EXSKALLERATE

ACCELERATE adoption of EXOSKELETONS for construction and manufacturing applications in the North Sea Region

Report: Benchmarking guidelines facilitating standardization framework

Benchmarking Metrics and Methodology

Start date of project: 01.01.2020

Duration: 41 months

Coordinator: Tom Verstraten
Leader: R&MM

Responsible person: Tom Verstraten

Revision: 1

Dissemination Level		
PU	Public (when it reaches the published state)	
CO	Confidential, only for members of the consortium	X
SCO	Confidential, Scientific Committee Only	

Document Contributions History

Version	Date	Author	Modifications
1.0	04/09/2020	Haibing Tian	Initial draft
1.1	21/9/2020	Tom Verstraten	Revision of initial content; lay-out
1.2	22/9/2020	Haibing Tian	Slight adjustments of the format
2.0	16/12/2020	Anand Varadharajan, Thierry Hubert	Update of benchmarking metrics based on US army reports
2.1	10/02/2021	Anand Varadharajan	Update of benchmarking metrics and test protocols
2.2	12/03/2021	Anand Varadharajan,	Update of benchmarking metrics based on feedback from partners (Fig. 3b, 3c)
2.3	04/05/2021	Anand Varadharajan, Thierry Hubert	Update of benchmarking metrics and test protocols
2.4	09/08/2021	Tom Verstraten	Revision of document
3.1	07/09/2022	Haibing Tian	Update of testing methods and benchmarking framework
3.2	19/11/2022	Haibing Tian	Update of benchmarking metrics based on feedback from partners
3.3	24/11/2022	Haibing Tian	Format update

Glossary

ADL	Activities of daily living
LBP	Low back pain
MSDs	Musculoskeletal disorders
ROM	Range of freedom
DOF	Degrees of freedom
LPP	Local perceived pain
RPE	Rate of perceived exertion

Executive Summary and report scope

This deliverable collects the performance metrics and testing methods for occupational exoskeletons found in literature and industry. On this basis, we present a list of benchmarking metrics and testing methodology for industrial exoskeletons, which contributes to comparing exoskeletons across different scenarios and creating benchmarking guidelines facilitating standardization framework. Finally, we come up with a benchmarking method and assessment framework to benchmark occupational exoskeletons. The content of the deliverable is:

- Chapter 1 presents the dilemma when evaluating the exoskeletons performance, and points out the importance and necessity of benchmarking evaluation which is helpful and useful to facilitate comparison of diverse industrial exoskeletons.
- Chapter 2 reports the search method that we used to form this report. Papers from Web of science database and the product specifications of commercially available occupational exoskeletons on the market were searched.
- Chapter 3 analyses the results we obtained in academia and industry. A total of 61 papers were considered eligible and 43 exoskeletons were included in our review. The survey of specifications of industrial exoskeletons yielded 44 exoskeletons.
- Chapter 4 compares the performance metrics and testing methods widely used for occupational exoskeleton evaluation in academic and industrial side. By comparison, we found that the research and development of industrial exoskeleton mainly focused on upper-body exoskeleton and back exoskeleton. In addition, both the academia and industry rate the weight and dimension high in terms of design metrics. The difference is that the industry pays more attention to customer-centered metrics than the academia. Furthermore, in the biomechanical metrics, the muscle activity and assistive torque/force gain much interest of people from both fields. With regard to physiological, functional and subjective metrics, the results from industrial circles are much fewer than that from academic side, which can be explained by intellectual property concerns, difficulty of implementation and a lack of interest from industrial side. As for testing methods, despite researchers tend to adopt customized testing task to assess the performance of the exoskeleton, the adoption of other testing methods, such as lifting, holding and overhead tasks, show great consistency in these two fields.
- Chapter 5 presents an exhaustive list of practical benchmarking metrics that accommodates a wide range of exoskeleton types and application domains. The metrics are grouped into five categories: exoskeletons design factors, physiological factors, biomechanical factors, functional and subjective factors. Besides, a basic library, which includes common industrial activities, is proposed to decompose complex tasks into basic tasks. With regard to the basic task, we provide a benchmarking method to evaluate the result. In the end, we propose a benchmarking framework which can contribute to comparing exoskeleton across different scenarios and creating benchmarking guidelines facilitating standardization and market adoption.

Table of Contents

1. Introduction

Although the robot technology and automation have spread rapidly in the EU in recent decades, there are still many occupational tasks to be handled manually, such as assembling, repetitive lifting/lowering, carrying. However, these tasks may expose workers to work-related health problems. According to European Risk Observatory Report, roughly three out of every five workers in the EU-28 report work-related musculoskeletal disorders (WMSDs) complaints (De Kok, Vroonhof et al. 2019). WMSDs may not only cause absence from work, but also lead to a great burden on the national finance. Hence adopting appropriate measures to alleviate WMSDs without affecting workers' flexibility and the productivity of enterprises gains much interest in industry.

An occupational exoskeleton is a wearable device designed to support users performing physical tasks in a working condition (de Looze, Bosch et al. 2016, Frisoli 2019). Occupational exoskeletons are designed to enhance workers' physical and cognitive capabilities without hindering their flexibility and agility. This assistance supported by exoskeletons exists in the body parts prone to fatigue and injury. Therefore, based on the different supportive regions, occupational exoskeletons can be divided into four different categories (Crea, Beckerle et al. 2021). Upper-body exoskeletons are developed to provide support in the upper body, while executing tasks such as lifting and overhead. Back exoskeletons aim to reduce compression force on users' lower back (L5-S1 disc) by providing an assistive torque. Lower-body exoskeletons are designed to decrease the load on the lower extremities for the completion of physical work. Full-body exoskeletons are intended to provide assistance at full-body level. Examples are shown in figure 1 based on this classification. In addition, due to the diverse actuation architecture, exoskeletons also can be classified into three groups (Crea, Beckerle et al. 2021): active, passive and quasi-passive exoskeletons. Active exoskeletons adopt actuators (electrical, hydraulic, or pneumatic) to generate torque to aid the wearers. Passive exoskeletons are characterized by using elastic elements (spring and damper) to store and release energy at different stages of wearer's movement to assist the human. Semi-active exoskeletons consist of elastic elements and actuators used to change the properties of these elements. At present, active exoskeletons are mainly used for rehabilitation purpose. The industrial application of these exoskeletons is relatively few. Compared with active exoskeletons, passive exoskeletons have been quickly deployed and adopted in industry due to their lightweight, safety, comfort, and low cost. Implementation of occupational exoskeletons has the potential to reduce perceived musculoskeletal effort and loading on the supported region thereby reducing the risk of WMSDs (de Looze, Bosch et al. 2016). Several studies have reported beneficial effect of exoskeletons on physical load when workers perform lifting/lowering (Spada, Ghibaudo et al. 2017, Spada, Ghibaudo et al. 2018, Yong, Yan et al. 2019, Kim, Little et al. 2020, Wei, Zha et al. 2020, Yin, Yang et al. 2021), overhead (Huysamen, Bosch et al. 2018), bending (Zhang, Kadrolkar et al. 2016, Di Natali, Chini et al. 2021, Yin, Yang et al. 2021) and field tasks (Liu, Hemming et al. 2018, Smets 2019, Pinho, Taira et al. 2020, Wang, Le et al. 2021). Although numerous exoskeletons assessments have been conducted, evaluating the performance of exoskeletons is not straightforward due to the fact that diverse testing methods and performance metrics are typically employed for specific robotic systems and functional scenarios. The lack of a consensus on the performance metrics and testing methods for the evaluation of occupational exoskeletons performance still exists (Lorenzo, Baojun et al. 2019, Babič, Laffranchi et al. 2021).

Benchmarking is one of the pillars of this issue as it is not only a tool to assess and compare the performance of different technologies (Elprama, Vanderborght et al. 2022), but also a way to define and support standardization and certification necessary to introduce exoskeletons to the market (Torricelli, Rodriguez-Guerrero et al. 2020). By now, the benchmarking in the context of industrial exoskeletons is still in its infancy, although some metrics and tests have been implemented over the recent years. Despite some efforts have been proposed to assess exoskeleton performance, the vast majority of the published work reports a few performance evaluation, ranging from biomechanical performance (Ranaweera, Gopura et al. 2018, Han, Du et al. 2019, Ji, Wang et al. 2020, Kim, Little et al. 2020, Pinho, Taira et al. 2020, Di Natali, Chini et al. 2021), kinematic performance (Madinei, Kim et al. 2021), physiological performance (Baltrusch, van Dieën et al. 2019), to a preliminary user-based validation (Groos, Fuchs et al. 2019, Hensel and Keil 2019, Baltrusch, Houdijk et al. 2021, Diana, McNamara. et al. 2022, Elprama, Vanderborght et al. 2022). Such a low-dimensional evaluation not only shows a limited portion of exoskeletons' performance but can also be misleading in its interpretation (Moltedo, Cavallo et al. 2019). All this points towards a clear need for standardized benchmarking tools of industrial exoskeletons. A performance index (or several ones) and an accompanying methodology need to be devised that are applicable across different scenarios while considering specific features of each device.

This study provides an overview on the current metrics and testing methods used to evaluate the performance of occupational exoskeletons in academic and industrial environments. The differences in metrics and testing methods between academia and industry were identified. On this basis, an exhaustive list of practical benchmarking metrics that accommodates a wide range of exoskeleton types and application domains has been proposed. Furthermore, we establish a basic tasks library considering numerous common activities in industry. By referring it, the actual task in the factory can be decomposed into several basic tasks. In addition, a calculation method to assess the performance of exoskeleton in each basic task is presented. Finally, an assessment framework to benchmark capabilities of occupational exoskeletons is reported.

Upper-body exoskeleton (Ottobock)

Back exoskeleton (Backx)

Lower-body exoskeleton (ChairlessChair)

Full-body exoskeleton (GuardianXO)

Figure 1. Examples of each exoskeleton based on diverse support regions.

2. Methods

In order to more comprehensively characterize and evaluate occupational exoskeleton, we collected data from academia and industry respectively. Literature review is the most direct way to gain an insight into what performance indicators and testing methods researchers use to evaluate the performance of occupational exoskeleton. Those indicators and testing procedures can be clearly obtained by reviewing full texts. Differently, many factors, such as material, technology, process, human-machine interaction, market and psychology, need to be considered in industrial circles when designing an industrial product (Baxter 2018) to achieve a certain functions. By investigating the specifications of commercially available exoskeleton, we are able to know the most commonly used parameters of manufacturers to show their exoskeleton performance and appeal customers. Despite the fact that suppliers do not disclose specific testing procedures and protocols for evaluating the performance of exoskeleton, they specify the applicable tasks of their products, such as lifting, holding task. Considering that manufacturers report on their websites the results and benefits of using their exoskeletons to perform these tasks, therefore, to some extent, the applicable task can be considered as their testing task. To investigate the difference of performance metrics and testing methods from academia and industry, we performed a thorough study consisting of two parts:

2.1 Survey of academic literature

This academic overview combines electronic literature search by using Web of Science and personal database search. Papers, on performance metrics and relevant evaluation tests, published in English and peer-reviewed journals from January 1995 to July 2022 were considered. More specifically, papers had to present testing metrics, methods and results employed in the assessment of occupational exoskeletons. It is noted that articles reporting exoskeletons applied for rehabilitation purposes or for assistance of disabled and/or weakened individuals were excluded. Additionally, papers that only presented exoskeleton prototypes without any evaluation experiments on participants were considered ineligible. The following search strategy was used: TS=((“exoskeleton” OR “assistive device” OR “exosuit” OR "wearable devices") AND ("test" OR "benchmark" OR "assess" OR "evaluate" OR "metric") NOT ("rehabilitation")). The specific selection process of articles can be found in figure 2.

2.2 Survey of industrial specifications

The investigation of commercially available industrial exoskeletons disclose the diverse characteristics of exoskeletons, reflecting the requirements of the industrial sector, which are different from the demands of academic circles. This search was completed by using Exoskeleton report website and personal database search. We reviewed the product specifications of commercially available industrial back, lower-, upper- or full-body exoskeletons. These results were obtained from the website of the respective manufacturers. The selection process of occupational exoskeletons is reported in figure 3.

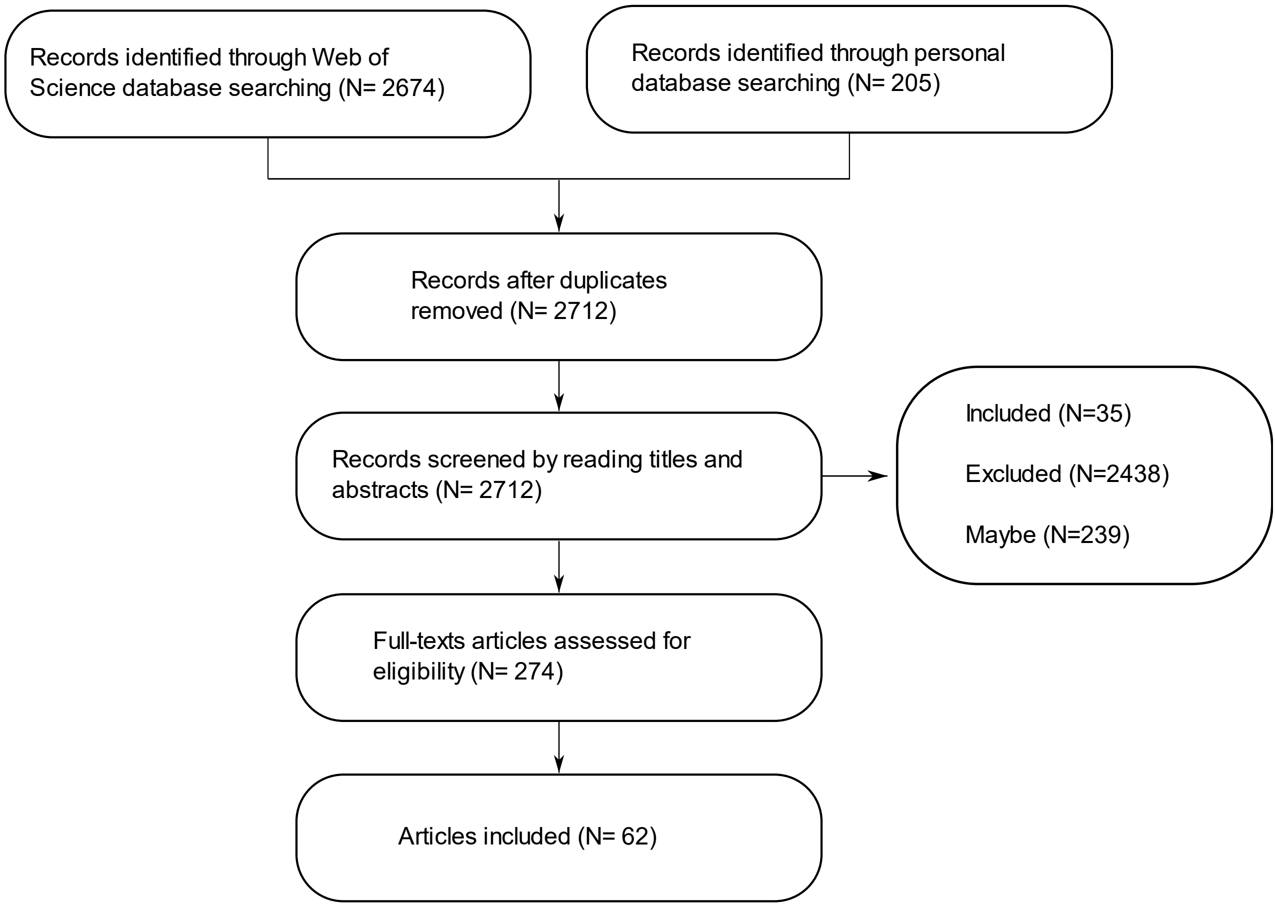


Figure 2. Selection process of academic literature

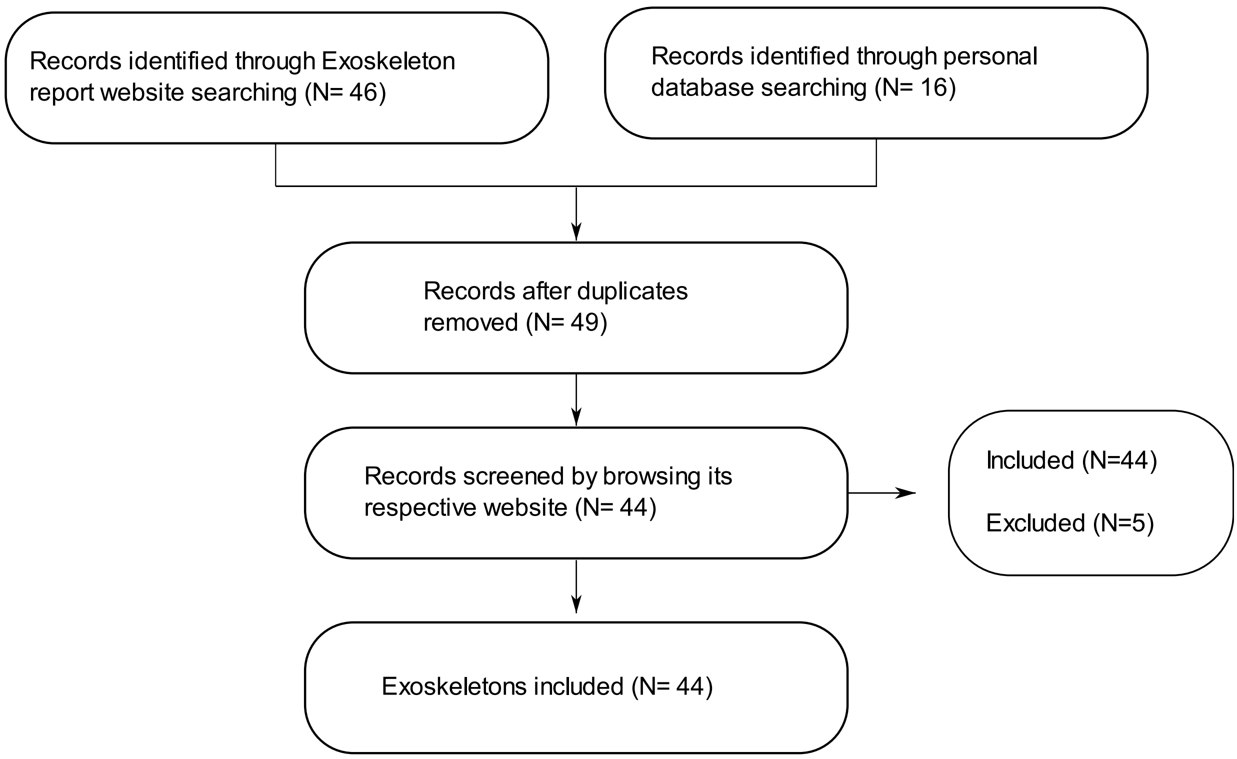


Figure 3. Selection process of industrial exoskeletons on the market.

3. Results:

3.1 Survey of academic literature

The results retrieved after the first search completed on July 18, 2022, were 2674. All works were imported into the Rayyan application. After removing the duplicates, by screening the titles and abstracts, articles that were not related to industrial exoskeletons and that only performed simulations without experimental assessments were considered to be excluded. The third selection was performed by reviewing full texts, which resulted in 61 studies included. It is worth noting that (Diana, McNamara. et al. 2022) evaluated both upper exoskeletons and back exoskeletons. Additionally, the K-SRD^{TM} was included in this review. Because (Hondzinski, Ikuma et al. 2018) researched the effect of K-SRD^{TM} on work-related tasks, though K-SRD^{TM} was originally designed for military applications. Results filtered from selected articles are provided in appendix 1.

3.2 Survey of industrial specifications

Most industrial exoskeletons on the market and their related classifications are reported in the Exoskeleton Report (ExoskeletonReport). Other industrial exoskeletons not included in the Exoskeleton Report but found on the market are also considered eligible. The detailed information of each occupational exoskeleton can be searched by browsing its respective website. This search was completed on July 22, 2022, and results obtained are reported in appendix 2.

3.3 Exoskeletons

All exoskeletons included in these surveys are anthropomorphic in design. Thus, based on the diverse supportive regions, occupational exoskeletons can be divided into different groups: arm-support, back-support, leg-support and full-body exoskeletons.

In total, 72 exoskeletons are assessed in 61 papers, as listed in appendix 1. Considering that the same exoskeleton could be assessed in several papers, we remove the duplicates and get the following non-redundant results. Finally, 44 occupational exoskeletons are included in this field. As for commercial exoskeletons, a total of 42 commercially available exoskeletons (appendix 2) are included. It should be noted that Muscle Upper and Shiva Exo can be used as both arm-support and back-support exoskeletons (ExoskeletonReport). They are included both in arm-support and back-support exoskeleton group. This explains why the sum of exoskeletons in industry is 44. A summary of those exoskeletons together with exoskeleton type and actuation type is represented in table 1.

3.4 Test types

Industrial circles have not disclosed details of the testing protocols of occupational exoskeletons. On the other hand, academic research papers clearly describes these details. Of the 61 articles, 7 field tests and 50 laboratory tests were carried out. Besides, there are four papers assessing exoskeletons under both laboratory and field conditions. Therefore, in academia, there are a total of 11 field tests and 54 laboratory tests evaluating occupational exoskeletons.

Table 1. Retrieved industrial exoskeletons: number, type and actuation type of exoskeletons.

Source	Number	Exoskeleton type	Actuation type
Academia	Exoskeleton (N=44)	Arm-support (N=18)	Active (N=4)
			Semi-Active (N=1)
			Passive (N=13)
		Back-support (N=16)	Active (N=8)
			Passive (N=8)
		Leg-body (N=8)	Active (N=1)
			Semi-Active (N=2)
			Passive (N=5)
		Full-body (N=2)	Active (N=2)
Industry	Exoskeleton (N=44)	Arm-support (N=17)	Active (N=2)
			Passive (N=15)
		Back-support (N=23)	Active (N=8)
			Passive (N=15)
		Leg-support (N=3)	Passive (N=3)
		Full-body (N=1)	Active (N=1)

3.5 Metrics

As one or more performance indicators of exoskeleton clearly pointed out by authors in each paper (Appendix 1), the exoskeleton company also reported related metrics on its website (Appendix 2). This led to multitudinous metrics used to characterize different features of exoskeletons. To make these metrics easier to distinguish and compare, we have grouped them into five categories:

- Design metrics: Design metrics are parameters that reflect the nature of an exoskeleton, such as, actuation type, applicable height, battery life. Those metrics are not only helpful in understanding the properties of an exoskeleton, but also useful to provide information for evaluating the performance of an exoskeleton.
- Physiological metrics: Physiological metrics indicate the overall effect of the exoskeleton on the human physiological response.
- Biomechanical metrics: Biomechanical metrics are used to assess the biomechanical effect of an exoskeletons on the user. These metrics not only indicate to what extent an exoskeleton hinders movements of the human body, but quantify to what extent an exoskeleton can reduce the mechanical loading on body structures.
- Functional metrics: Task-related metrics to assess the exoskeleton performance.
- Subjective metrics: Subjective metrics, usually obtained through questionnaires, scales, and interviews, are used to show the individuals' perception of an exoskeleton.

Based on the metric classification mentioned above, we collected the metrics used to evaluate the exoskeleton in laboratory tests, field tests, and industry separately. Following this, we added up these indicators separately. Eventually, the performance metrics for evaluating different types of exoskeletons in different fields are depicted in figure 4.

3.6 Testing methods

The comparison of testing methods performed to evaluate exoskeleton performance in laboratory tests, field tests, and industry is presented in figure 5. It is worth noting that customized task, including multiple test phases, is designed specifically by the researchers to assess performance aspects of exoskeletons (Spada, Ghibaudo et al. 2017, Liu, Hemming et al. 2018, Spada, Ghibaudo et al. 2018, Baltrusch, van Dieën et al. 2019, Baltrusch, Houdijk et al. 2021). One of the most widely used customized tasks is the test battery proposed by Baltrusch et all, which includes 12 different tests to evaluate an exoskeleton.

4. Discussion

4.1 Exoskeletons: types and proportion in industry and academia

4.1.1 Based on body parts supported

As it can be seen in table 1, in the academic literature, arm-support and back-support exoskeletons account for a large proportion of all exoskeletons considered eligible, making up 40.91% and 36.36%, respectively. In contrast, the percentage of leg-support (18.18%) and full-body (4.55%) exoskeletons is far less than that of previous two types. A similar trend can be observed in industry. Arm-support, back-support, leg-support and full-body exoskeletons occupy respectively 38.64%, 52.27%, 6.82% and 2.27% of all commercial industrial exoskeletons. With regard to occupational exoskeletons, both academia and industry place much emphasis on the arm-support and back-support exoskeletons, compared with leg-support and full-body exoskeletons. Multiple research institutions and companies are currently developing such exoskeletons for occupational use. It is not only because arms and back are involved in the vast majority of industrial work and become vulnerable, and these exoskeletons have the possibility to reduce the WMSDs, but these exoskeletons are of good compactness and versatility. On the other hand, numerous joints and degrees of freedom in anthropomorphic full-body exoskeletons inevitably result in a high complexity of the device and a correspondingly high volume and weight. The latter are considered an important barrier for the large-scale adoption of full-body exoskeletons in actual factories. Leg-support exoskeletons are seldomly used in industry. they are primarily developed for assisting the disabled or the elderly. Although the Chairless Chair, as one of the outstanding examples of lower limb exoskeletons for industry, can help workers with prolonged standing, its application is greatly limited by the working environment, especially for the work that needs to be moved frequently. Besides, compared with versatile back-support and arm-support exoskeletons, it is relatively narrowly researched and utilized due to its limited function.

4.1.2 Based on actuation type

Passive exoskeletons are lighter, cheaper, simpler, and easier to implement than active exoskeletons, which might explain why passive exoskeletons are more prevalent (60.47% in academic literature, 75% in commercial exoskeletons). However, active exoskeletons can provide auxiliary force in a more flexible way, as their assistance is not related to the wearer's kinematics as in a passive exoskeleton. This makes them potentially much more functional in industrial environments. Still, active exoskeletons are rarely adopted in the working environment. The main obstacles are their elevated prices, large mass and volume, lack of relevant standards, lack of insurance support and increased number of potential failure modes and corresponding safety hazards. Most of these issues are due to the implementation of actuators. In anticipation of drastically lighter and more compact actuation solutions, which are still an important research focus, semi-active devices may provide a good middle ground between the functionality of active devices and the lower mass (Babič, Laffranchi et al. 2021, Verstraten and Lefeber 2021). These semi-active devices typically allow assistance to be switched of when it is not desired. In some cases, they also allow for minor adjustments to the assistance level.

4.2 Test types: laboratory tests and field tests

As mentioned in 3.4 section, in the retrieved literature, the number of laboratory tests (n=54) are considerably more than that of field tests (n=11). According to the Technology Readiness Levels (TRLs), both laboratory and field tests are needed in the process of product development. The laboratory tests mainly tend to assess feasibility and to quickly verify design parameters with small sample sizes, so as to facilitate design parameters as an input for subsequent rapid design iterations Tests carried out under controllable laboratory condition are easy to conduct and take a short time to obtain results. Once those obtained results become stable as well as reliable, field tests become necessary. Different from lab tests, field tests, more prone to environmental disturbances but reflecting real exoskeleton performance in the actual workplace setting, are more difficult to perform because of limitations of working environment, potential hazards, infeasible experimental instrument and other factors. Despite field tests are indeed important, there are relatively few evaluations (16.92%) and reports on them. In turn, it is also the lack of field verification, especially long-term verification (3.08%), that hinders the widespread adoption of industrial exoskeletons (Crea, Beckerle et al. 2021, Elprama, Vanderborght et al. 2022). In this review, only two relatively long-term assessments have been conducted to evaluate the effect of back exoskeletons (3 weeks) (Amandels, Eyndt et al. 2019) and upper-body exoskeletons (3 months) (Smets 2019) in an industrial working setting. Strictly speaking, Smets et al. collected the use patterns of exoskeletons during 3 months, which is not equivalent to workers using exoskeletons for three months. Long-term validation is still of great significance and necessity to investigate the effect of exoskeletons on wearers. Long-term verification is not only an approach to test exoskeletons, but also a tool to pave the path for occupational exoskeletons benchmarking. With regard to tests in industry, manufacturers clearly indicate which type of tasks can be assisted by using their device, for example preventing back injuries for lifting tasks. However, the specific tests that are carried out to come to this conclusion are not described. In the future, providing protocols to assess these technologies would be of great benefit towards benchmarking. Similarly to the automotive industry in which testing protocols are set. Ideally, such protocols would replicate with as much fidelity as possible the real-world situations. Only when long-term and large sample data verification proves to be effective, will government, insurance companies and practitioners support the adoption of occupational exoskeletons.

4.3 Performance metrics

4.3.1 Differences in performance metrics between laboratory tests and field tests

Since more tests are conducted in laboratory environments (N=54) than in the field (N=11), it is no surprise that laboratory tests contribute more performance metrics than field tests (figure 4). Especially in design metrics, conveying user basic product information, much fewer results can be found in field tests. It can be explained that field tests are usually conducted with mature devices to avoid hazards and unforeseen circumstances, thus researchers tend to pay more attention to test results instead of accessible product specifications. The only design metric mentioned in field tests is the weight of exoskeleton. Weight is considered a very important indicator when designing exoskeletons (Schnieders and Richard 2018). It directly affects wearer's comfort and acceptance (Toxiri, Sposito et al. 2019). However, the weight should be taken into account together with the assistance offered by the exoskeleton. Clearly, the passive exoskeleton is lighter, but correspondingly, it may provide less physical assistance than the active exoskeleton.

Figure 4. Metrics used to report performance of different types of exoskeleton from academia and industry

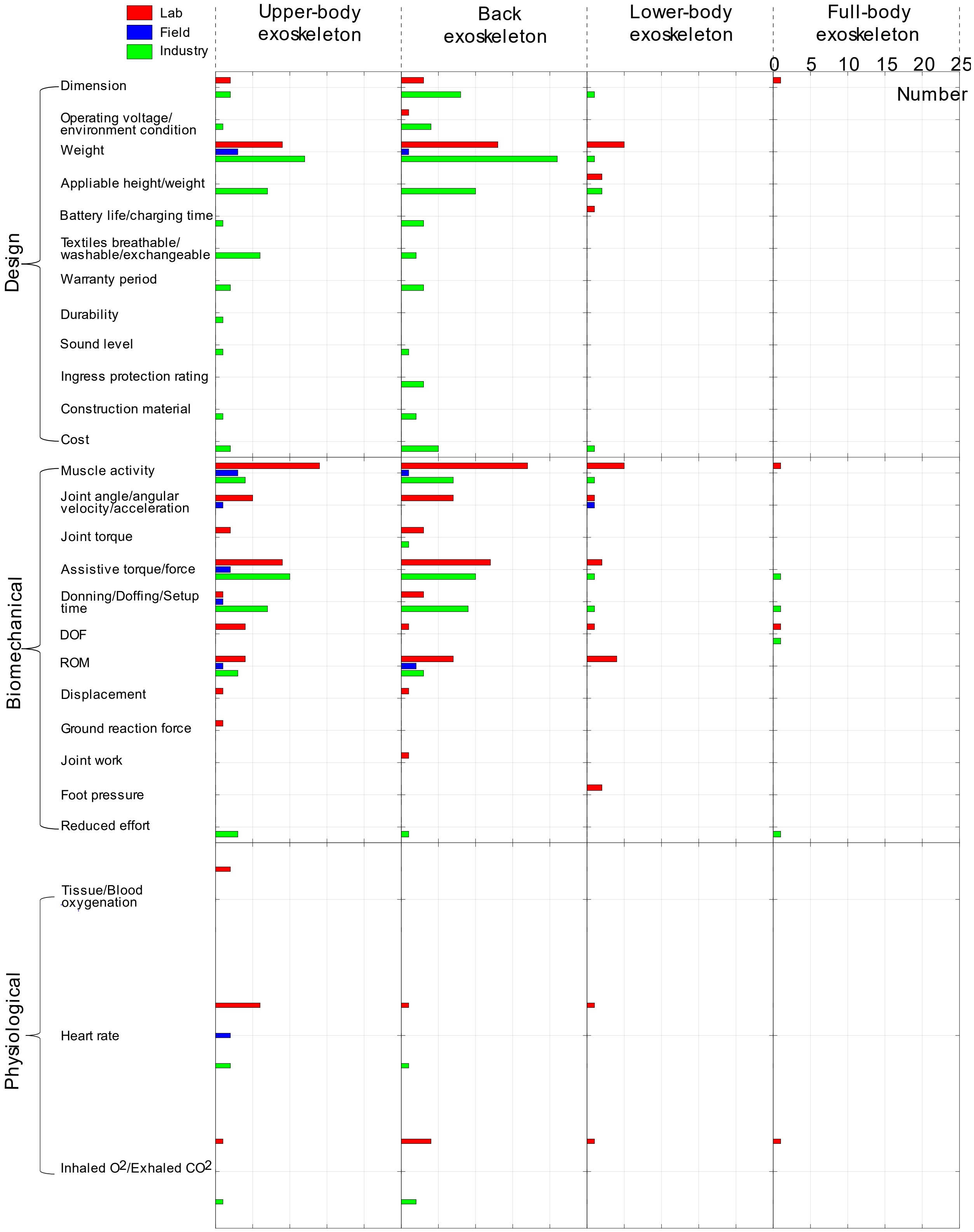


Figure 4. (Continued) metrics used to assess performance of different types of exoskeleton from academia and industry

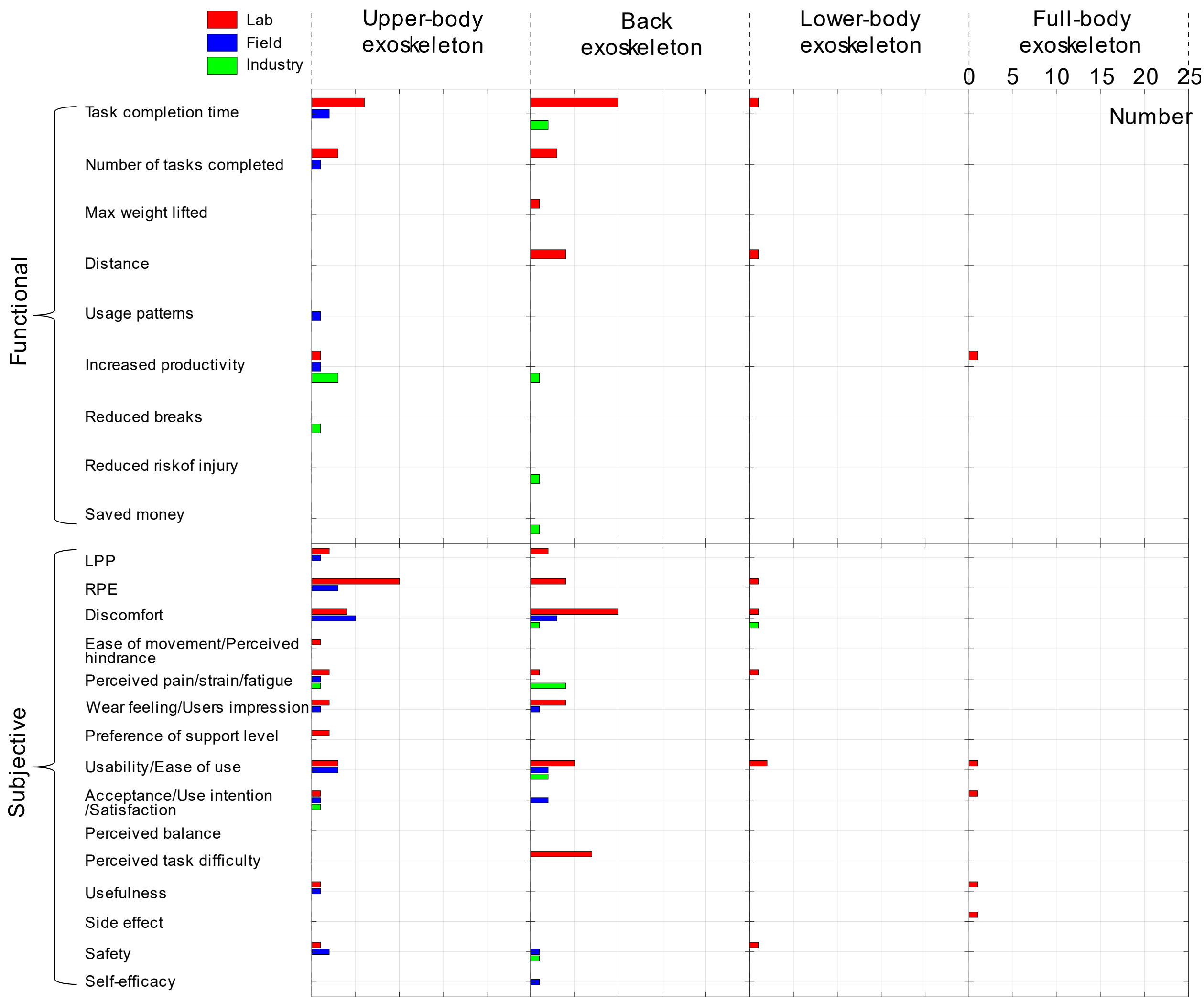
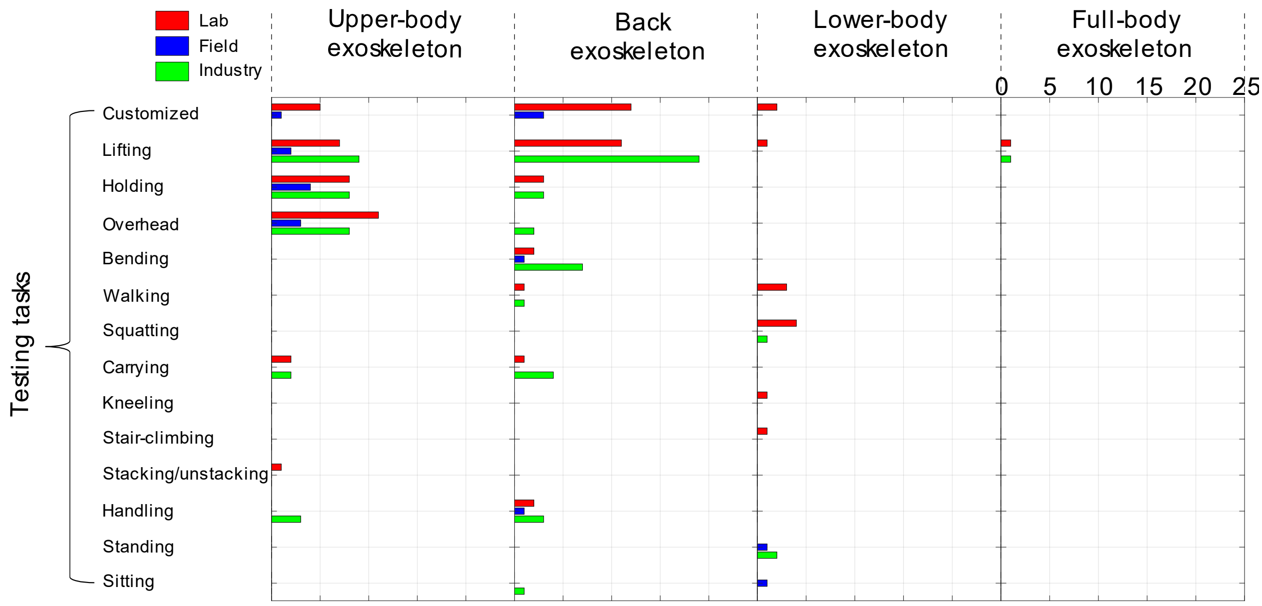


Figure 5. Testing tasks used to report performance of different types of exoskeleton from academia and industry.



With regard to biomechanical metrics, various and numerous metrics are introduced in the evaluation of exoskeleton performance under laboratory condition (figure 4). It is found that muscle activity, joint angle/angular velocity/angular acceleration, assistive torque/force, ROM, and donning/doffing/setup time are the most widely used metrics in lab tests. These metrics are also adopted in field tests, despite the number of these metrics is less. As a frequently used indicator, muscle activity delivers user the information on the influence of exoskeleton on individual muscles. However, the implementation of EMG signal measurement is not straightforward, because it requires much preparation, for instance, it needs an expert to guide how to place electrodes correctly. And the electromagnetic interference in a manufacturing environment is another issue that may affect EMG signal measurement. Furthermore, wearing exoskeleton will inevitably change the working postures as well as movements of workers to a certain extent. In any case, the exoskeleton should not seriously hamper workers during performing tasks. ROM and kinematics of joints can be used to assess the hindrance caused by exoskeleton during tasks. In the design process, it is important for designer to decrease ROM and kinematics limitation, cover more applicable tasks, and so as to improve the diversity of exoskeleton. Although only one study reported donning/doffing/setup time, this metric is crucial to the adoption of an exoskeleton. Prolonged exoskeleton wearing and setting operations will compromise the normal working schedule. Another important metric is the assistive torque/force, which indicates to what extent the exoskeleton is able to help wearers. Other parameters, such as ground reaction force and foot pressure, gain less interest in field tests. It would be possible to measure these metrics only if the user wears dedicated insoles. However, it is challenging to implement it in a factory environment. The ground reaction force is also needed to calculated the joint torque. It is clear that there is no assessment of joint torque in field tests since there is no result of ground reaction force in field tests..

Compared with other metrics, the results of physiological metrics are less. In these few results, we found that lab testing has a high priority in heart rate and inhaled O_2/exhaled CO_2. The decrease in heart rate and oxygen consumption indicates the effectiveness of exoskeleton during the task from a physiological point of view. However, only 2 field tests studied the effect of exoskeleton on user's heart rate. To measure the oxygen uptake, the user has to wear a mask during performing the task, which may interfere with specific work. This may explain why there are fewer tests on these indicators in the field.

In functional metrics, both lab and field tests report the task completion time and the number of tasks completed. The number of tasks completed in a certain time or the time required to complete a certain number of tasks is measured in those tests. Those two metrics are focused on task performance. By comparing those metrics with and without exoskeleton, the change of productivity for a certain task can be obtained. Additionally, another metric, the usage pattern, was only assessed in the field tests. The use patterns of exoskeletons during 3 months were recorded to reflect the user's acceptance. Other task-related metrics were only evaluated in the researcher's customized testing tasks. Due to the site constraints, indicators related to these customized tasks are easier to obtain in a flexible laboratory environment rather than in the field. For instance, Yong-Ku et al. studied the changes of muscle activity with and without exoskeleton during a simulated harvesting task at different working heights (Kong et al., 2021). In this case, the independent variables (working height and muscle type) are easier to control in laboratory conditions than in the field..

It is interesting that both lab and field tests provide much information of subjective metrics to assess the exoskeleton. In figure 4, we can find that the rating of perceived exertion (RPE), discomfort and usability are the most widely used parameters in laboratory tests. Likewise, these metrics are also applicable to field tests, although the reported number of same indicator in the field is far less than that in laboratory. It makes sense because considerably more tests are carried out under laboratory condition than under field condition. However, these indicators all account for a dominant proportion in their respective fields. RPE is measured based on the BORG-CR10 scale to indicate how hard an individual feels in an activity. The discomfort may affect the task execution. Ideally, the exoskeleton would reduce discomfort in targeted region without bring discomfort in other body areas. Another significant indicator is usability, because the complicated operation may hinder the use of device. The remaining indicators cover many categories, but the number of each category is quite small. These indicators gain less interest from researchers, although some metrics are important, such as safety and usefulness. It is the lack of consensus on subjective assessment that results in numerous and multifarious metrics. It is urgent to have a set of benchmarking metrics which can be used for researchers to compare diverse exoskeletons.

4.3.2 Differences in performance metrics between academia (lab and field tests) and industry

As shown in figure 4, by comparing design metrics from academia with metrics from industry, it is clear that metrics reported by industry are more customer-centered. Characteristics such as applicable height, applicable weight, warranty period, detachable and washable textile parts, battery life and battery charging time are sometimes neglected in the context of academia, yet prove to be decisive in industrial contexts. It is interesting that both academia and industry pay much attention to weight. The exoskeleton would be worn by the user to engage in a specific task during work. Thus, keeping the exoskeleton as light as possible is critical. Besides, dimension is another important parameter, because bulky equipment cannot be used in some environments with limited space. The implementation of these equipment may hinder the movement or work of staff.

With regard to biomechanical metrics, EMG gains the interest of practitioners both from academia and industry. Muscle activity was one of the most common metrics used to evaluate the exoskeleton. Through comparison of the muscle activities before and after using exoskeleton, the impact of exoskeleton on human body can be intuitively observed. Kinematics is another significant parameter for researchers to investigate the position and orientation of the body segments, the angles of the joints and the corresponding velocities and acceleration. Normally, these metrics can be obtained through the laboratory experiment by using motion capture device. However, in the factory, the implementation of camera based motion capture equipment is relatively limited due to obstruction. Using a IMUs-based (inertial measurement units) motion capture suit to measure these metrics in industry might be a good solution, but this is not always possible because of data drift and noise. This could explain why these metrics are less assessed in industry. Similarly, DOF and ROM, which are also kinematic metrics, are less important in industry, however, these metrics directly show the hindrance of an exoskeleton to individuals. Besides, assistive force, which accounts for almost the same proportion in academia and industry, is another important kinetic parameter, because despite the weight of the active exoskeleton is being higher than that of the passive exoskeleton, it provides higher force correspondingly. Time is very significant in the manufacturing industry, especially in assembly line. Therefore, people in industry pay more attention to this criteria, donning/doffing/setup time, than those in academia. Other metrics, such as foot pressure, joint work, joint torque, are only researched only in academia.

Although physiological metrics play an important role in reflecting the effectiveness of the exoskeleton on the user, they are seldom assessed under industrial condition. This may be because it is difficult to measure the physiological parameters of workers in a real factory environment. In retrieved papers, heart rate is found to be the most widely used physiological parameter to characterize the physiological benefit of the exoskeleton. This is followed by inhaled O_2/exhaled CO_2, which are examined to calculate the metabolic cost. Tissue and blood oxygenation are barely used to evaluate exoskeleton performance. For each of these two metrics, only one paper investigated the effects of the use of industrial exoskeleton on the parameter.

In functional metrics, it is clear that researchers are more concerned with task-specific parameters, for instance, task completion time, the number of tasks completed and carrying/walking distance. Task completion time is the most frequently evaluated index by scholars followed by the number of tasks completed, carrying/walking distance and maximum weight lifted. These indicators show the performance of exoskeleton from the view of effectiveness. In contrast, the industry on its website is more willing to disclose the benefits of its exoskeleton from a company perspective. Four interesting factors which have been listed are increased productivity (EVO, Ottobock Shoulder, Exorise, BackX), reduced breaks (EVO), reduced risk of injury (FLx ErgoSkeleton) and saved money (FLx ErgoSkeleton). These performance indicators are a direct measure of an exoskeleton's effectiveness on a company-level, and are therefore appealing to practitioners. However, it is worth noting that these results are debatable, because they are reported on their respective websites without relevant details and evidence. More proofs are needed to have a deeper insight.

On the other hand, subjective metrics are more often considered in academia than in industry. The company tends to use other indicators on its website to indicate the subjective benefits that its product can achieve. For example, the manufacturer would describe the soft texture and padding (design metric), and large ROM (biomechanical metric) to reflect the comfort of exoskeleton. Similarly, they would use increased productivity (functional metric) to show the usefulness of their exoskeletons. Differently, various assessment scales and interview questions are generally used to evaluate the user feeling or feedback on exoskeleton. Therefore, there are numerous redundant parameters in academia. There is for example an overlap between pain and discomfort when subjectively describing the feeling of exoskeleton during tests. Similar metrics can often be simplified by using a unified term: intention of use and acceptance could be combined, and so could perceived task difficulty and effectiveness. All this points towards a clear need for a concise and comprehensive list of industrial exoskeleton performance metrics that is applicable across different scenarios.

4.4 Testing methods

As shown in figure 5, it is interesting that customized tasks are largely used in lab tests, especially in the evaluation of upper-body and back exoskeletons. As mentioned in section 3.6, those customized tasks, adapted to specific conditions, are a combination of several tasks (lifting, holding, etc.). Laboratory, differently from industry, is a controllable and flexible environment with dedicated facilities. Various types of experiment are able to be performed in this setting, which lead to extensive numbers of customized tests. Furthermore, it is interesting that holding, lifting and overhead work tasks are largely used to test upper-limb exoskeletons both in academia (lab and field tests) and industry. This makes sense, since upper-body exoskeleton is developed for supporting upper body prone to fatigue and injury while executing related industrial tasks. It is reasonable to adopt these tasks to assess the performance of upper-body exoskeleton. Many studies have also proved that exoskeletons play a role in holding (Spada et al., 2017; Spada et al., 2018; Du et al., 2020), lifting (Theurel et al., 2018; Pinho et al., 2020) and overhead tasks (Sylla et al., 2014; Van Engelhoven et al., 2019). Compared with tasks mentioned above, the number of carrying task, which means walking while holding loads (Theurel et al., 2018; Poliero et al., 2020b), is fewer. This may be because the carrying task is more related to the movement of the lower limbs, and is rarely considered and used to test the upper body exoskeleton. In addition, only one paper has reported the stacking and unstacking task used in the evaluation of the Exhauss exoskeletons (Theurel et al., 2018). In this test, the subject needs to lift the box, rotate the torso to 90 degrees, and place the box in the designated place. This testing task is not reported in industrial side, but they present the manual handling task. The manual handling task includes carrying, lifting, lowering, pushing and pulling loads by hands. In lab and field tests, researchers clearly state testing procedures and tasks to assess exoskeletons so that it is easy to identify and classify these tasks. However, manufacturer reports the manual handling task on its website without disclosing details. This makes it difficult to know which specific type of testing task is used in exoskeleton evaluation.

The lifting task occupies the largest proportion of all testing tasks of the back exoskeleton in the academic and industrial field. Loads lifting is the one of the most common activities in the workplace, and the incorrect lifting technologies could expose workers to low back pain. Thus, testers place much emphasis on this task to evaluate the back exoskeleton. Several studies have shown that back exoskeletons may be helpful in reducing spinal muscle fatigue (de Looze et al., 2016) and muscle activities (Yong et al., 2019; Wei et al., 2020; Qu et al., 2021) during loads lifting work. Furthermore, the development of customized tasks for evaluating back exoskeletons is also widely used in the academic field. Customized test methods, applicable to specific environments and easier to implement under laboratory conditions, gains the favor of researchers. Additionally, the number of holding tasks used in academic investigations is same as that found in industrial field. This seems to indicate that academia and industry attach equal importance to them. It is worth mentioning that although walking task used in both in academic and industrial settings, their assessment focuses are different. For the former, this testing task is conducted to investigate whether the back exoskeleton would hinder user's walking. Instead, Cray X is an active exoskeleton and provides walking assistance. Interestingly, there is one commercially available back exoskeleton designed to help wearer sit in a good posture, while the sitting task has not been utilized in the evaluation of the back exoskeleton in academia. As for overhead, bending, handling and carrying tasks, the number of these tasks in academia is less than that in industry. This does not mean that those testing methods are less used in the academic community. Actually, those tasks are usually evaluated as part of customized tasks, rather than a single task. For instance, functional performance tests proposed by Baltrusch et al (Baltrusch et al., 2018), a very famous customized test, includes twelve tasks to provide further insights in exoskeleton's versatility. Instead, there are relatively few studies that only use a single task to evaluate the performance of exoskeleton. It is the diversity of test methods that makes it difficult to compare test results. At the same time, the lack of benchmarks for exoskeleton evaluation has also seriously hindered researchers to compare exoskeletons among studies (De Bock et al., 2022). Therefore, to better compare exoskeletons and accelerate the introduction of exoskeletons into the market, it is necessary to develop a standardized benchmarking methods for occupational exoskeleton.

With regard to lower-body exoskeletons, results found in the literature and industry are much fewer. Lab experiments prefer to adopt walking and squatting task to test lower-body exoskeletons. These methods mainly assess whether the exoskeleton will interfere with the normal movement of the wearer during use. There is a field study that uses standing and sitting methods to evaluate physical load during the use of Chairless Chair 2.0 (Onofrejova, Balazikova et al. 2022). An improvement in the posture of workers when using this exoskeleton has been reported. There are three commercially available lower-body exoskeletons on the market. Among which, two exoskeleton report their positive impact on standing activity. The remaining one claims its advantage of reducing muscle strain reduction around the knee joint while squatting. There are also some results using other methods to evaluate the lower-body exoskeleton, such as lifting, kneeling, stair-climbing tasks, but the number is quite few. At present, the research on lower limb exoskeleton mainly focuses on rehabilitation, while the industrial research is still in its infancy.

Two full-body exoskeletons (Li, Li et al. 2021, Bai, Islam et al. 2022) were found in the literature and one exoskeleton (Guardian XO) was found on the market. These three full-body exoskeletons are all active exoskeletons, and they are designed to help user to lift heavy loads. It is reasonable to use lifting task to test this kind of exoskeleton. In addition to using the lifting task, Bai et.al also adopted the carrying task to evaluate their exoskeleton.

To conclude, customized tasks gain much interest of researchers. Such tasks can be designed according to the needs of designers, so that the performance of exoskeletons can be tested from multiple perspectives. Furthermore, lifting and holding tasks are very common for the assessment of arm-support and back-support exoskeletons. These two types of exoskeletons are both helpful to reduce muscle strain. The only difference is that is that they benefit different regions of the human body. In addition, bending and carrying tasks are more used to evaluate the back-support exoskeleton than the arm-support exoskeleton. However, the overhead task is just the opposite. It is mainly because the assistive region of these exoskeletons are different, resulting in different test methods related to this assistive regions. In the evaluation of the leg-support exoskeleton, there are more research results involving the activities of the lower limbs of the human body, such as stair-climbing, standing, squatting and walking. Although the retrieved results of full-body exoskeleton are quite few, it is shown that lifting and carrying tasks are the most common methods to assess these exoskeletons.

5. Benchmarking framework

5.1 Proposed benchmarking metrics

As mentioned above, a wide range of performance indicators have been proposed by researchers and manufacturers, resulting in numerous overlapped metrics. Additionally, some metrics have been underestimated and ignored. In this section, bringing together metrics used to in the evaluation of industrial exoskeletons from both academia (lab tests and field tests) and industry (specification sheets), together with several metrics proposed by the authors, we have compiled an exhaustive list of benchmarking metrics of exoskeletons for all relevant application domains, as shown in table 2.

Table 2. Benchmarking metrics for occupational exoskeletons

Metrics types	Metrics		Units
Design metrics	Active, semi-active and passive exoskeletons	Weight	kg
		Dimensions (Length × Width × Height)	m
		Applicable body weight	kg
		Applicable body height	m
		Price	€
		Warranty period	year
		Construction material	unitless
		Actuation type	unitless
		Ingress protection rating	unitless
		Operating temperature conditions	°C
		Operating humidity conditions (relative air humidity)	%
		Operating air pressure conditions	hPa
		Sound level	dB
		Attachment points	unitless
		Padding exchangeable/washable/breathable?	Yes/No
		Operator interface	unitless
	Active exoskeletons	Additional apparatus required?	unitless
		Power source type	unitless
		Operating voltage	V
		Operating current	A
		Battery life	h
		Battery charging time	h
Physiological metrics		Metabolic consumption	ml/(kg*min)
		Heart rate	bpm
Functional metric		Productivity booster	%
		Task repetition number	unitless
		Carrying/Walking distance	m
		Task completion time	s
Biomechanical metric		ROM	°
		Setup/Donning/Doffing time	s
		Muscle activity reduction	%
		DOFs	unitless
		Maximum assistive force	N
		Joint moment reduction	%
Subjective metrics		Comfort	unitless
		Usefulness	
		Perceived effort	
		Aesthetics	
		Ease of use	
		Visibility	
		Versatility	
		Perceived quality	
		Perceived safety	

Below, we provide explanations about a selected list of metrics reported in figure 3. Common or self-explanatory metrics have been left out for brevity.

- Price: Euro signs will be used to indicate price instead of a real price, e.g., € means low price (5000 euros), €€ means relatively low price (5000 – 10000 euros), €€€ means medium price (10000 - 20000 euros), €€€€ means relatively high price (20000-40000 euros), €€€€€ means high price (>40000 euros).

- Construction material: The main materials used for the construction of the device (mostly the links).
- Ingress protection rating: Level of sealing effectiveness of electrical enclosures against intrusion from foreign bodies (tools, dirt etc.) and moisture, as described in the standard IEC 60529.
- Sound level: The sound level generated by exoskeleton at maximum torque or speed output.
- Attachment points: The contact part between exoskeleton and human body (shown in the form of pictures).
- Operator interface: The way to operate the device, e.g. buttons, touchscreen, smartphone, voice.
- Productivity booster: The extent to which the exoskeleton can boost users' productivity. The ratio of the number of tasks completed with and without exoskeleton in a period of time.
- Perceived effort: Cognitive feeling of work associated with voluntary actions.
- Aesthetics: To what extent the exoskeleton looks aesthetically pleasing.
- Versatility: To what extent the exoskeleton can adapted to different tasks.
- Perceived quality: the customer's perception of the overall quality of the device.
- Visibility: The degree to which the product attracts people's attention. It is better to design the exoskeleton less visible. User might be unwilling to be seen when wearing the exoskeleton at the workplace (Baltrusch, Houdijk et al. 2021).

These metrics that almost accommodate all the relevant scenarios can be utilized to assess the several aspects of multifaceted performance, allowing a truthful comparison of each device.

5.2 Basic task library

Although there are many industrial exoskeletons on the market and in the lab, there is still no unified standard to guide people how to test them in a standardized way. At present, one of reasons that hinder the comparison of exoskeletons is that their target tasks are complex. Inspired by the Therblig Analysis (Aft 2000), which consists of 18 elemental movements in the workplace, we aim to establish a basic task library that includes common industrial tasks. If complex industrial tasks are a combination of several simple basic tasks, it would be easier to compare each basic task after decomposing complex tasks into basic tasks. To make each basic task comparable, a standardized testing protocol is indispensable. Functional performance tests proposed by S.J. Baltrusch et al (Baltrusch, van Dieen et al. 2018) provide a good base to establish our basic task library. However, our basic library (as shown in appendix 3) contains a total of 20 tasks, such as pushing, pulling, lifting and carrying tasks, accommodating common scenarios and movements involved in industry.

5.3 Required equipment for basic task tests

The equipment required to measure the aforementioned benchmarking metrics consists of the following:

- Weighing scales
- Meter
- Force sensor
- Torque sensor
- Stopwatch
- Multimeter (or separate Voltmeter/Amperemeter)
- Thermometer
- Hygrometer
- Barometer
- Sound level meter
- Gas analysis system
- Pulse oximeter
- EMG device
- Motion capture system
- Force plates
- Angular measurement (e.g. encoder, potentiometer)

- Metronome

We conducted a survey among the EXSKALLERATE consortium members to determine which hardware is available in each of the labs. The results of this survey are summarized in appendix 6. As can be seen in this table, nearly all labs are capable of recording the proposed metrics.

5.4 Benchmarking method

Although the benchmarking metrics we proposed are very important for evaluating the performance of exoskeleton, not every metric can be compared, especially design metrics, whose main function are to provide the user with details of the product. With regard to metrics can be compared, as shown in table 3, we propose a method to evaluate them. IF refers to the importance factor of related metric, which should be determined by relevant individuals before the test, such as users, researchers, industry managers, etc. It is an average value, reflecting the importance of this metric from different aspects. The more significant a performance metric of the exoskeleton is, the higher the value of IF corresponding to this metric. In the European EXSKALLERATE project, we discussed the importance factors of each metric with more than 50 experts from different fields, and eventually obtained the following data.

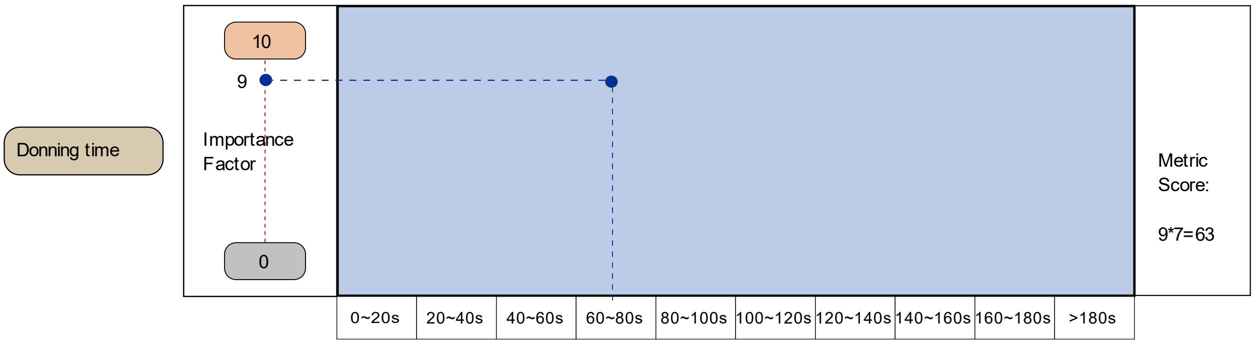
Table 3 the importance factor and score of related metric

Category	Metrics	Importance factor (IF)	Reference score (RS)	Score
Biomechanical benefits	Maximum assistive force	8	Appendix 5	Score A
	Muscle activity reduction	10		
	Joint moment reduction	10		
Physiological benefits	Metabolic consumption	3	Appendix 5	Score B
	Heart rate	3		
Ease of use	Donning/Doffing/Setup time	9	Appendix 5	Score C
	Weight	9	Appendix 5	
	ROM	10	Appendix 4	
Productivity benefits	Productivity booster	2	Appendix 5	Score D
Broader acceptance	Comfort	10	Appendix 4	Score E
	Usefulness	8		
	Perceived effort	7		
	Aesthetics	3		
	Visibility	3		
	Versatility	10		
	Perceived quality	5		
	Perceived safety	8		

Reference score can be obtained by referring to the reference datasheet or Borg CR scale after completing every basic task. It must be mentioned that the reference score of broader acceptance metrics and ROM is obtained by interviewing subjects with the Borg CR scale (appendix 4). The reference datasheet of other metrics is reported in appendix 5. If multiple muscles are involved, each should refer to the reference datasheet to obtain its score. Therefore, based on the importance factor and reference score, the metric score can be calculated using the below formula (1). An example of calculating the metric score of donning time is presented in figure 6.

$$Metric\ score = IF * RS \tag{1}$$

Figure 6. Example of calculating the mertic score of donning time



The score of each metric category can be calculated by using the following weight average equation:

$$Score = \frac{\sum_1^m Metric\ score}{\sum_1^m IF(m)} \tag{2}$$

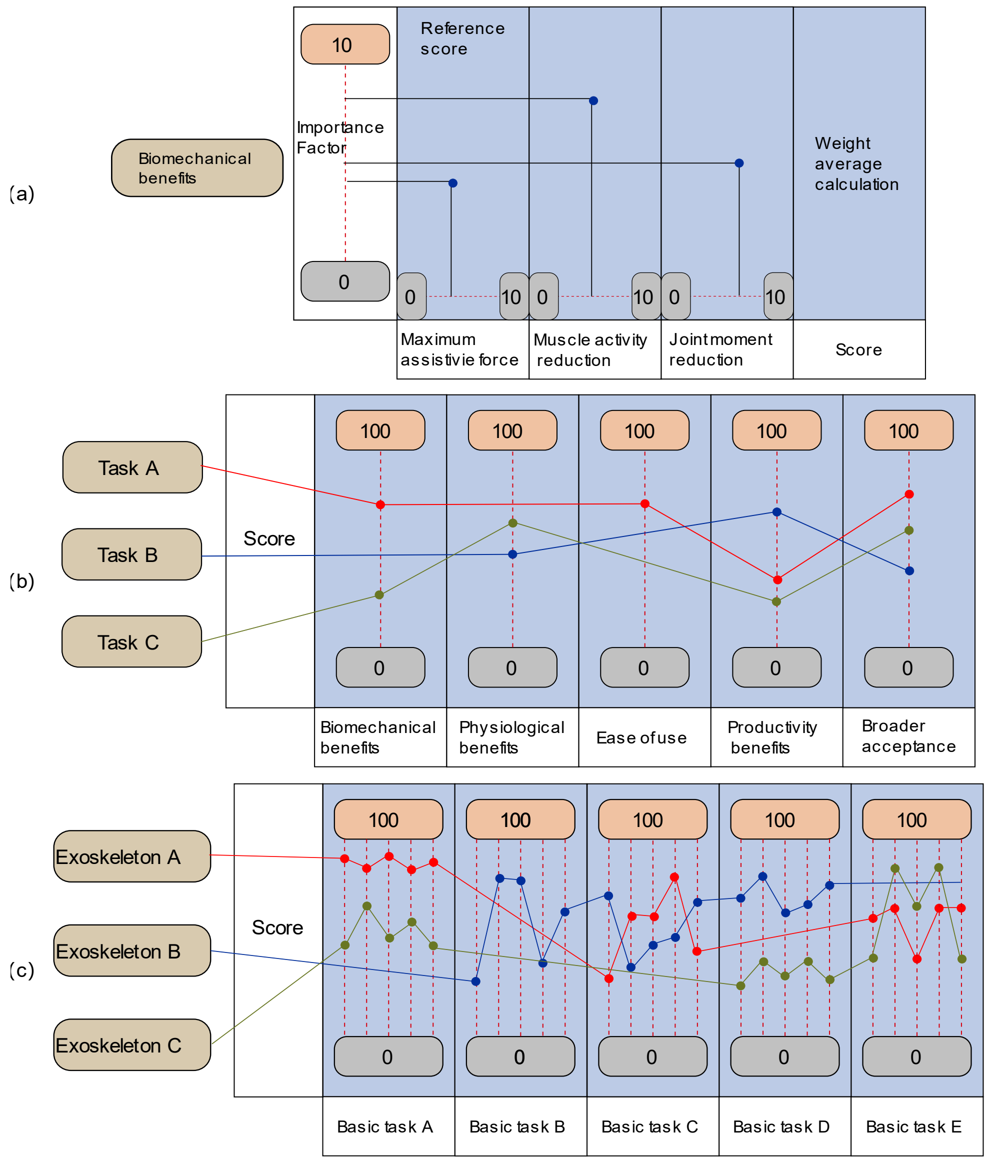
Therefore, each basic task proposed in our library can be assessed by using metrics (table 3) and methods (equation 1-2) mentioned above.

5.5 Benchmarking framework

The holistic framework of our benchmarking methods is presented in figure 7. Firstly, we are supposed to communicate with relevant practitioners and experts to determine which basic tasks are involved and necessary in real working environment. Then, following the testing protocol in appendix 3, tester can assess the basic task by referring to appendix 4 and 5 and obtain reference scores. Combined with formulas we proposed, the category score (score A-E) can be got in the assessment of each basic task, as shown in figure 7 (a). Furthermore, multiple tasks (figure 7 (b)) can be evaluated using the same method. Finally, testing results are described in figure 7 (c), with which customers can easily compare and select the appropriate product according to their needs.

This benchmarking framework provides a clear method for intuitive comparison of exoskeletons. It enables exoskeleton manufacturers to visually disclose the performance of their products, rather than providing users with numerous and cumbersome parameters. In addition, another benefit is that individuals can choose exoskeletons based on the target usage.

Figure 7. Benchmarking framework



5.6 Example of benchmarking framework implementation

We take Laevo and Cray-X for an example to specify the benchmarking process, as shown in figure 8.

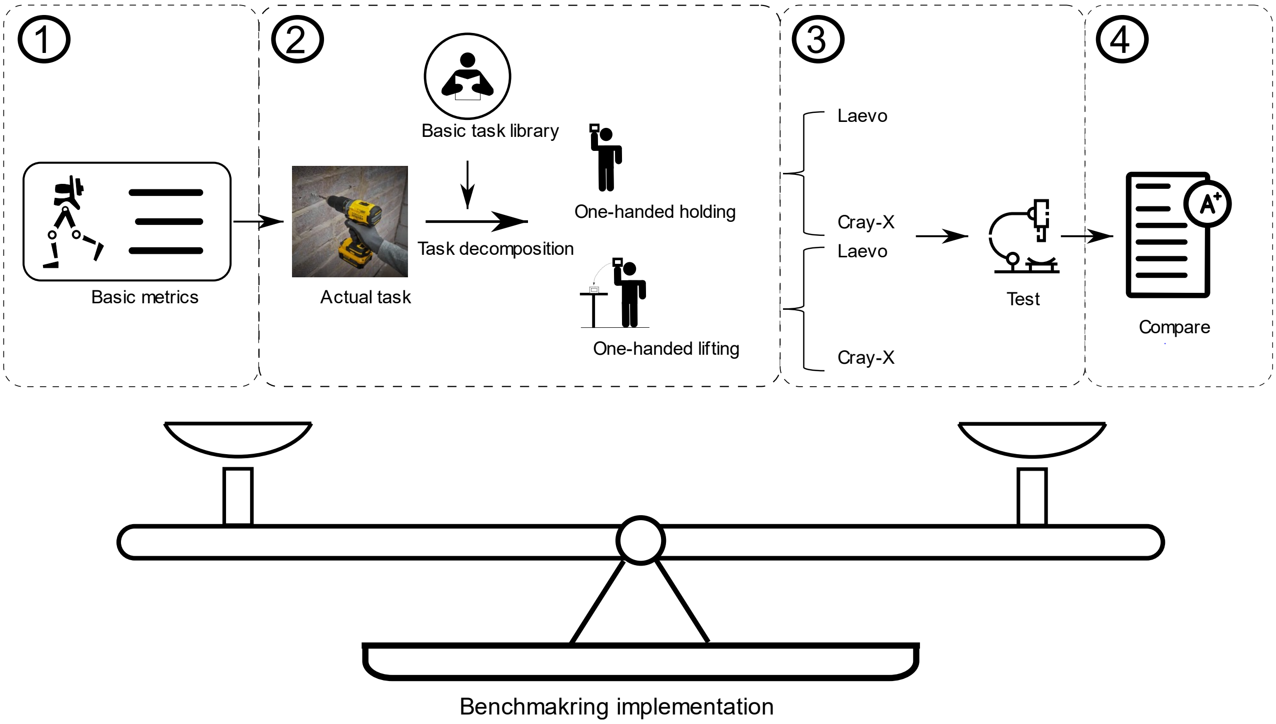


Figure 8. Benchmarking implementation process.

Step 1: Basic information of design parameters is supposed to be disclosed by manufacturers. This information can help users quickly view and compare products, such as price and applicable weight. Below, we provide the information collected on the websites of Laevo (passive exoskeleton) and Cray-X (active exoskeleton). It can be seen that some information is missing, therefore we write 'to be provided' in the corresponding column.

Table 4. Design parameters of Laevo.

Metrics types	Metrics	Result
Design	Dimensions (Length × Width × Height)	To be provided
	Applicable body weight	To be provided
	Applicable body height	150cm and 180cm
	Price	€
	Warranty period	To be provided
	Construction material	Steel, plastic
	Actuation type	Passive
	Ingress protection rating	To be provided
	Operating temperature conditions	To be provided
	Operating humidity conditions (relative air humidity)	To be provided
	Operating air pressure conditions	To be provided
	Sound level	To be provided
	Attachment points	Figure 9
	Padding exchangeable/washable/breathable?	Yes
	Operator interface	NA
	Additional apparatus required?	No

Figure 9. Attachment points between Laevo and wearer.

Table 5. Design parameters of Cray-X.

Metrics types	Metrics	Value
Design	Dimensions (Length × Width × Height)	TBD
	Applicable body weight	TBD
	Applicable body height	150cm and 180cm
	Price	€
	Warranty period	TBD
	Construction material	Steel, plastic,
	Actuation type	Active
	Power source type	Electric
	Operating voltage	40 V
	Operating current	TBD
	Battery life	TBD
	Battery charging time	TBD
	Ingress protection rating	IP54
	Operating temperature conditions	TBD
	Operating humidity conditions (relative air humidity)	TBD
	Operating air pressure conditions	TBD
	Sound level	TBD
	Attachment points	See below
	Padding exchangeable/washable/breathable?	Yes
	Operator interface	Button
	Additional apparatus required?	No

Figure 10. Attachment points between Cray-X and wearer.

Step 2: Look through the basic task library and decompose the actual task into one or several basic tasks. We take the screwing task for an example, it includes two basic task (one-handed holding and one-handed lifting task).

Step 3: For each basic task (one-handed holding and one-handed lifting task), it is supposed to complete the tests of the following metrics. It is worth noting that the following listed data are only used to demonstrate the implementation of the benchmark framework and do not represent actual experimental results.

Table 6. Hypothetical experimental results of Laevo and Cray-X

Category	Metrics	Results of Laevo	Results of Cray-X
Biomechanical benefits	Maximum assistive force	50N	300N
	Muscle activity reduction	15%	28%
	Joint moment reduction	10%	20%
Physiological benefits	Metabolic consumption reduction	5%	2%
	Heart rate reduction	3%	2%
Ease of use	Donning/Doffing/Setup time	35s	50s
	Weight	2.8 kg	7 kg
	ROM	8	8
Productivity benefits	Productivity booster	13%	15%
Broader acceptance	Comfort	8	9
	Usefulness	6	8
	Perceived effort	8	9
	Aesthetics	8	8
	Visibility	3	2
	Versatility	7	6
	Perceived quality	8	8
	Perceived safety	7	5

Step 4: Based on results listed in table 6, the reference score (RS) of Laevo and Cray-X can be got by referring appendix 5. Since the importance factor and reference score have been obtained, the equation (1-2) listed in section 5.4 can be used to calculate final score for each exoskeleton. The performance of different aspects of Laevo and Cray-X can be evaluated and compared.

Table 7. Performance of Laevo and Cray-x (based on hypothetical experimental results)

Category	Metrics	Importance factor (IF)	RS of Laevo	RS of Cray-X	Score of Laevo	Score of Cray-X
Biomechanical benefits	Maximum assistive force	8	3	10	1.9	5
	Muscle activity reduction	10	2	3		
	Joint moment reduction	10	1	3		
Physiological benefits	Metabolic consumption	3	1	1	3	3
	Heart rate	3	1	1		
Ease of use	Donning/Doffing/Setup time	9	9	8	8.3	6.4
	Weight	9	8	3		
	ROM	10	8	8		
Productivity benefits	Productivity booster	2	2	2	2	2
Broader acceptance	Comfort	10	8	9	7.1	7.2
	Usefulness	8	6	8		
	Perceived effort	7	8	9		
	Aesthetics	3	8	8		
	Visibility	3	3	2		
	Versatility	10	7	6		
	Perceived quality	5	8	8		
	Perceived safety	8	7	5		

6. Conclusion

In this report, we analyze and compare the performance metrics obtained in literature and industry. We found the research and development of industrial exoskeleton mainly focused on upper-body exoskeleton and back exoskeleton. The investigation of lower-body and full-body exoskeleton for industrial purpose is considerably fewer. In addition, both the academia and industry evaluate the weight and dimension highly in terms of design metrics. The only difference is that the industry pays more attention to customer-centered metrics than the academia. It makes sense that those metrics appeal customers than technical parameters. Furthermore, in the biomechanical metrics, the muscle activity and assistive torque/force gain much interest of people from both fields. With regard to physiological, functional and subjective metrics, the results from industrial circles are much fewer than that from academic side, which can be explained by intellectual property concerns, difficulty of implementation and a lack of interest from industrial side. On the other hand, the lack of field tests is also a barrier for the development of suitable benchmarking metrics.

As for testing methods, despite researchers tend to adopt customized testing task to assess the performance of the exoskeleton, the adoption of other testing methods, such as lifting, holding and overhead tasks, show great consistency in these two fields. Additionally, numerous and various testing methods are applied in the laboratory, especially customized testing tasks. Different types of exoskeletons indeed require different test methods, and even within the same type of exoskeletons test methods may differ depending on the task they are designed for. This makes it difficult to define a general benchmarking method for occupational exoskeletons. Such methods would however provide a great impetus to the establishment of exoskeleton standards. Through our investigation and research, we classify common activities of workers to inform the development of relevant, unified testing methods (basic task library). On this basis, we propose a benchmarking framework which can contribute to comparing exoskeleton across different scenarios and creating benchmarking guidelines facilitating standardization and market adoption. Finally, our analysis of literature shows that exoskeleton suppliers publish test results, but the test methods by which they are acquired are poorly reported. More transparency would be needed in this regard.

7. References

Aft, L. S. (2000). Work measurement and methods improvement, John Wiley & Sons.

Amandels, S., et al. (2019). Introduction and Testing of a Passive Exoskeleton in an Industrial Working Environment. Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)**:** 387-392.

Babič, J., et al. (2021). "Challenges and solutions for application and wider adoption of wearable robots." Wearable Technologies **2**.

Backx https://www.suitx.com/backx.

Bai, S., et al. (2022). "User-centered development and performance assessment of a modular full-body exoskeleton (AXO-SUIT)." Biomimetic Intelligence and Robotics **2**(2): 100032.

Baltrusch, S. J., et al. (2021). "Passive Trunk Exoskeleton Acceptability and Effects on Self-efficacy in Employees with Low-Back Pain: A Mixed Method Approach." Journal of Occupational Rehabilitation **31**(1): 129-141.

Baltrusch, S. J., et al. (2019). The Effect of a Passive Trunk Exoskeleton on Functional Performance and Metabolic Costs. Wearable Robotics: Challenges and Trends**:** 229-233.

Baltrusch, S. J., et al. (2019). "SPEXOR passive spinal exoskeleton decreases metabolic cost during symmetric repetitive lifting." European Journal of Applied Physiology **120**(2): 401-412.

Baltrusch, S. J., et al. (2018). "The effect of a passive trunk exoskeleton on functional performance in healthy individuals." Appl Ergon **72**: 94-106.

Baltrusch, S. J., et al. (2020). "Testing an Exoskeleton That Helps Workers With Low-Back Pain: Less Discomfort With the Passive SPEXOR Trunk Device." IEEE Robotics & Automation Magazine **27**(1): 66-76.

Baxter, M. (2018). Product design, CRC press.

Blanco, A., et al. (2022). "The Effect of an Active Upper-Limb Exoskeleton on Metabolic Parameters and Muscle Activity During a Repetitive Industrial Task." Ieee Access **10**: 16479-16488.

Bosch, T., et al. (2016). "The effects of a passive exoskeleton on muscle activity, discomfort and endurance time in forward bending work." Appl Ergon **54**: 212-217.

ChairlessChair "Chairless Chair " https://www.noonee.com/.

Crea, S., et al. (2021). "Occupational exoskeletons: A roadmap toward large-scale adoption. Methodology and challenges of bringing exoskeletons to workplaces." Wearable Technologies **2**.

De Bock, S., et al. (2021). "Passive Shoulder Exoskeletons: More Effective in the Lab Than in the Field?" Ieee Transactions on Neural Systems and Rehabilitation Engineering **29**: 173-183.

De Kok, J., et al. (2019). Work-related musculoskeletal disorders: Prevalence, costs and demographics in the EU. European Agency for Safety and Health at Work.

de Looze, M. P., et al. (2016). "Exoskeletons for industrial application and their potential effects on physical work load." Ergonomics **59**(5): 671-681.

Di Natali, C., et al. (2021). "Equivalent Weight: Connecting Exoskeleton Effectiveness with Ergonomic Risk during Manual Material Handling." International Journal of Environmental Research and Public Health **18**(5).

Diana, S., et al. (2022). "Exploratory Field Testing of Passive Exoskeletons in Several Manufacturing Environments Perceived Usability and User Acceptance." Iise Transactions on Occupational Ergonomics & Human Factors: 1-12.

Du, Z. H., et al. (2020). Mechanical Design and Preliminary Performance Evaluation of a Passive Arm-support Exoskeleton. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Electr Network.

Elprama, S. A., et al. (2022). "An industrial exoskeleton user acceptance framework based on a literature review of empirical studies." Appl Ergon **100**: 103615.

ExoskeletonReport. from https://exoskeletonreport.com/product-category/exoskeleton-catalog/industrial/.

Frisoli, A. (2019). Wearable Robots. Encyclopedia of Robotics. M. H. Ang, O. Khatib and B. Siciliano. Berlin, Heidelberg, Springer Berlin Heidelberg: 1-8.

Gonsalves, N. J., et al. (2021). "ASSESSMENT OF A PASSIVE WEARABLE ROBOT FOR REDUCING LOW BACK DISORDERS DURING REBAR WORK." Journal of Information Technology in Construction **26**: 936-952.

Grazi, L., et al. (2020). "Design and Experimental Evaluation of a Semi-Passive Upper-Limb Exoskeleton for Workers With Motorized Tuning of Assistance." Ieee Transactions on Neural Systems and Rehabilitation Engineering **28**(10): 2276-2285.

Groos, S., et al. (2019). Determination of the Subjective Strain Experiences During Assembly Activities Using the Exoskeleton "Chairless Chair", Washington, DC.

GuardianXO https://www.sarcos.com/products/guardian-xo-powered-exoskeleton/.

Han, B., et al. (2019). Mechanical Framework Design with Experimental Verification of a Wearable Exoskeleton Chair.

Hensel, R. and M. Keil (2019). "Subjective Evaluation of a Passive Industrial Exoskeleton for Lower-back Support: A Field Study in the Automotive Sector." IISE Transactions on Occupational Ergonomics and Human Factors **7**(3-4): 213-221.

Hondzinski, J. M., et al. (2018). "Effects of exoskeleton use on movement kinematics during performance of common work tasks: A case study." Work-a Journal of Prevention Assessment & Rehabilitation **61**(4): 575-588.

Huysamen, K., et al. (2018). "Evaluation of a passive exoskeleton for static upper limb activities." Applied Ergonomics **70**: 148-155.

Huysamen, K., et al. (2018). "Assessment of an active industrial exoskeleton to aid dynamic lifting and lowering manual handling tasks." Applied Ergonomics **68**: 125-131.

Hwang, J., et al. (2021). "Effects of passive back-support exoskeletons on physical demands and usability during patient transfer tasks." Applied Ergonomics **93**: 103373.

Ji, X. Y., et al. (2020). "SIAT-WEXv2: A Wearable Exoskeleton for Reducing Lumbar Load during Lifting Tasks." Complexity **2020**.

Kim, H. K., et al. (2021). "Analysis of Active Back-Support Exoskeleton During Manual Load-Lifting Tasks." Journal of Medical and Biological Engineering **41**(5): 704-714.

Kim, Y. G., et al. (2020). "A voice activated bi-articular exosuit for upper limb assistance during lifting tasks." Robotics and Computer-Integrated Manufacturing **66**.

Ko, H. K., et al. (2018). "Waist-assistive exoskeleton powered by a singular actuation mechanism for prevention of back-injury." Robotics and Autonomous Systems **107**: 1-9.

Kong, Y. K., et al. (2021). "Guidelines for Working Heights of the Lower-Limb Exoskeleton (CEX) Based on Ergonomic Evaluations." International Journal of Environmental Research and Public Health **18**(10).

Kozinc, Z., et al. (2020). "Reliability of a battery of tests for functional evaluation of trunk exoskeletons." Applied Ergonomics **86**.

Kozinc, Z., et al. (2021). "Short-Term Effects of a Passive Spinal Exoskeleton on Functional Performance, Discomfort and User Satisfaction in Patients with Low Back Pain." Journal of Occupational Rehabilitation **31**(1): 142-152.

Latella, C., et al. "Analysis of Human Whole-Body Joint Torques During Overhead Work With a Passive Exoskeleton." Ieee Transactions on Human-Machine Systems.

Li, X., et al. (2021). "Method, Design, and Evaluation of an Exoskeleton for Lifting a Load In Situ." Applied Bionics and Biomechanics **2021**.

Liu, C., et al. (2020). "Functional Evaluation of a Force Sensor-Controlled Upper-Limb Power-Assisted Exoskeleton with High Backdrivability." Sensors **20**(21).

Liu, S. L., et al. (2018). "Solving the surgeon ergonomic crisis with surgical exosuit." Surgical Endoscopy and Other Interventional Techniques **32**(1): 236-244.

Lorenzo, G., et al. (2019). "Towards methodology and metrics for assessing lumbar exoskeletons in industrial applications." MetroInd4.0&IoT: 400-404.

Luger, T., et al. (2021). "Using a Back Exoskeleton During Industrial and Functional Tasks-Effects on Muscle Activity, Posture, Performance, Usability, and Wearer Discomfort in a Laboratory Trial."

Madinei, S., et al. (2020). "Biomechanical Evaluation of Passive Back-Support Exoskeletons in a Precision Manual Assembly Task: "Expected" Effects on Trunk Muscle Activity, Perceived Exertion, and Task Performance." Human Factors **62**(3): 441-457.

Madinei, S., et al. (2021). "Effects of back-support exoskeleton use on trunk neuromuscular control during repetitive lifting: A dynamical systems analysis." Journal of Biomechanics **123**.

Moltedo, M., et al. (2019). "Variable stiffness ankle actuator for use in robotic-assisted walking: Control strategy and experimental characterization." Mechanism and Machine Theory **134**: 604-624.

Moyon, A., et al. (2019). "Development of an Acceptance Model for Occupational Exoskeletons and Application for a Passive Upper Limb Device." Iise Transactions on Occupational Ergonomics & Human Factors **7**(3): 291-301.

Naf, M. B., et al. (2018). "Passive Back Support Exoskeleton Improves Range of Motion Using Flexible Beams." Frontiers in Robotics and AI **5**.

Ogunseiju, O., et al. (2021). "Evaluation of postural-assist exoskeleton for manual material handling." Engineering Construction and Architectural Management **29**(3): 1358-1375.

Onofrejova, D., et al. (2022). "Ergonomic Assessment of Physical Load in Slovak Industry Using Wearable Technologies." Applied Sciences-Basel **12**(7).

Ottobock "Ottobock Shoulder." https://ottobockexoskeletons.com/automotive-exoskeletons/?lang=en

Pinho, J. P., et al. (2020). A comparison between three commercially available exoskeletons in the automotive industry: an electromyographic pilot study, New York City, NY.

Poliero, T., et al. (2020). "Applicability of an Active Back-Support Exoskeleton to Carrying Activities." Front Robot AI **7**: 579963.

Qu, X. S., et al. (2021). "Effects of an industrial passive assistive exoskeleton on muscle activity, oxygen consumption and subjective responses during lifting tasks." Plos One **16**(1): e0245629.

Ranaweera, R., et al. (2018). "Design and Evaluation of Passively Powered Knee Exoskeleton (PPKE) for Squat Lifting." 55-59.

Schnieders, T. M. and T. Richard (2018). Ranking Importance of Exoskeleton Design Aspects. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications Sage CA: Los Angeles, CA.

Smets, M. (2019). "A Field Evaluation of Arm-Support Exoskeletons for Overhead Work Applications in Automotive Assembly." IISE Transactions on Occupational Ergonomics and Human Factors **7**(3-4): 192-198.

Spada, S., et al. (2018). Passive Upper Limb Exoskeletons: An Experimental Campaign with Workers, Florence, ITALY.

Spada, S., et al. (2017). Investigation into the applicability of a passive upper-limb exoskeleton in automotive industry, Modena, ITALY.

Sylla, N., et al. (2014). "Ergonomic contribution of ABLE exoskeleton in automotive industry." International Journal of Industrial Ergonomics **44**(4): 475-481.

Theurel, J., et al. (2018). "Physiological consequences of using an upper limb exoskeleton during manual handling tasks." Applied Ergonomics **67**: 211-217.

Torricelli, D., et al. (2020). "Benchmarking Wearable Robots: Challenges and Recommendations From Functional, User Experience, and Methodological Perspectives." Front Robot AI **7**: 561774.

Toxiri, S., et al. (2019). Towards Standard Specifications for Back-Support Exoskeletons. Wearable Robotics: Challenges and Trends**:** 219-223.

Tu, Y., et al. (2022). "Design and Experimental Evaluation of a Lower-Limb Exoskeleton for Assisting Workers With Motorized Tuning of Squat Heights." Ieee Transactions on Neural Systems and Rehabilitation Engineering **30**: 184-193.

Van Engelhoven, L., et al. (2018). "Evaluation of an adjustable support shoulder exoskeleton on static and dynamic overhead tasks." Proceedings of the Human Factors and Ergonomics Society Annual Meeting **62**(1): 804-808.

Van Engelhoven, L., et al. (2019). "Experimental Evaluation of a Shoulder-Support Exoskeleton for Overhead Work: Influences of Peak Torque Amplitude, Task, and Tool Mass." Iise Transactions on Occupational Ergonomics & Human Factors **7**(3): 250-263.

Verstraten, T. and D. Lefeber (2021). "Compliant Actuation for Wearable Robotics." Novel Bioinspired Actuator Designs for Robotics.

Wang, H. M., et al. (2021). "Evaluation of a Passive Upper-Limb Exoskeleton Applied to Assist Farming Activities in Fruit Orchards." Applied Sciences-Basel **11**(2).

Wang, Z., et al. (2021). "A Semi-active Exoskeleton Based on EMGs Reduces Muscle Fatigue When Squatting." Frontiers in Neurorobotics **15**.

Wei, W., et al. (2020). "A Hip Active Assisted Exoskeleton That Assists the Semi-Squat Lifting." Applied Sciences-Basel **10**(7).

Weston, E. B., et al. (2022). "A physiological and biomechanical investigation of three passive upper-extremity exoskeletons during simulated overhead work." Ergonomics **65**(1): 105-117.

Yamada, Y., et al. (2020). "TasKi: Overhead Work Assistance Device with Passive Gravity Compensation Mechanism." Journal of Robotics and Mechatronics **32**(1): 138-148.

Yin, P., et al. (2021). "The Effect of Mobile Wearable Waist Assist Robot on Lower Back Pain during Lifting and Handling Tasks." Mobile Networks & Applications **26**(3): 988-996.

Yin, P., et al. (2021). "Development of an ergonomic wearable robotic device for assisting manual workers." International Journal of Advanced Robotic Systems **18**(5).

Yong, X., et al. (2019). "Ergonomic Mechanical Design and Assessment of a Waist Assist Exoskeleton for Reducing Lumbar Loads During Lifting Task." Micromachines **10**(7).

Zhang, H. H., et al. (2016). "Design and Preliminary Evaluation of a Passive Spine Exoskeleton." Journal of Medical Devices-Transactions of the Asme **10**(1).

APPENDIX 1

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Upper-body	Levitate AIRFRAME	Passive	Functional performance tests (customized tasks), holding tasks	Lab	-	-	-	Task completion time	Discomfort, fatigue, pain, user impression	(Liu, Hemming et al. 2018)
			Holding tasks	Field						
	Personal device	Active	Manipulating tasks (customized tasks), lifting, holding tasks	Lab	Weight	-	Muscle activity, joint angle	-	-	(Kim, Little et al. 2020)
	EksoVest, AIRFRAME, ShoulderX	Passive	Overhead tasks	Lab	-	tissue oxygenation	Joint torque	-	Discomfort	(Weston, Alizadeh et al. 2022)
	PUES	Active	Holding, lifting tasks	Lab	Weight	-	Muscle activity	-	-	(Yin, Yang et al. 2021)
	TasKi	Passive	Overhead tasks, ROM tasks (customized tasks)	Lab	Dimension	-	ROM, muscle activity, assistive force	-	Ease of movement, wear feeling, fatigue	(Yamada, Arakawa et al. 2020)
	ShoulderX	Passive	Overhead tasks	Lab	-	-	Assistive torque, muscle activity	-	RPE, preferred support level	(Van Engelhoven, Poon et al. 2019)
	PULE	Passive	Simulated thinning and pesticide spraying tasks (holding and lifting tasks)	Lab	Weight	-	Joint angle, assistive torque, muscle activity	-	RPE, LPP	(Wang, Le et al. 2021)
			Fruit thinning and pesticide spraying tasks (holding and lifting tasks)	Field						
	EXHAUSS	Passive	Stacking-unstacking, lifting, carrying tasks	Lab	Weight	Heart rate	ROM, muscle activity, the displacement of centre of pressure	Task completion time	RPE	(Theurel, Desbrosses et al. 2018)
	ABLE	Active	Overhead tasks	Lab	-	-	Ground reaction force, joint torque, joint angle, DOF	Task completion time	-	(Sylla, Bonnet et al. 2014)

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Upper-body	Levitate AIRFRAME	Passive	Precision tasks (customized tasks), holding, lifting tasks	Lab	-	-	-	Task completion time, the number of tasks	RPE	(Spada, Ghibaudo et al. 2017)
	Mate	Passive	Precision tasks (customized tasks), holding, lifting tasks	Lab	-	-	-	Task completion time, the number of tasks	RPE	(Spada, Ghibaudo et al. 2018)
	ShoulderX, Mate, Paexo	Passive	Three assembly tasks (overhead, holding tasks)	Field	-	-	Muscle activity	-	-	(Pinho, Taira et al. 2020)
	Skelex	Passive	Simulated finishing tasks (overhead tasks)	Lab	-	Heart rate	Donning and doffing time	Task completion time, the number of tasks	Ease of use, RPE, satisfaction, ease of learning, utility, safety, efficiency, comfort	(Moyon, Poirson et al. 2019)
			Finishing tasks (overhead tasks)	Field						
	Personal device	Active	Carrying tasks	Lab	Weight, dimension	-	DOF,joint angle, muscle activity, assistive force	-	-	(Liu, Liang et al. 2020)
	Robo-Mate	Passive	Overhead tasks	Lab	Weight	-	Muscle activity	-	RPE, LPP, usability	(Huysamen, Bosch et al. 2018)
	Personal device	Passive	Overhead tasks	Lab	-	Heart rate, blood oxygenation	Muscle activity	-	RPE	(Hondzinski, Ikuma et al. 2018)
	H-PULSE	Semi-Active	Overhead tasks	Lab	Weight	Heart rate	DOF, ROM, assistive torque, muscle activity	-	RPE	(Grazi, Trigili et al. 2020)
	ShoulderX, Skelex	Passive	Lifting, holding tasks	Lab	Weight	Heart rate	Assistive torque, muscle activity	-	RPE, discomfort, usability	(De Bock, Ghillebert et al. 2021)
			Order picking tasks (lifting, holding tasks)	Field						
	ShoulderX	Passive	Overhead tasks	Lab	-	-	Assistive torque, muscle activity	-	Preferred support level	(Van Engelhoven, Poon et al. 2018)
	EksoVest™	Passive	Overhead assembly tasks	Field	-	-	-	Usage patterns	Discomfort	(Smets 2019)
	PAEXO	Passive	Overhead tasks	Lab	Weight	-	Joint torque, assistive torque	-	-	(Latella, Tirupachuri et al.)

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Upper-body	Personal device	Passive	Holding tasks	Lab	-	-	Muscle activity	-	-	(Du, Yan et al. 2020)
	Personal device	Active	Overhead tasks	Lab	-	Heart rate, heart rate variability, inhaled O2, cardiorespiratory response	DOF, ROM, assistive torque, muscle activity	-	-	(Blanco, Catalan et al. 2022)
	Levitate Airframe™ and EksoVest™	Passive	17 tasks (customized tasks)	Field	Weight	-	ROM	-	Safety, usability, comfort	(Diana, McNamara. et al. 2022)
Back	Laevo	Passive	"Functional performance tests	Lab	-	-	-	Max weight lifted, distance, task completion time	Perceived task difficulty, discomfort, user impression	(Baltrusch, van Dieen et al. 2018)
	Laevo	Passive	Simulated industrial tasks (customized tasks)	Lab	Weight	Heart rate	Joint angle, muscle activity	Task completion time	Perceived task difficulty, usability , comfort	(Luger, Bär et al. 2021)
	Laevo	Passive	Functional performance tests (customized tasks)	Lab	-	Energy expenditure	-	-	Perceived task difficulty	(Baltrusch, van Dieën et al. 2019)
	SPEXOR	Passive	Functional performance tests (customized tasks)	In-field	-	-	-	-	Self-efficacy, acceptability	(Baltrusch, Houdijk et al. 2021)
	Personal device	Passive	Bending tasks	Lab	-	-	Muscle activity , joint torque	-	-	(Zhang, Kadrolkar et al. 2016)
	Personal device	Active	Lifting tasks	Lab	Weight	-	Donning time, joint angle, angular velocity, acceleration, muscle activity	-	-	(Yong, Yan et al. 2019)

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Back	MeBot-EXO	Active	Lifting tasks	Lab	Weight	Inhaled O2, exhaled CO2	Joint torque, assistive torque, joint angle, angular velocity, muscle activity	-	-	(Wei, Zha et al. 2020)
	MWWAR	Active	Lifting, handling, holding tasks	Lab	Weight	-	Angular velocity, center of gravity displacement, muscle activity	-	-	(Yin, Yang et al. 2021)
	IPAE	Passive	Lifting tasks	Lab	-	Inhaled O2	Donning time, muscle activity	-	RPE, LPP, usability	(Qu, Qu et al. 2021)
	SPEXOR	Passive	ROM tasks (customized tasks), Functional performance tests (customized tasks)	Lab	-	-	ROM, assistive torque,	-	Users impression, perceived task difficulty, discomfort	(Naf, Koopman et al. 2018)
	BackXTM AC and LaevoTM V2.5	Passive	Lifting tasks	Lab	-	-	ROM, angular velocity	-	-	(Madinei, Kim et al. 2021)
	BackXTM AC and LaevoTM V2.5	Passive	Precision manual assembly tasks (customized tasks)	Lab	-	-	Muscle activity	Task completion time	RPE	(Madinei, Alemi et al. 2020)
	SPEXOR	Passive	Functional performance tests (customized tasks)	Lab	-	-	ROM, assistive torque	The number of tasks , distance, task completion time	Perceived task difficulty, discomfort, user impression	(Kozinc, Baltrusch et al. 2021)
	H-WEX	Active	Walking, lifting tasks	Lab	Weight, dimension, operating voltage	-	Assistive torque, DOF, joint angle, angular velocity, muscle activity	-	Usability	(Ko, Lee et al. 2018)
	Personal device	Active	Mobility tests (customized tasks), lifting tasks	Lab	Weight	-	Assistive force, muscle activity	Task completion time	RPE	(Kim, Hussain et al. 2021)
	SIAT-WEXv2	Active	Lifting tasks	Lab	Weight	-	Assistive torque, muscle activity	-	-	(Ji, Wang et al. 2020)

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Back	FLx ErgoSkeleton, V22 ErgoSkeleton, Laevo V2.5	Passive	Simulated patient transferring tasks (customized tasks)	Lab	Weight, dimension	-	ROM, hand pull force, muscle activity	-	Usability	(Hwang, Yerriboina et al. 2021)
	Robo-Mate	Active	Lifting tasks	Lab	-	-	Assistive torque, setup time, muscle activity	-	Contact pressure, LPP, RPE, usability	(Huysamen, de Looze et al. 2018)
	XoTrunk	Active	Holding tasks	Lab	Weight	-	Assistive torque, muscle activity	-	-	(Di Natali, Chini et al. 2021)
	SPEXOR	Passive	Functional performance tests (customized tasks)	Lab	-	-	Assistive torque	The number of tasks, distance, task completion time	Discomfort, perceived task difficulty, users impression	(Baltrusch, van Dieen et al. 2020)
	SPEXOR	Passive	Lifting tasks	Lab	-	Inhaled O2, exhaled CO2	Assistive torque, muscle activity, joint work, joint angle	-	-	(Baltrusch, van Dieen et al. 2019)
	FLx ErgoSkeletonTM	Passive	Handling tasks	Lab	Weight, dimension	-	ROM	Task completion time	Discomfort, pain	(Ogunseiju, Olayiwola et al. 2021)
	XoTrunk	Active	Lifting, carrying tasks	Lab	Weight	-	Assistive torque, ROM, muscle activity	-	Discomfort	(Poliero, Lazzaroni et al. 2020)
	Laevo	Passive	Multitask on the shop floor (customized tasks)	Field	-	-	ROM, muscle activity	-	Discomfort, users impression	(Amandels, Eyndt et al. 2019)
	Laevo	Passive	Bending, holding tasks	Lab	-	-	ROM, muscle activity	Task completion time	Discomfort	(Bosch, van Eck et al. 2016)
	Laevo	Passive	Bending, handling tasks	Field	-	-	-	-	Discomfort , usability, acceptance	(Hensel and Keil 2019)

Table 8. (continued) Results filtered from selected articles

Types	Names	Actuations	Tasks	Tests	Metrics					References
					Design	Physiological	Biomechanical	Functional	Subjective	
Back	SPEXOR	Passive	Functional performance tests (customized tasks)	Lab	Weight	-	-	The number of tasks, distance, completion time	Discomfort, perceived task difficulty	(Kozinc, Baltrusch et al. 2020)
	BackX™ S	Passive	Simulated rebar tasks (customized tasks)	Lab	Weight	-	Assistive force, muscle activity	Task completion time	Discomfort	(Gonsalves, Ogunseiju et al. 2021)
	Laevo™ v2.5 and SuitX™ backX™ AC	Passive	17 tasks (customized tasks)	Field	Weight	-	ROM	-	Safety, usability, comfort	(Diana, McNamara. et al. 2022)
Lower-body	Personal device	Semi-Active	Walking, squatting tasks	Lab	Weight, appliable height	Inhaled O2, exhaled CO2	ROM, DOF, foot pressure, muscle activity	-	-	(Wang, Wu et al. 2021)
	PPKE	Passive	Squatting tasks	Lab	Weight	-	Assistive force, ROM, muscle activity	-	-	(Ranaweera, Gopura et al. 2018)
	CEX	Passive	Simulated harvesting tasks (customized tasks)	Lab	Weight, appliable height	-	ROM, muscle activity	-	Discomfort	(Kong, Park et al. 2021)
	K-SRDTM	Active	Kneeling, lifting , stair-climbing tasks	Lab	Weight, Battery life	Heart rate	ROM	Task completion time, distance	RPE, usability	Jan M. Hondzinski 2018 (Hondzinski, Ikuma et al. 2018)
	HUST-EC	Passive	Squatting tasks	Lab	-	-	Assistive force, muscle activity	-	-	(Han, Du et al. 2019)
	Chairless Chair	Passive	Walking, simulated screwing and assembly tasks (customized tasks)	Lab	-	-	-	-	Safety, strain, usability, discomfort	(Groos, Fuchs et al. 2019)
	E-LEG	Semi-Active	Walking, squatting tasks	Lab	Weight	-	Joint angle, foot pressure, muscle activity	-	-	(Tu, Zhu et al. 2022)
	Chairless Chair 2.0	Passive	Assembly tasks (standing and sitting tasks)	Field	-	-	Joint angle	-	-	(Onofrejova, Balazikova et al. 2022)

Table 8. (continued) Results filtered from selected articles

Full-body	Personal device	Active	Lifting tasks	Lab	Dimension	Inhaled O2,	DOF, muscle activity	-	Usefulness, side effect, intention of use, ease of use, facilitating condition	(Li, Li et al. 2021)
	AXO-SUIT	Active	Lifting, carrying tasks	Lab	Weight, Battery life	-	DOF, ROM, muscle activity	Task completion time	RPE, Discomfort, effectiveness, satisfaction	(Bai, Islam et al. 2022)

APPENDIX 2

Table 9: (continued) A List of Commercially Available Exoskeletons

Types	Names	Actuations	Tasks	Metrics					Websites
				Design	Physiological	Biomechanical	Functional	Subjective	
Upper-body	AIRFRAME	Passive	Lifting, overhead, holding tasks	-	-	Reduced effort, ROM	-	-	https://www.levitatetech.com/airframe/
	Armorman 3.0	Passive	Holding tasks	Weight, dimensions	-	-	-	-	https://tilta.com/shop/armor-man-3-0-gimbal-support-system/
	BESK	Passive	Lifting, overhead, holding tasks	Weight	-	Assistive force	-	-	https://en.cyberhs.eu/besk
	CDYS	Passive	Lifting tasks	Weight, applicable body height, textiles breathable, washable, exchangeable	-	-	-	-	https://www.c-dyn.com/product?
	EVO	Passive	Lifting tasks	Warranty period,	-	Assistive force, donning time	Reduced breaks, increased productivity	-	https://eksobionics.com/ekso-evo/
	Exhauss	Passive	Lifting, holding tasks	Weight	-	Assistive force	-	-	http://www.exhauss.com/index.html
	Exy ONE	Passive	Lifting, overhead, holding tasks	Weight, durability	-	Reduced effort, assistive force, donning time	-	-	https://www.exy9br.com/
	LIGHT	Passive	Handling tasks	Weight, textiles exchangeable and washable, applicable body height	-	Assistive force, donning time	-	-	https://www.hmt-france.com/
	MATE-XT	Passive	Overhead tasks	Applicable body height, textiles breathable, construction material	-	ROM, donning time, muscle activity	-	-	https://mate.comau.com/

Table 9: (continued) A List of Commercially Available Exoskeletons

Types	Names	Actuations	Tasks	Metrics					Websites
				Design	Physiological	Biomechanical	Functional	Subjective	
Upper-body	Muscle Upper	Active	Handling tasks	Weight, dimension, operating temperature condition, applicable body height, sound level, warranty period	-	Assistive force	-	-	https://innophys.jp/
	Ottobock Shoulder	Passive	Overhead tasks	Applicable body height, weight, textiles exchangeable and washable	Inhaled O2, heart rate	Reduced effort, ROM, donning/doffing time, assistive force, muscle activity	Increased productivity, cost	Acceptance	https://paexo.com/?lang=en
	Plum'	Passive	Lifting, overhead, holding tasks	Weight, textiles exchangeable and washable	-	Assistive force, donning/setup time	-	-	https://www.hmt-france.com/en/ourExoskeletons/plum
	Shiva Exo	Passive	Overhead work, lifting and carrying tasks	Applicable body height	-	-	-	-	https://www.shivaexo.fr/en/
	ShoulderX	Passive	Overhead tasks	Weight, textiles breathable, applicable body height	-	Muscle activity, donning/doffing time	-	Muscle fatigue	https://www.suitx.com/shoulderx
	CarrySuit	Passive	Carrying, holding tasks	-	Heart rate	Muscle activity	Cost	-	https://www.auxivo.com/
	AGADEXO	Active	Handling tasks	Weight, battery life	-	Assistive force	-	-	https://agade-exoskeletons.com/en/agadexo-eng/
	Exorise	Passive	Lifting and holding tasks	Weight	-	Assistive force	Increased productivity	-	https://exorise.com/
Back	ALDAK	Passive	Bending, lifting tasks	Weight	-	Assistive force, donning time	-	-	https://en.cyberhs.eu/_files/ugd/64cc92_b3dd6bd227f04228b46d695025f25038.pdf
	HeroWear Apex	Passive	Lifting tasks	Weight, applicable body height, textiles breathable	-	Donning/doffing time, muscle activity	Cost	Muscle fatigue, comfort, usability	https://herowearexo.com/

Table 9: (continued) A List of Commercially Available Exoskeletons

Types	Names	Actuations	Tasks	Metrics					Websites
				Design	Physiological	Biomechanical	Functional	Subjective	
Back	AWN-12	Active	Lifting tasks	Weight, battery life, dimensions, applicable body height, operating temperature conditions	-	Assistive force	-	-	https://atoun.co.jp/en/products/atoun-model-y/
	BackX	Passive	Bending, lifting tasks	Weight	Inhaled O2, exhaled CO2	Muscle activity	Increased productivity, task completion time	Muscle fatigue, usability	https://www.suitx.com/backx
	BionicBack	Passive	Bending, holding, lifting tasks	Weight	-	Muscle activity	Task completion time	-	https://en.htrius.com/wissen
	Atlas	Passive	Carrying tasks	Weight, applicable body height	-	Muscle activity, assistive force	-	-	https://www.exomys.com/
	Cray X	Active	Lifting, walking tasks	Weight, ingress protection rating	-	Assistive force	-	-	https://www.germanbionic.com/en/cray-x-solution/#fourth-generation
	DARWING PA-Jacket	Active	Lifting tasks	Weight, applicable body height	-		-	-	http://www.daiyak.co.jp/product/detail/403
	Enyware	Passive	Sitting tasks	Weight, dimensions, construction material, warranty period, applicable body height	-	Donning time	Cost	-	https://astride.io/products/enyware-seat
	ExoAtlant Torso	Passive	Lifting, bending, carrying and holding tasks	Weight, construction material	-	Reduced effort, assistive force, setup time	-	-	https://exoatlet.lu/exoatlant/
	FLx ErgoSkeleton	Passive	Rotating, lifting tasks	Dimensions	-		Reduced risk of injury, saved money, cost	Safety	https://www.robotcenter.co.uk/products/flx-ergoskeleton
	H-WEX	Active	Overhead, lifting tasks	Weight, dimensions	-	Assistive force	Cost	-	https://www.hyundai.news/eu/articles/press-releases/hyundai-develops-wearable-vest-exoskeleton-for-overhead-work.html

Table 9: (continued) A List of Commercially Available Exoskeletons

Types	Names	Actuations	Tasks	Metrics					Websites
				Design	Physiological	Biomechanical	Functional	Subjective	
Back	HAL Lumbar Support	Active	Lifting tasks	Weight, dimensions, applicable body height, applicable body weight, battery life, battery charging time, ingress protection rating, operating temperature conditions	-	-	-	-	https://www.cyberdyne.jp/english/products/Lumbar_LaborSupport.html
	HAPO MS	Passive	Handling, holding tasks	Weight	-	Muscle activity, assistive force, ROM, doffing/setup time	-	Muscle fatigue	https://ergosante.fr/exosquelette-hapo-ms/
	Japet.W	Active	Lifting, handling tasks	Weight, battery life	Heart rate, inhaled O2	Muscle activity, ROM, joint force	-	Pain, muscle fatigue	https://www.japet.eu/exosquelette-entreprise/
	Laevo V2	Passive	Bending, lifting tasks	Weight, applicable body height	-	Donning/doffing time, ROM	-	-	https://www.laevo-exoskeletons.com/
	Laevo FLEX	Passive	Bending, lifting tasks	Weight	-	Donning/doffing time	-	-	https://www.laevo-exoskeletons.com/flex
	LiftSuit	Passive	Bending, lifting tasks	Weight	-	Muscle activity	-	-	https://www.auxivo.com/liftsuit
	Muscle Suit Every	Active	Lifting tasks	Weight, dimensions, applicable body height, operating temperature conditions, warranty period, ingress protection rating	-	Assistive force, donning time	-	-	https://innophys.jp/
	Muscle Upper	Active	Handling tasks	Weight, dimension, operating temperature condition, applicable body height, sound level, warranty period	-	Assistive force	-	-	https://innophys.jp/
	Paexo Back	Passive	Lifting tasks	Weight, textiles washable, exchangeable	-	Assistive force, donning/doffing time	Cost	-	https://paexo.com/paexo-back/?lang=en

Table 9: (continued) A List of Commercially Available Exoskeletons

Types	Names	Actuations	Tasks	Metrics					Websites
				Design	Physiological	Biomechanical	Functional	Subjective	
Back	Shiva Exo	Passive	Overhead work, lifting and carrying tasks	Applicable body height	-	-	-	-	https://www.shivaexo.fr/en/
	V22 ErgoSkeleton	Passive	Lifting, carrying tasks	Weight, dimensions	-	-	-	-	https://multimedia.3m.com/mws/media/1182159O/strongarm-v22-sell-sheet.pdf
Lower-body	Archelis	Passive	Standing tasks	Applicable body heights, applicable body weight, dimensions, weight	-	-	-	-	https://www.archelis.com/product-fx
	Chairless Chair	Passive	Standing tasks	Applicable body heights		Donning time	-	-	https://www.noonee.com/the-chairless-chair-2-0/?lang=en
	LegX	Passive	Squatting tasks	-	-	Assistive force, muscle activity	Cost	Discomfort	https://www.suitx.com/legx
Full-body	Guardian® XO®	Active	Lifting tasks	-	-	Assistive force, DOF, donning time, reduced effort	-	-	https://www.sarcos.com/products/guardian-xo-powered-exoskeleton/

APPENDIX 3

Table 10: (continued) Basic Task Library

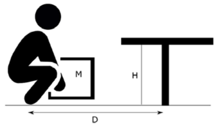
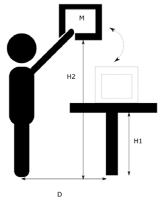
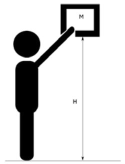
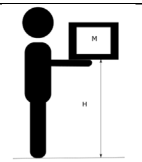
Test	Schematic sketch	Procedure
Repetitive lower lifting		The subject lifts the load (5 kg) from floor level to table level (1075 mm) for 5 minutes. The number of lifting times (N) is recorded. The recording should be stopped when the subject feels fatigue, uncomfortable.
Repetitive higher lifting		The subject lifts the load (5 kg) from table level (1075 mm) to target level (15 mm above stature) for 5 minutes. The number of lifting times (N) is recorded. The recording should be stopped when the subject feels fatigue, uncomfortable.
Overhead holding		The subject holds the load (5 kg) at target level (15 mm above stature). The time recording (T) should be stopped when the subject feels fatigue, uncomfortable.
Holding		The subject holds the load (5 kg) at chest level. The time recording (T) should be stopped when the subject feels fatigue, uncomfortable

Table 10: (continued) Basic Task Library

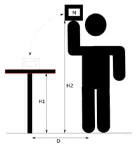
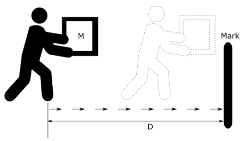
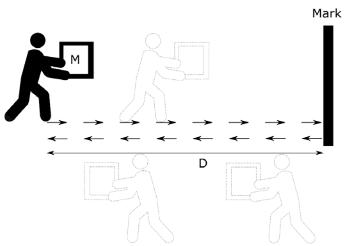
Test	Schematic sketch	Procedure
One-handed holding		The subject holds the load (5 kg) at target level (15 mm above stature) by using one hand. The time recording (T) should be stopped when the subject feels fatigue, uncomfortable
One-handed lifting		The subject lifts the load (5 kg) from table level (1075 mm) to target level (15 mm above stature) for T minutes by using one hand. The number of lifting times (N) is recorded. The recording should be stopped when the subject feels fatigue, uncomfortable.
One-way carrying		The subject carries the load (5 kg) to a target distance (5000 m). The time recording should be stopped when the subject passes the mark.
Shuttle carrying		The subject carries the load (5 kg) to a target distance (5000 mm) then goes back to the starting place. The distance that the subject walks within 5 minutes is recorded.

Table 10: (continued) Basic Task Library

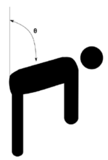
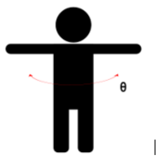
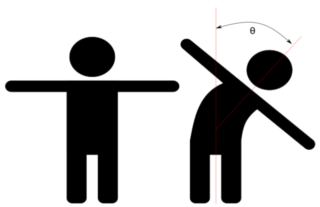
Test	Schematic sketch	Procedure
Maximum bending		The subject bends as much as possible with knees extended 5 times. The maximum angle (θ) of bending is recorded.
Bending over to perform the task		The subjects bends (with knees extended) to perform a screwing task on a table (1075 mm high from floor and 500 mm away from the subject) within 5 minutes. The number of screwing nuts (N) that the subject completes is recorded.
Rotation of trunk		The subject rotates the trunk 5 times. The maximum angle of trunk rotation is recorded.
Lateral flexion		The subject performs lateral flexion 5 times. The maximum angle of trunk flexion is recorded.

Table 10: (continued) Basic Task Library

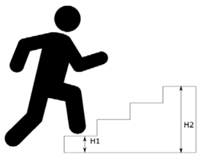
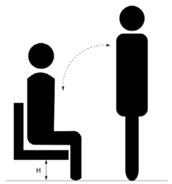
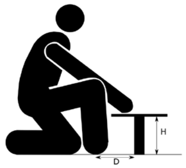
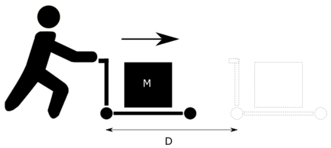
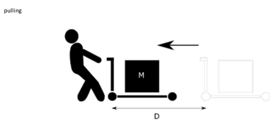
Test	Schematic sketch	Procedure
Squatting		The subject squats down to the floor as much as possible 5 times. The time of squatting is recorded.
Wide stance		The subject stands apart as much as possible 5 times. The maximum distance (D) is recorded.
Stair climbing		The subject climbs up- and downstairs (150 mm high for each stair, 750 mm high for all stairs) as fast as possible for 5 minutes. The number of stairs (N) that the subject walks is recorded.
Sit to stand		The subject sits down on a chair and gets up 5 times. The subject starts in sitting position and the time recording (T) stops when the subject sits down the 5th time.

Table 10: Basic Task Library

Test	Schematic sketch	Procedure
Single-knee kneeling		The subject performs a screwing task on a table (250 mm high from the floor and 100 mm away from the subject) for 5 minutes by using a single-knee kneeling posture. The number of screwing task (N) that the subject completes is recorded.
Pushing		The subject pushes the load (M kg) to a target distance (5000 mm). The time recording (T) should be stopped when the subject passes the mark.
Pulling		The subject pulls the load (M kg) to a target distance (5000 mm). The time recording (T) should be stopped when the subject passes the mark.
Crouching		The subject performs a screwing task on a floor level for T minutes by holding a crouching posture. The number of screwing task (N) that the subject completes is recorded.

APPENDIX 4

Table 11. Borg CR10 scale

Rating	description
0	Not at all
0.5	Very, very light
1	Very light
2	Fairly light
3	Moderate
4	Somewhat strong
5	Strong
6	
7	Very strong
8	
9	
10	Very, very strong

(1) To what extent do you think this exoskeleton is comfortable?
(2) To what extent do you think this exoskeleton is useful?
(3) To what extent do you think this exoskeleton is helpful to reduce perceived effort?
(4) To what extent do you think this exoskeleton looks aesthetically pleasing?
(5) To what extent do you think this exoskeleton is visible?
(6) To what extent do you think this exoskeleton is of good versatility?
(7) To what extent do you think this exoskeleton is of good quality?
(8) To what extent do you think this exoskeleton is of good safety?
(9) To what extent do you think this exoskeleton restrict your movement?

APPENDIX 5

Table 12. Reference Datasheet of Benchmarking Metrics

Weight	Range	<1 kg	1~2 kg	2~3 kg	3~4 kg	4~5 kg	5~6 kg	6~7 kg	7~8 kg	8~9 kg	>9 kg
	Reference score	10	9	8	7	6	5	4	3	2	1
Setup/Donning/Doffing time	Range	<20 s	20~40 s	40~60 s	60~80 s	80~100 s	100~120 s	120~140 s	140~160 s	160~180 s	>180 s
	Reference score	10	9	8	7	6	5	4	3	2	1
Cost	Range	<5 k	5~10 k	10~15 k	15~20 k	20~25 k	25~30 k	30~35 k	35~40 k	40~40 k	>40 k
	Reference score	10	9	8	7	6	5	4	3	2	1
Metabolic consumption reduction/ Heart rate reduction/ Productivity booster/ Muscle activity reduction/Joint moment reduction	Range	< -90%	-90%~-80%	-80%~-70%	-70%~-60%	-60%~-50%	-50%~-40%	-40%~-30%	-30%~-20%	-20%~-10%	-10%~-0
	Reference score	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
	Range	0~10%	10%~20%	20%~30%	30%~40%	40%~50%	50%~60%	60%~70%	70%~80%	80%~90%	> 90%
	Reference score	1	2	3	4	5	6	7	8	9	10
Maximum assistive force	Range	<20 N	20~40 N	40~60 N	60~80 N	80~100 N	100~150 N	150~200 N	200~250 N	250~300 N	>300 N
	Reference score	1	2	3	4	5	6	7	8	9	10

APPENDIX 6

Table 13. Required equipment

Metrics	Hardware	field lab							
		AAU	TNO	CITA	CSIC	U. Twente	UGAV	HAWK	KU Leuven
Weight	weighing scales	yes		yes	yes	yes	yes	yes	
Dimensions (Length × Width × Height)	meter	yes		yes	yes	yes	yes	yes	
Applicable body weight	weighing scales	yes		yes	yes	yes	yes	yes	
Applicable body height	meter	yes		yes	yes	yes	yes	yes	
Payload	weighing scales	yes		yes	yes	yes	yes	yes	
Reach	meter	yes		yes	yes	yes	yes	yes	
DOFs	-	yes		?	yes	yes	yes	?	
Setup/Donning/Doffing time	stop watch	yes		yes	yes	yes	yes	yes	
Transparency	-	yes		?	tbc	yes	-	-	
cost	-	yes		yes	yes	yes	yes	-	
construction material	-	yes		yes	yes	yes	yes	-	
Actuation and power									
Actuation type	-	yes		yes	yes	yes	yes	no	
Power source type	-	yes		yes	yes	yes	yes	no	
Operating voltage	voltmeter	yes		yes	yes	yes	yes	no	
Operating current	amperemeter	yes		yes	tbc	yes	-	no	
Battery run time	timer	yes		yes	yes	yes	yes	no	
Battery charging time	timer	yes		yes	tbc	yes	yes	no	
Operating conditions									
Ingress Protection Rating	-			?	tbc			?	
Operating temperature conditions	thermometer	yes		yes	yes	yes	yes	yes	
Operating humidity conditions (relative air humidity)	hygrometer	yes		yes	yes	yes	yes	yes	
Operating air pressure conditions	barometer	yes		yes	yes	yes	yes	yes	
Sound level	Sound-level meter	yes		yes	yes	yes	yes	no	
Interface									
Attachment points	-	yes		yes	yes	yes	yes	-	
Padding exchangeable?	-	yes		yes	yes	yes	yes	-	
Padding breathable?	-	yes		yes	yes	yes	yes	-	
Padding washable?	-	yes		yes	yes	yes	yes	-	
Operator interface	-	yes		yes	yes	yes	yes	yes	
Additional apparatus required?	-	yes		yes	yes	yes	-	-	
Metabolic consumption	gas-analysis systems	no		no	no	yes		yes	yes
Relative heart rate	pulse Oximeter	yes		no	no	yes		yes / not a pulse Oximeter	yes
Muscle activity	EMG device	yes	yes	no	yes	yes		yes	yes
Duration of stance, single/double support and swing phase	motion capture devic	yes	yes	yes	yes	yes		-	yes
Walking speed	motion capture devic	yes	yes	yes	yes	yes		-	yes
Walking stride length	ion capture device, m	yes	yes	yes	yes	yes		-	yes
Step width	ion capture device, m	yes	yes	yes	yes	yes		-	yes
force	plate, motion capture	yes	yes	Mocap	yes	yes		yes	yes
Lifting speed	counter, stopwatch	yes	yes	yes	yes	yes		yes	yes
Productivity (rate) booster	counter, stopwatch	yes	yes	yes	yes	yes		yes	