Infermedica

Clinical validation of pediatric tools

Methods and results

Document authors:

Loh I. M.D., Kurkiewicz P., Palczewski M. M.D. PhD, Chwiećko J. M.D., Banaszak F. M.D., Kuszczyński K. M.D., Kwiecień A.

Last update:

July 4, 2022

PM/EN/2021/06/01/1

Table of contents

Introduction to the clinical validation of pediatric tools

Pediatric tools, available within the Infermedica platform, combine trusted medical knowledge and carefully designed technologies. Each of them was created by teams of experienced medical experts and engineers specialized in medical solutions, working according to the highest standards described in ISO 13485:2016 Medical devices—Quality management systems—Requirements for regulatory purposes.

While creating pediatric tools, we have strongly focused on their security and accuracy, making them safe to use by their end-users. Since the very beginning of the development processes, our teams were thoroughly testing and validating every element. The tests started over 2 years ago and have been successively expanded with external tests throughout the past 9 months.

Below, we present the core phases of these validation processes that included:

- internal validation processes (continuously ongoing)
- external validation with pediatricians (2020, Q3)
- external validation with partners (2021, Q2)

Each time, we ran an in-depth analysis of the results, discussed them, and proposed changes that led us to improve the pediatric tools above our initial expectations. The external physicians assessed that the new tool is safe in 96.7% of verified cases. Finally, in the internal pre-release tests, our pediatric tools achieved 99.2% of symptom analysis accuracy, and we released it to the public with this score.

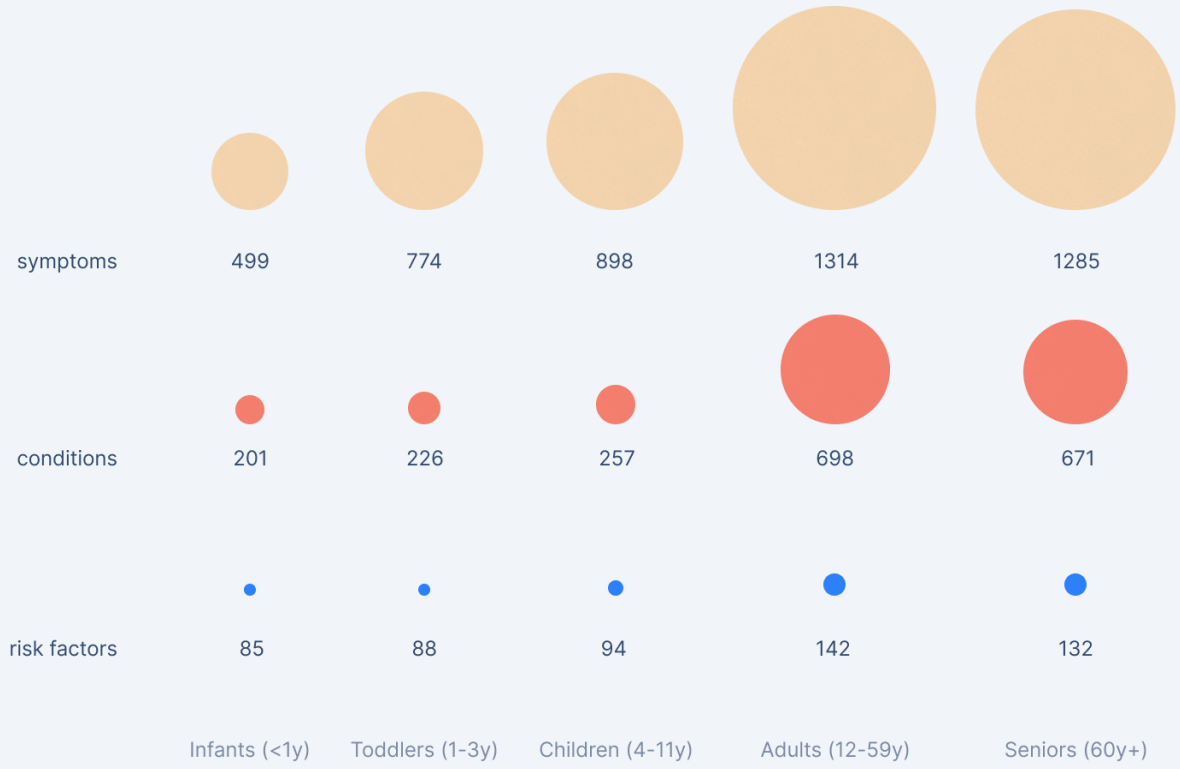
Piotr Orzechowski
CEO at Infermedica

Internal validation processes

Pediatric tools, like any other solutions at Infermedica, were and are a subject of continuous validation processes. Infermedica, as a certified medical device producer, established detailed procedures that help us ensure the highest quality of the medical content used in our symptom checkers, medical chatbots, and other tools.

This process is presented in detail here, and every medical concept added to our medical knowledge base must pass it.

In total, we have added over 1,300 new medical concepts—conditions, symptoms, and risk factors, all adjusted to three age groups in the pediatric population.



The number of medical concepts within Infermedica's medical knowledge base.

Pediatric medical content covers the majority of pediatric specialties and **90% of the pediatric** cases listed by Global Burden of Disease for OECD countries. The database of medical content is constantly being expanded.

In comparison to the medical content for adults, the pediatric one is much more complex, and it has inspired us to propose several new mechanisms, like tuples or detailed age groups, to organize medical content and analyze it.

TRIAGE TUPLES: a special mechanism for identifying health- and life-threatening combinations of symptoms that signal an upcoming rapid deterioration of the patient's health. This mechanism was introduced in order to increase patient safety.

AGE LAYER: an additional layer of algorithms that help us recognize differences between different age groups. That is why interviews for a 3-day-old newborn and a 13-year-old teenager will be significantly different. The layer enables a separate description and probability of symptoms, risk factors, and conditions depending on the exact user age.

At the time when the pediatric content was developed, we established a special Clinical Validation Team, consisting of 7 experienced physicians. In parallel to the Medical Team that was working in the rigorous content creation process, the new team focused and took responsibility for the careful validation of pediatric content and features.

Our pediatrics experts (pediatricians, family doctors, and pediatric surgeons) from the Clinical Validation Team performed subsequent tests. Each condition was individually assessed, and standard acceptance test cases (saTC) were created. It was confirmed that all typical pediatric patients with each list of probable diseases would get accurate results—99.2% of calculated diseases were correct (see details in the last chapter).

The multidimensional approach to the validation of medical content for pediatrics helped achieve an "infinite loop" of feedback and improvements of this new tool.

The process of internal validation started over 2 years ago and continues to be an integral part of pediatric tools today.

Validation with external pediatricians

Along with the intensive tests conducted internally, we have decided to set up 2 types of tests with external experts. The first type was conducted in the mid-development phase of the pediatric tools, falling in September 2020, when we invited a group of experienced pediatricians for collaboration.

This group consisted of 4 experienced pediatricians with over 10 years of practical experience working in 4 large children's hospitals. Medical experience was the most important criterion, followed by an understanding of new technologies.

The tests with external pediatricians covered:

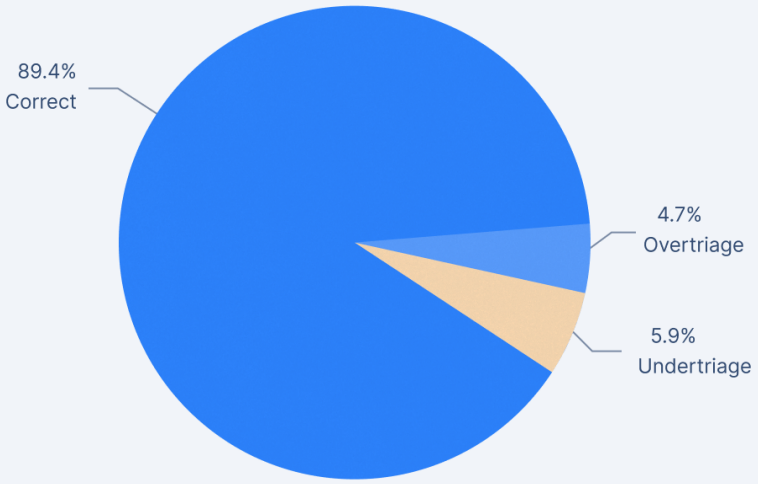
- 150+ medical conditions for three age groups
- 475 test cases
- 120 pediatricians working hours

The goal was to verify the quality of the existing medical content and the level of its triage accuracy. Validation with practicing pediatricians allowed us to gather valuable feedback and use it to adjust solutions to physicians' needs.

Total: 475 test cases

Pediatricians, marked above as P1-P4, conducted 475 test cases in total. The physicians analyzing test cases representing the following specializations: P1 - pediatrician, P2 - pediatrician, neonatologist, P3 - pediatrician, P4 - pediatrician, children gastroenterologist.

Within the tests, pediatricians used 475 manually added test cases to check the triage safety of the pediatric tools. At the end of each interview (symptom check), they marked if the proposed triage is correct, overestimated, or underestimated.



Triage safety in the tests with pediatricians.

As a result of validation with external physicians, our pediatric tools achieved over 94% of safe triage recommendation. This number is defined by the correct triage recommendations and those with overtriage, as in both cases, the patient is directed to the right help.

Pediatrics features gained a positive opinion from independent pediatricians who tested this new feature in clinical set-ups.

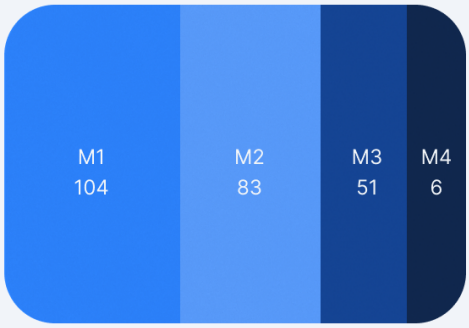
> **NOTE:** At Infermedica, we put patient safety first. Of course, we look for the best way to optimize our triage recommendations, however, when overtriage occurs, we know that patients will be well taken care of. In some cases, this will cause additional costs. We believe this cost is justified.

Validation with 3rd party healthcare organizations

The other type of external tests included clinical validation by physicians working with Infermedica's partners and clients willing to implement pediatrics soon after its release. These tests took place between February and April 2021. Companies taking part in these tests had prior experience with Infermedica solutions, yet the mechanisms and medical content used in pediatric tools were new for them.

Tests conducted by these selected companies were particularly thorough, as all companies were testing both the accuracy of symptom assessment and triage safety of the pediatric tools, as well as the business fit of this new solution.

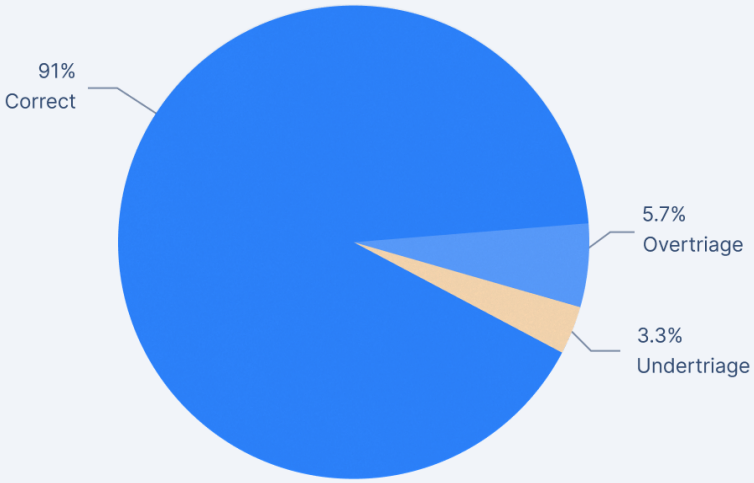
Four independent medical teams from 3 continents took part in the tests. All of them consisted of experienced pediatricians working in the field. In total, they analyzed and rated over 240 test cases.



Total: 244 test cases

Medical teams representing the partner companies performed over 240 test cases in the number of M1 - 104, M2 - 83, M3 - 51, M4 - 6.

Testing teams analyzed test cases representing over 145 conditions, marking their match with the expected triage level.



Triage safety in the tests with partners.

In the results, 91% of test cases were marked with the correct triage and 5.7% of them were over-triaged. **The physicians rated our recommendations as safe in 96.7% of the cases.**

What is the opinion of testing companies?

"I have been working on the pediatrics algorithm, and it has been extremely interesting, it's amazing!" - Infermedica's client

"Please let us know if you require something else. We are very enthusiastic about continuing supporting you all as much as possible." - Infermedica's client

"The tool is great and very innovative!" - Infermedica's client

"The new tool and the job you have been developing is amazing. Thank you!" - Infermedica's client

"The testing is completed from my end. Thank you for the experience!" - Infermedica's client

Pre-release tests of pediatrics

Conducted testing processes helped us to spot multiple chances to improve the medical content and mechanisms used in pediatrics. Analyses and discussions on its results always led us to develop better algorithms used by Infermedica AI.

We are also satisfied with the feedback provided by external pediatricians and partnering companies, as it allowed us to build the new pediatric tools in accordance with the needs of practicing physicians.

Throughout the time the pediatric tool was tested, we used over 1,700 test cases derived from well-established medical journals (e.g., Nelson Textbook of Pediatrics, BMJ, NEJM), guidelines from international organizations (e.g., WHO and CDC), and in-person experience of our physicians.

The pre-release tests conducted internally by Infermedica took place in May 2021. We used them to verify the accuracy of symptom assessment and triage safety of the new tool. The tests were run against 238 test cases from the group of Infermedica's standard acceptance test cases (saTC). Some of them are published among Infermedica's public test cases and are available to review.

The accuracy tests verified the percentage of correct diagnoses, measured as a presence of a tested condition in the results ranking in the TOP 3 or TOP 5 (according to disease incidence). In these tests, the pediatric tools achieved an accuracy level of 99.2%.

Safe triage was obtained for 93.7% of investigated test cases. Within a group of undertriaged cases, the remaining 6.3%, we noted 2 situations. Firstly, where more than one triage would be proper, and secondly, where undertriage could lead a patient to some treatment delay but still pointed clearly that medical help is promptly needed.
All of these undertriage cases are being reassessed as part of the feedback training loops used by Infermedica AI.

Continuous improvement and expansion of the pediatric tools are planned.

Accuracy scores are continuously monitored and improved as part of quality processes at Infermedica per ISO and CE certifications.

Quality System certification

The quality system implemented at Infermedica was audited by an independent certification body, TÜV Nord. The auditor thoroughly analyzed the medical content creation process, as well as clinical validation. Based on the review of documentation and review of product samples, the TÜV Nord audit concluded that Infermedica ensures the high quality of the company's processes and monitors their appropriate parameters.

For more information and more detailed data used for this report, please contact us →

www.infermedica.com