents an advance in an area where relatively little is known. My view would be that while this a step forward, and is publishable as such, much more remains to be done and my comments should be taken in that spirit.

In terms of the general approach, I would highlight that there is a considerable amount of uncertainty in the inputs but this uncertainty is not apparent in the output numbers (Figs 2 to 4). It is recognised in the structure of the model that managing a portfolio is about managing risk but risk is only captured through the probability of success. I would argue that managing a portfolio also requires managing the cost risk esp as Table 2 shows that cost uncertainty about particular products can be quite substantial. Squeezing out the uncertainty by averaging seems to me a missed opportunity and loses managerially relevant information.

The most obvious gap from the point of view of building confidence in the underlying figures is to close the validation loop, and undertake prospective research to validate these numbers. To some extent this is recognised in the conclusion but this takes the form of relatively weak recommendations to “gather more data”. This data gathering would have to be coordinated by some powerful actor in the field (such as a funder) to ensure methodological consistency and accuracy. Are the authors able to make more concrete recommendations on this validation can be undertaken and confidence built, including saying who should do this and how? I also feel that the paper could be strengthened by clearer explanation and justification of the methods in various places and this is my reason for my “partly” reply to the questions about replicability and also about source data.

Table 2. It is possible to give a statement of what is included in the “costs”? For example, are the

costs simply the direct costs of the staff working on the product, and materials required for development, or do they also include estate overheads including for shared facilities and if so how apportioned? Management overheads? Cost of capital? Etc etc. Do they include payments to third parties in which case the prices will reflect not only the cost to the third party but also the surplus which the 3rd party can command due to its market position? I understand that these are top down estimates rather than bottom up micro costed numbers but it would be good to know more about what the people giving the estimated thought they had been asked to estimate.

“Research and development costs were approximate, and were based either on information from the companies themselves or from persons with knowledge of the companies; additionally, a number of the products were not yet marketed, thus final numbers for research and development costs and timelines were not available.” – is it possible to show on the table which numbers came from the companies and which from outside experts; and where the numbers are forecasts and where they are for realised actuals?

Is it possible to say what the ranges in table 2 mean? Can they be given a probabilistic interpretation, such as an 95% credence interval (“I am 95% sure that the true value lies within this interval”).

Table 3.

Are the individual criterion level risk scores simply added up to come to a composite risk score? In that case has the assumption that the criteria all contribute equally been validated? What is intended profile of the person who is making this assessment, as many of the criteria would require quite a bit of contextual knowledge to assess (e.g experience with the intended customer), and what is the process by which this assessment would be made? Conceptually isn't there an overlap between financial strength and market development criteria?

Probability of success

How was the information from the different interviews aggregated to come up with a probability of success number? Is there a reason why the product are grouped into classes and then a probability of success is defined for the class rather than having the probability success defined as a function of the risk score, which would make better use of all of the information in the risk scoring? Also, trivially but also usefully for the avoidance of unnecessary confusion would it not be more appropriate to call the risk score a “confidence score” as a high score means low risk?

Potential impact

I can see the rationale for the criteria. However, I would like to know about how performance against these criteria can be robustly assessed in practice as for many if not most of the criteria this may be quite challenging to do. I'd also argue that some of the criteria are instrumental (e.g. improvement in guideline adherence) and some relate to more health gain which is a more fundamental objective (eg. the change in disease incidence and change in mortality and morbidity). Given the differing nature of the criteria, just summing up scores may not much sense (Morton, 2017). I'm also puzzled that on fig5 the discovery archetype products seem to be low impact and the extension archetype products are high impact – shouldn't it be the other way around? Maybe it would be good to label the quadrants of the chart as high/ low risk and high/ low impact to make the message clearer.

References

1. Morton A: Treacle and Smallpox: Two Tests for Multicriteria Decision Analysis Models in Health

Technology Assessment. *Value Health*. **20** (3): 512-515 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I do methodological and applied research in resource allocation, decision analysis and health economics relating to health interventions and products, contracting for innovation in new technologies and general public health and health financing.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research